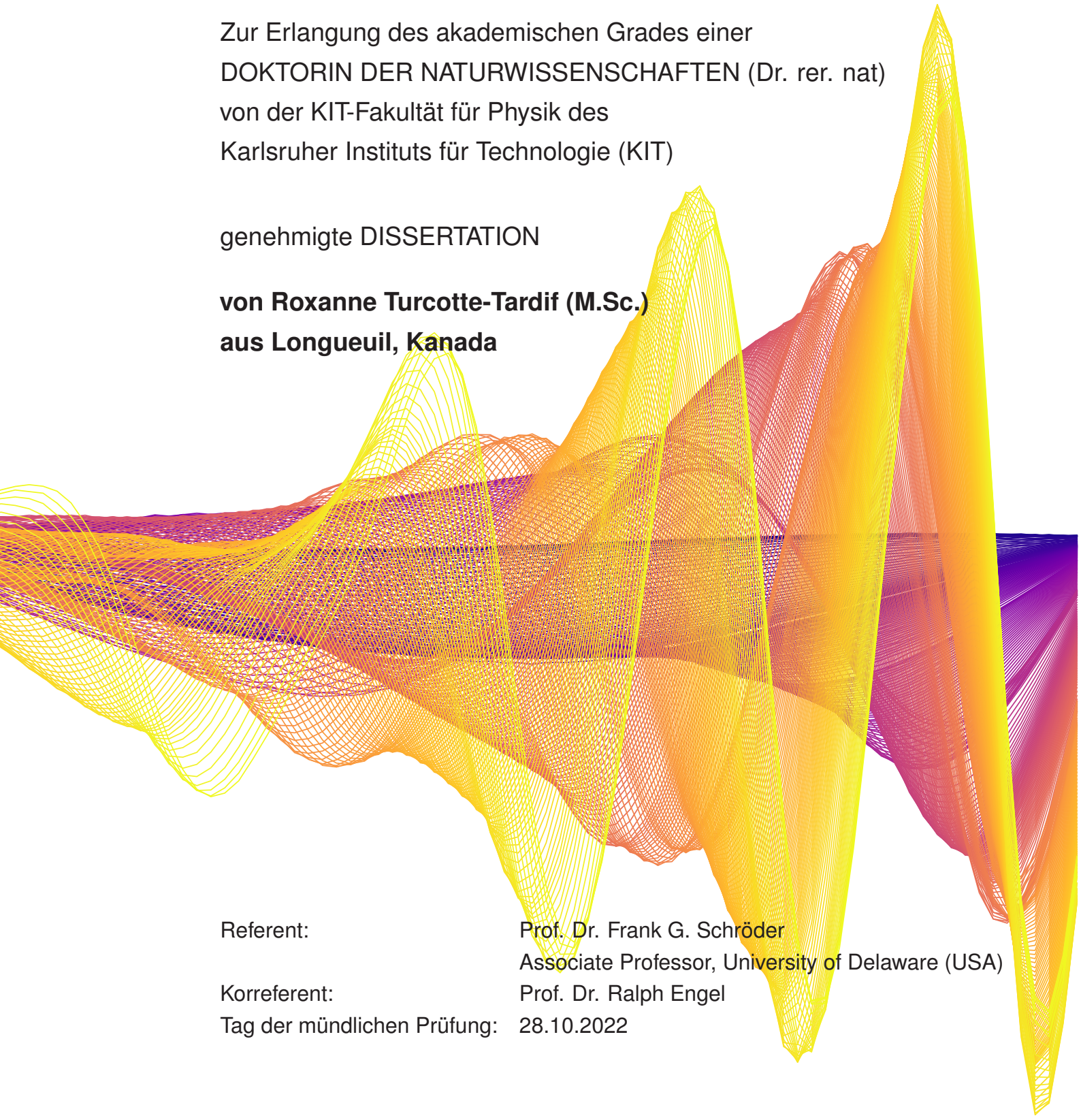
Radio Measurements of Cosmic Rays at the South Pole

Zur Erlangung des akademischen Grades einer
DOKTORIN DER NATURWISSENSCHAFTEN (Dr. rer. nat)
von der KIT-Fakultät für Physik des
Karlsruher Instituts für Technologie (KIT)

genehmigte DISSERTATION

von Roxanne Turcotte-Tardif (M.Sc.)
aus Longueuil, Kanada



Referent:	Prof. Dr. Frank G. Schröder Associate Professor, University of Delaware (USA)
Korreferent:	Prof. Dr. Ralph Engel
Tag der mündlichen Prüfung:	28.10.2022

Abstract

Ultra-high energy cosmic rays can be measured on the ground by arrays of detectors via the massive cascades of particles (air showers) they instigate in the Earth's atmosphere. Amongst the diverse detectors utilized, radio antennas became more popular in the last decade as these offer a unique opportunity to probe these air showers. The radio emission, which is created during the air shower development, mainly by the deflection of the electrons and positrons in the cascade by the Earth's magnetic field, encompasses information about the type of the particle that initiated the shower. Indeed, radio antennas, alongside fluorescence telescopes, are able to reconstruct the atmospheric depth of the maximum development of the air shower (X_{max}). This observable is sensitive to the type of atomic nucleus of the cosmic ray which initiated the air showers. Knowledge about the cosmic ray type, in turn, increases the understanding of the powerful acceleration processes of astrophysical sources in our universe.

The IceCube Neutrino Observatory, located at the geographic South Pole, is a multi-purpose detector, able to detect astrophysical neutrinos as well as air showers, especially with its complementary surface detector, IceTop. In order to improve IceTop as a cosmic ray detector and mitigate the effect of snow accumulation, a hybrid array composed of elevated scintillation panels and radio antennas is planned to be deployed in the coming years. This array will be composed of 32 stations, each comprising 8 scintillation panels and 3 antennas, covering an area of 1 km^2. The radio array uses a higher frequency band than the other radio experiments, nominally from 70 to 350 MHz instead of 30 to 80 MHz. The first complete prototype of a hybrid station was deployed in January 2020. In this work, the hardware of the prototype station and of the prospective planned stations, the commissioning of the data from the prototype station, and an energy and X_{max} reconstruction method developed based on measured events and Monte Carlo simulation, will be presented.

Specifically, a structure for raising the antennas over the snow was designed, built, field-tested, and produced, alongside with a radio front-end board for the pre-preprocessing of the analog signal received from the antennas. The calibration of the other radio signal components at various temperatures reaches an uncertainty on the amplitude of only 3.9%, significantly lower than the 10% uncertainty requirement on the radio signal chain. The functioning of the array was confirmed by analysing the radio background at the site, utilizing the developed radio data processing pipeline. Ultimately, a total of 121 air showers were detected, 5 of which are shown to also have an in-ice counterpart. Finally, 16 air showers are used to develop the first energy and X_{max} reconstruction method for the radio component of the surface array enhancement.

This reconstruction method is based on the state-of-the-art technique for radio arrays. A complete study concerning the influence of the background radio noise on the signal is performed. Afterwards, the commonly used χ^2 minimization method is converted to a log-likelihood minimisation with a complete parametrisation of the noise. The technique is shown to work with the measured data. Furthermore, for the same reconstructed events, the high frequency band gives rise to a significantly better reconstruction accuracy than the traditional low frequency band, even with only the three antennas of the prototype station. Once the planned array is completed the expected reconstruction precision is estimated at 15 $\mathrm{g/cm}^2$ for X_{max} and less than 10% in energy.

Zusammenfassung

Die ultrahochenergetische kosmische Strahlung, die in der Erdatmosphäre massive Teilchenkaskaden (ausgedehnt Luftschauer) auslöst, kann am Erdboden mit Hilfe von Detektorfeldern gemessen werden. Unter den verschiedenen Detektoren, die zum Einsatz kommen, haben Radioantennen im letzten Jahrzehnt an Bedeutung gewonnen, da sie eine einzigartige Möglichkeit bieten diese Luftschauer zu untersuchen. Die Radioemission, die während der Entwicklung des Luftschauers hauptsächlich durch die Ablenkung der Elektronen und Positronen in der Teilchenkaskade durch das Erdmagnetfeld entsteht, enthält Informationen über die Art der Teilchen, die den Schauer ausgelöst haben. Insbesondere können Radioantennen zusammen mit Fluoreszenzteleskopen die Position des Maximums der Entwicklung des Luftschauers (X_{max}) rekonstruieren. Dieser rekonstruierte Parameter ist abhängig von der Art des primären Atomkerns der kosmischen Strahlung, die den Luftschauer ausgelöst hat. Die Kenntnis des Typs der kosmischen Strahlung wiederum trägt zu einem besseren Verständnis der Beschleunigungsprozesse astrophysikalischer Quellen in unserem Universum bei.

Das IceCube Neutrino Observatorium am geografischen Südpol ist ein Mehrzweckdetektor, der sowohl astrophysikalische Neutrinos, als auch Luftschauer nachweisen kann, insbesondere mit seinem Oberflächendetektor, IceTop. Um IceTop als Detektor für kosmische Strahlung zu verbessern und die Auswirkungen der Schneeansammlung abzuschwächen, soll in den kommenden Jahren ein hybrider Dektector aus anhebbaren Szintillationsplatten und Radioantennen installiert werden. Dieser Sub-Detektor wird aus 32 Stationen bestehen, die jeweils 8 Szintillationspaneele und 3 Antennen umfassen und eine Fläche von 1 km^2 abdecken. Die Radioantennen nutzen mit 70 bis 350 MHz statt 30 bis 80 MHz ein höheres Frequenzband als bisher üblich. Der erste vollständige Prototyp einer Hybridstation wurde im Januar 2020 in Betrieb genommen. Diese Arbeit behandelt die Hardware der Prototyp-Station und der zukünftigen geplanten Stationen, die Inbetriebnahme der Daten der Prototyp-Station sowie eine Methode zur Energie- und X_{max}-Rekonstruktion, die auf der Grundlage gemessener Ereignisse und Monte-Carlo-Simulationen entwickelt wurde.

Insbesondere wurde eine Struktur zum Anheben der Antennen über dem Schnee entworfen, gebaut, im Feld getestet und produziert, zusammen mit einer Radio-Frontend-Platine für die analoge Vorverarbeitung des von den Antennen empfangenen Signals. Die Kalibrierung der anderen Radiosignalkomponenten bei verschiedenen Temperaturen erreicht eine Amplitudenunsicherheit von nur 3,9%, was deutlich unter der geforderten Unsicherheit von 10% für die Radio-Signalkette liegt. Die Funktionsweise der Detektoren wurde durch die Analyse des Radio-Untergrunds unter Verwendung der entwickelten Radio-Datenanalysekette bestätigt. Es wurden insgesamt 121 Luftschauer nachgewiesen, von denen 5 auch durch die anderen Detektoren nachgewiesen wurden. Sechszehn Luftschauer wurden verwendet, um die erste Energie- und X_{max}-Rekonstruktionsmethode für die Radiokomponente der Detektorerweiterung zu entwickeln.

Diese Rekonstruktionsmethode basiert auf dem neuesten Stand der Technik für Radio-Detektoren. Es wurde eine Analyse des Einflusses des Radio-Untergrundes auf das Signal durchgeführt. Anschließend wird die üblicherweise verwendete Methode der χ^2 -Minimierung durch eine Log-Likelihood-Minimierung mit einer Parametrisierung des Rauschens ersetzt, und es wird gezeigt, dass diese Technik mit den gemessenen Daten funktioniert. Darüber hinaus zeigt sich, dass bei denselben rekonstruierten Ereignissen das Hochfrequenzband mit den nur drei Antennen der Prototypstation eine deutlich bessere Genauigkeit als das traditionelle Niedrigfrequenzband aufweist. Sobald der gesamte Detektor fertiggestellt ist, wird die erwartete Rekonstruktionsgenauigkeit auf 15 $\mathrm{g/cm}^2$ für X_{max} und besser als 10% für die Energie geschätzt.

Contents

CHAPTER 1

Introduction

Cosmic rays have been of interest since Victor Hess confirmed an increase in the flux of ionizing radiation at high altitudes. Since then, techniques for detecting cosmic rays have continuously improved, from early instruments such as cloud chambers and particle counters to more advanced detectors such as fluorescence telescopes and radio antennas. These advancements have greatly increased our understanding of extraterrestrial particles. Cosmic rays have played a key role in discovering new particles and are promising candidates for understanding the most energetic sources and processes in our universe.

The IceCube Observatory located at the South Pole is a one-of-its-kind detector, designed to measure the first very energetic astrophysical neutrinos. In addition, the combination of its surface array, IceTop, and its in-ice volume, makes it an excellent cosmic-ray detector. However, snow accumulated over the years at a rate of roughly 20 cm per year, slowly covering the ice-Čerenkov tanks, leading to a rise in the uncertainties as well as in the detection threshold of the surface detectors. To mitigate the snow effect, an array of scintillation panels was proposed. Additionally, a complementary array made of antennas was incorporated into the design. This radio array will improve the measurement accuracy of the depth of the shower maximum, X_{max} , an observable widely used for distinguishing the type of particle that entered our atmosphere and initiated a detectable air shower. The ability to distinguish the type of the primary particle, i.e., the mass composition of cosmic rays, is a crucial tool in modern cosmic-ray detectors, as it enables a better understanding of the sources and processes that created such energetic particles. A review of the current status of the field constitutes the first part of this work.

To achieve this goal, the surface array enhancement is composed of hybrid stations, each of which consists of a central data acquisition module, eight scintillation panels, and three antennas. The deployment of the enhancement is planned for the next few years. Once the deployment is completed, the array will have a total of 32 stations covering an area 1 km^2, equivalent to that of IceTop.

For the radio array, a structure must be developed that can keep the receiving antennas above the snow throughout the year and allow for raising the antennas over time. The structure also needs to withstand harsh weather and be easy to set up at the site, given the limited personnel available. Not only the mechanical hardware has strict requirements, but also the electronics of the radio signal chain. It is paramount that the total uncertainty on the amplitude of the radio signal is less than 10% in order to be able to reconstruct the energy of the primary particle with that accuracy. The design, tests, and calibration of the mechanical and electrical radio hardware compose the second part of this thesis. In particular, a radio front-end board, for the analog pre-processing of the radio signal is designed, which filters the radio data to the desired frequency band, amplifies the signal as required, and splits the signal for the digital recording. The calibration of the principal components in the radio signal chain is presented along with total uncertainty in the system. Both systems, the antenna structure, and the front-end board, underwent two iterations in design. Their last version is currently the one planned for the larger scale deployment.

A complete hybrid station underwent a testing phase at the site after the initial design was concluded.

This started in January 2020, when a complete prototype station with the central data acquisition, the eight scintillation panels, and three antennas was installed at the South Pole. This prototype station is running continuously since then. The commissioning of the radio component of this prototype station and the basic checks on the received data to ensure proper operation constitutes the third part of this work. In addition to these checks, the first air showers were detected with the hybrid station, demonstrating the possibility of coincident measurements between all the detector types of the IceCube Observatory.

Finally, the last part of the thesis involves the development of a reconstruction technique for X_{max} and energy using both simulations and data. A reconstruction method based on the state-of-the-art technique in the field of cosmic-ray physics using radio arrays is developed, implemented, and tested with data from the prototype station. The reconstruction method expands the current state-of-the-art technique by replacing the minimization method from χ^2 to a more general log-likelihood (LLH). The implementation is based on an extensive study of how noise affects the measured radio quantities and is outlined in Chapter 9. This method is the first reconstruction method implemented for the future surface array enhancement and uses the developed tools in RadCube, a module from the IceCube standard software.

Along with the thesis work, the production and calibration of everything related to the radio component of the surface array enhancement, including the antenna structure, the radio front-end board, and the data acquisition system, was completed. This culminated in the production and calibration of seven complete stations in 2021, which are now ready for deployment at the IceCube Neutrino Observatory at the South Pole.

1.1 Outline

The chapters of this thesis are organized as follows:

Chapter 1 introduces the work.

Chapter 2 is a theoretical overview of multi-messengers physics with a focus on cosmic rays.

Chapter 3 discusses extensive air showers and their detection.

Chapter 4 focuses on the IceCube Neutrino Observatory and the future surface array enhancement.

Chapter 5 reviews the SKALA-v2 antennas used and the mounting structure for the deployment.

Chapter 6 describes the electronics required for the prototype station and the future array.

Chapter 7 elaborates on the calibration of the radio electronics.

Chapter 8 presents the commissioning of the radio data of the prototype station at the South Pole.

Chapter 9 details the study of noise relevant to radio measurements.

Chapter 10 explains the developed method for energy and X_{max} reconstruction and the results.

Chapter 11 concludes this work and gives an outlook into possible further development.

CHAPTER 2

Multimessenger Astronomy

Electromagnetic radiation, cosmic rays, neutrinos, and gravitational waves are the four messengers usually included in the emerging field of multimessenger astronomy. This new field appeared from the desire to probe the most extreme phenomena in our universe by all possible means, hence combining the characteristics and advantages of different messengers. Indeed, synthesizing the characteristics of these messengers will improve the understanding of the astronomical sources and of the acceleration processes producing particles with energy up to 10^{20} eV.

The last decade has seen new inter-collaborations analyses combining observations of different messengers. Just to pick two famous examples of these inter-collaboration efforts, which involve the IceCube Neutrino Observatory, are the TXS0506+056 blazar in 2017 [1] and more recently, in 2020, the tidal disruption event (TDE) AT2019dsg [2]. The first example is a $\sim 3\sigma$ significance of a high-energy neutrino event of $\sim$290 TeV in spatial and temporal coincidence with a flaring object reported by the Fermi Large Area Telescope Collaboration. The reconstructed neutrino direction was coincident with a cataloged blazar TXS0506+056 which was in a flaring state at the moment of the high-energy neutrino measurement. The MAGIC γ-rays telescope later followed up on the alert and detected periods where the γ-ray flux from the blazar reached energy up to 400 GeV [1]. After this discovery, the IceCube Collaboration analyzed its archives and found an excess of neutrinos, around the year 2015, of 13 ± 5 events in the direction of the blazar prior to the flaring with a significance of approximately 3.5σ [3]. The second example is the measurement of a spatial and temporal coincident high-energy neutrino by IceCube during a TDE which was recorded by the Zwicky Transient Facility. The probability of finding an unrelated high-energy neutrino coincident with a radio-emitting TDE is only 0.5% [2]. This first coincident detection of a TDE and a very high energy (VHE) neutrino opened the door to investigate TDEs as an important source of potential high-energy neutrino emission.

Each of the messengers provides unique insights into the universe and has its own advantages and disadvantages. Therefore, they are all necessary pieces of the astronomical puzzle. This chapter provides a brief overview of the four messengers with a deeper focus on cosmic rays.

2.1 The complementarity of the messengers

In order to make further progress, particularly in the field of cosmic rays, it will be necessary to apply all our resources and apparatus simultaneously and side-by-side; an effort which has not yet been made, or at least, only to a limited extent.
– Victor Francis Hess

The four messengers are closely related and complement each other. An interesting connection is the creation of gamma rays and neutrinos from the interaction of cosmic rays with the cosmic microwave background (CMB), as discussed later in this chapter. Similar creation processes in the vicinity of cosmic accelerators predict fluxes of gamma rays and neutrinos, which are highly energetic and can

point to the source. Neutrinos, in particular, are very promising messengers as they are not affected by magnetic fields and can travel large distances. However, their low cross-section makes them challenging to detect, requiring extensive and sensitive detectors. Gamma rays, on the other hand, are rapidly absorbed by the CMB. Cosmic rays are expected to be the first products of these accelerators, but they are deflected by the galactic and extra-galactic magnetic fields. Finally, the new field of gravitational waves complements the field by adding heavy sources to the catalogs and enhancing the understanding of extremely energetic processes that could potentially create ultra high energy (UHE) particles.

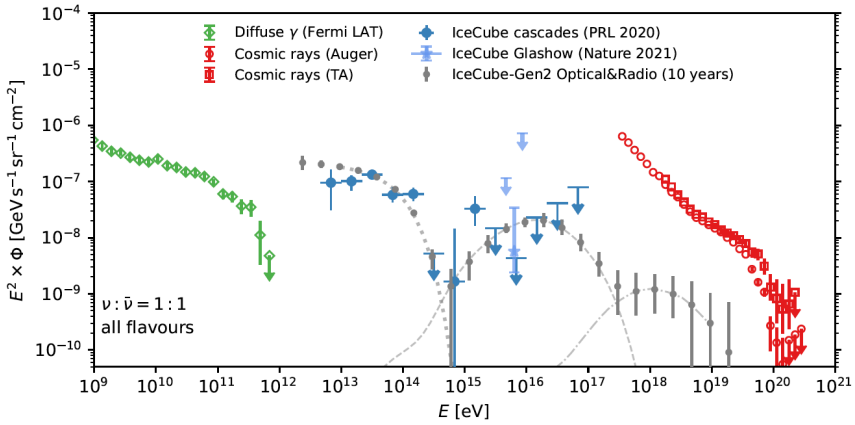


Figure 2.1: Relation between the energy spectra of γ-rays, astrophysical neutrinos and UHECR. Taken from [4]

Figure 2.1 shows quantitatively how the fluxes of these messengers are closely related to one another. In fact, the energy density of GeV γ-rays, TeV neutrinos, and EeV cosmic rays are of the same order of magnitude, accompanied by a suppression in γ-rays similar to the one in cosmic ray physics.

This plot will become clearer as the chapter progress. Nevertheless, with all these messengers adding tools into the experimenter box, this is the dawn of a new era in understanding our universe.

2.2 Cosmic rays

At two-tenths the speed of light, dust, and atoms might not do significant damage even in a voyage of 40 years, but the faster you go, the worse it is–space begins to become abrasive. When you begin to approach the speed of light, hydrogen atoms become cosmic-ray particles, and they will fry the crew. ...So 60,000 kilometers per second may be the practical speed limit for space travel.

– Isaac Asimov

Chapter 3 describes in more detail the principle of extensive air shower, radio emission as well as detection techniques. Here, the focus is on the recent results in cosmic ray astrophysics and on the open questions in the field.

Contrarily to what its name suggests, a cosmic ray is not a ray but is defined as a charged particle or nucleus traveling through the universe at relativistic speed. Cosmic rays can be either accelerated in/or around an astronomical source or be a product of interactions in space. Due to the extensive energy losses of electrons and positrons compared to hadrons, the former have been measured on Earth only up to $\sim 20\,\mathrm{TeV}$ [5]. From this point forward, the term "cosmic ray" will refer to protons and heavier nuclei. To understand cosmic rays, one can look at the energy spectrum, displaying the incoming flux at Earth as a function of the primary energy of the cosmic ray. The first observation from the spectrum, shown in Figure 2.2, is that the cosmic ray energies are vast, spanning several orders of magnitude and steeply ending at an energy of around a few $10^{20}\,\mathrm{eV}$. The high-energy end of the spectrum comprises particles with momentum far greater than what is achievable by human-made particles accelerator like

the large hadron collider in Geneva. Another important observation from the energy spectrum is the rapid decrease of the flux with energy. Note that differential energy spectra, as shown here, are often misleading because the displayed flux is usually scaled with the energy. To convey a feeling of how the exponential decrease affects the measurements, here are three integral fluxes, i.e. the number of particles above a certain energy detected in a given area and in a given time [6]:

- $\mathcal{F}(E > 10^9 \text{eV}) \approx 1000$ particles s^{-1} m^{-2}
- $\mathcal{F}(E > 10^{15} \text{eV}) \approx 1$ particle year^{-1} m^{-2}
- $\mathcal{F}(E > 10^{20} \text{eV}) \approx 1$ particles century^{-1} km^{-2}

Below 10–100 TeV, detectors located in satellites can measure cosmic rays directly, reaching high precision on the chemical composition of the cosmic rays and on their energies. Above 10–100 TeV, experiments need to rely on indirect detection with arrays of detectors on the ground. This type of detection is detailed in Chapter 3. With indirect detection, however, the reconstruction of energy and mass becomes more complex.

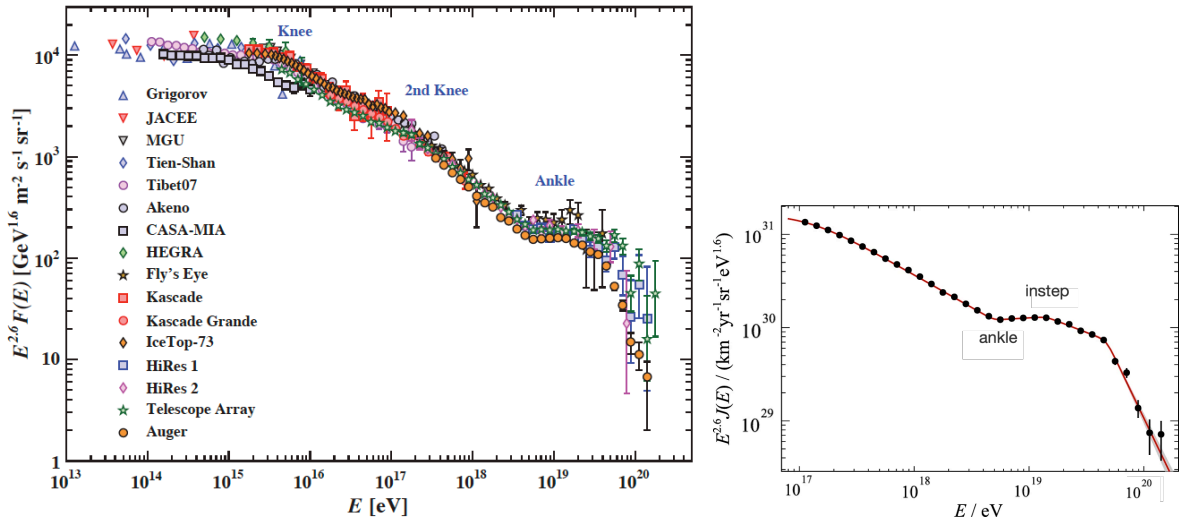


Figure 2.2: On the left, the cosmic ray energy spectrum spanning several orders of magnitude where three main features are noted, the knee, the second knee and the ankle. [7]. On the right, the new "instep" reported by the Auger Collaboration with a change of spectral index from $\gamma \sim 2.51$ to $\gamma \sim 3.05$ at 1.3×10^{19} eV . Slightly modified from [8]

The rapid decrease in flux follows a power-law function with a spectral index γ between ~2.7 and ~3.0 depending on the region of the spectrum. Those regions are defined between these four following features: the knee (~3 PeV), the second knee (~200 PeV), the ankle (~5 EeV), and the suppression (~50 EeV). Recently, after years of gathering data, the Auger collaboration published the discovery of a new feature, the so-called *instep*, at ~13 EeV shown on the right plot of Figure 2.2 [8]–[10].

Another important consideration to keep in mind when studying cosmic ray (CR) is that below an energy of $\sim 10^{18}$ eV, they are substantially affected by the galactic magnetic fields, experiencing the so-called galactic confinement. The particles stay trapped and are diffused inside an entrapment region as long as their Larmor radius (r_L) remains smaller than the size of the confinement region. They thus remained confined until they reached energies large enough energies to escape, which can be calculated with $E[\text{PeV}] = r_\text{L}\,[\text{pc}]\,Z\,B[\mu\text{G}]/1.08$. This is called the Hillas criterion [11]. The acceleration processes in the galaxy are mostly agreed upon being principally via shock acceleration in supernovae remnants and second-order Fermi acceleration, at least up to PeV energies [12]. The origin of the most energetic Galactic CR, reaching energies up to 10^{17}–10^{18} eV, remains, however, unclear. In our Galaxy, with the Hillas criterion, cosmic rays stay on average $\sim 10^7$ years [6]. Above 10^{18} eV, the cosmic rays can escape

their entrapment. Naturally, these are statistical processes and thus more energetic comic ray from the galaxy can reach the Earth, and vice-versa, lower energy cosmic ray can escape their entrapment. As a general rule, this confinement renders the identification of sources in the galactic region accelerating cosmic rays at energies below 10^{18} eV very challenging.

Nevertheless, anisotropy studies of low-energy cosmic rays can help, amongst other methods, to increase the understanding of the galactic magnetic fields. And as the knowledge about magnetic fields increases, the understanding of the trajectories of cosmic rays improves as well. For example, a TeV to PeV energy anisotropy study, as performed by IceCube, with more statistics could also probe theoretical models of propagation, such as the hypothetical superposition of fluxes from a few nearby sources [13]. A confirmation inferred from anisotropy studies of higher energies is the disfavor of CR sources above 10^{18},eV from the galaxy. If such sources exist, there would be a surplus of energetic CRs in the galactic plane, due to the fact that they would not be diffused by the galactic field.

Derived from the above-mentioned explanation, it is generally accepted by the community that the CRs from above the ankle are extra-galactic and that the transition between the galactic to the extra-galactic regime of cosmic rays production happens somewhere between the second knee and the ankle.

An interesting hypothesis for the end of the galactic cosmic ray spectrum is that it could be explained by a Peters cycle, which describes the acceleration and propagation in a magnetic field [14]. The maximal energy reachable is directly related to the atomic number of the cosmic ray via the rigidity R where $R = pc/Ze$. In this scenario, the proton would be the first to experience a cut-off followed by higher atomic nuclei via the relation:

$$E_{\max}(Z) = Ze \times R_c = Z \times E_{\max}(Z = 1) \tag{2.1}$$

Therefore, heavier elements can reach higher energies, assuming that the first knee is composed dominantly of a proton distribution steeply falling at 3,PeV, and that the second knee is composed of a smaller, yet more energetic, iron distribution falling at 80,PeV. This leads to a tempting calculation that $E^{\mathrm{Fe}}_{\max} = 26 * E^{\mathrm{p}}_{\max} \approx 80$,PeV, which is consistent with experimental observations by KASCADE-Grande [15]. However, it is still unclear at the moment whether a more energetic Galactic population exists, which could extend up to a few EeV in energy.

At the highest energies, Greisen, Zatsepin, and Kuzmin already predicted a cut-off in 1966 [16], [17]. The principle is simple, the UHE protons would interact with the CMB photons, via

$$\gamma_{CMB} + p \longrightarrow \Delta^+ \longrightarrow p + \pi^0 \tag{2.2}$$

$$\gamma_{CMB} + p \longrightarrow \Delta^+ \longrightarrow n + \pi^+ \tag{2.3}$$

Where the neutral pions further decay into two γ and the charged pions into muons and neutrinos, hence creating the connection between cosmic ray, γ-ray, and neutrinos. The first calculations predicted a cut-off for the proton starting at around 10^{20} eV and peaking at 2.5×10^{20} eV because of the Δ^+ resonance, assuming that the peak of the CMB blackbody distribution is at $T = 3.0K$ with a rapid increase for the proton cross-section at the approach of the Δ^+ resonance. These values are very close to the suppression seen in the experimental data, as shown in Figure 2.2. Furthermore, these photo-nuclear reactions with heavier nuclei also lead to similar cut-offs. It is also interesting to note that with this hypothesis the steepness of the energy spectrum suppression would relate directly to the steepness of the tail of the blackbody radiation, the distribution dictating the behavior of relic photons.

Still, although experimental data has a suppression in the expected energy range, it cannot be excluded so far that the end of the spectrum could be caused at least partially by the limitation of the acceleration mechanisms of cosmic sources via the Peters cycle explained previously [18].

Figure 2.3 illustrates theoretical distributions of four elements: proton, helium, nitrogen, and iron. These distributions are weighted to fit the spectrum. The shape of the distribution is described by the Peters cycle on the left, and by the GZK cut-off on the right. The two different models can explain, within the uncertainties, the experimental data from the Pierre Auger Observatory. They will be distinguishable by better measurements of the mass composition as planned with the AugerPrime upgrade[19].

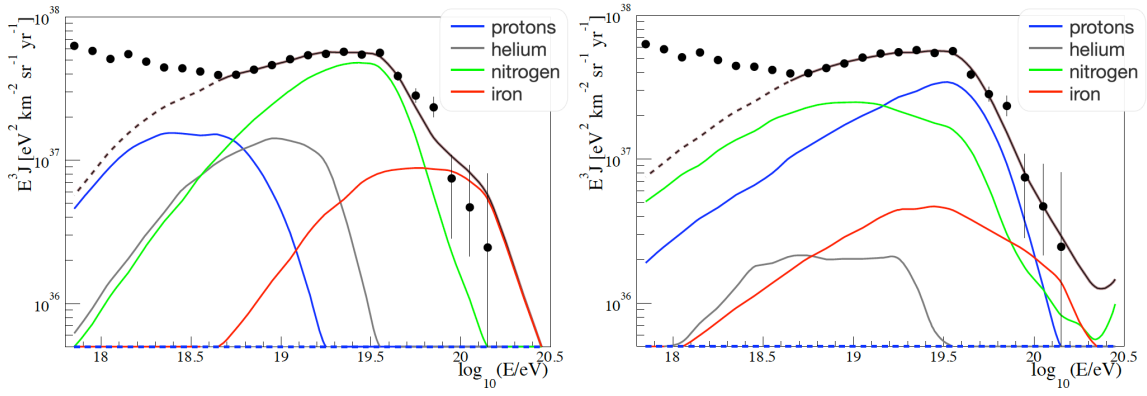


Figure 2.3: Two scenarios favored by Auger for describing the end of the spectrum. The left one illustrates a Peters cycle (maximum rigidity) and the right one a GZK cut-off (photo-disintegration); protons in blue, helium in gray, nitrogen in green, and iron in red. Slightly modified from [19].

2.2.1 Open questions in UHECR

The exact definition of ultra high energy cosmic rays (UHECR) changes from author to author e.g. [4], [20]. Here, we will define UHECR as a cosmic ray whose primary energy is above 10^{15} eV and extend up to the end of the spectrum. This range is of interest here because it contains the expected transition from galactic to extra-galactic origin of cosmic rays and also includes the range where the radio detection technique is relevant.

Energy spectrum and mass composition

As explained above, the UHECR spectrum is still not completely understood and subject to debates, especially as to where the galactic to extra-galactic transition happens and what causes the end of the spectrum. Improvements in the measurement of the mass composition and source identification will help clarify these questions.

A recent summary of the mass composition of cosmic rays results is shown in Figure 2.4, which tends to point towards a Peters cycle scenario for galactic cosmic rays as described before. Around the knee, the distribution leans to a lighter composition, whereas around the second knee the distribution becomes heavier. However, with the accuracy of the measurements, at the moment, most of the distributions overlap in their uncertainties, especially for middle elements such as helium and oxygen in this plot. Also, from a single Peters cycle, one would expect a clear composition going from dominated by proton, to dominated by helium and so on until iron, and that does not seem to be the case. Outside that shift from proton and iron distributions, most of all the elements overlap in the region between the knee and the second knee.

Further down the spectrum towards the highest energies, protons dominate again around the ankle supporting the extra-galactic sources hypothesis, or at least a second dominant population taking over at higher energy. It is interesting to observe at and after the ankle, a rotation of the dominant element that would arise from a Peters cycle tends to be present. First, the proton population dominates, followed by the helium nuclei, ended by the oxygen nuclei. Only the iron nuclei do not follow this trend, but this might be due to systematic uncertainties or to a flux simply too faint to be observed.

Another noteworthy observation on the energy spectrum is the discrepancy between the Pierre Auger Collaboration and the Telescope Array (TA) Collaboration at the extreme end of the spectrum, also depicted in Figure 2.4. While TA observes a longer flattening at the highest energy before experiencing a suppression, Auger sees an earlier suppression. The cause of this difference in the spectra is not well understood. Some propose that different sources between the northern and the southern hemisphere

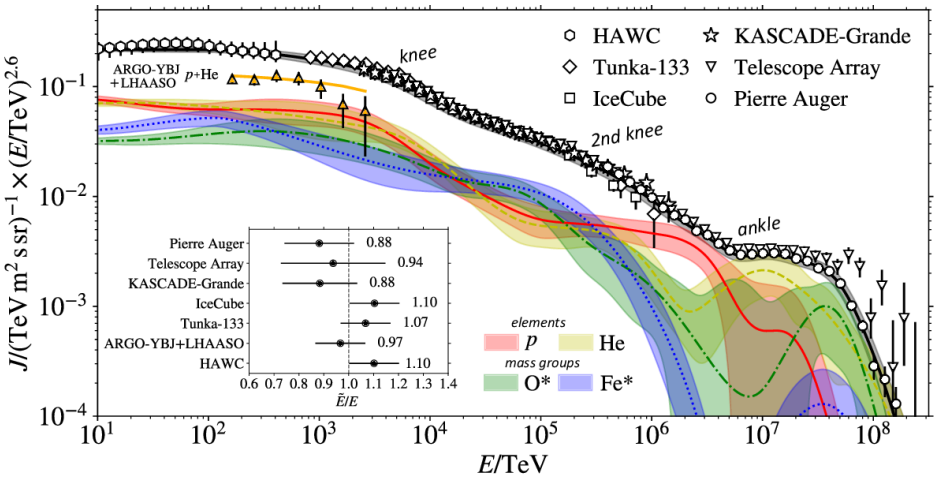


Figure 2.4: Cosmic-ray energy spectrum and mass composition combining the results from several experiments and adjusting the energy scale of all the experiments to fit the main features together. Taken from [21].

could cause this behavior, others look into new physics, and some point towards potential experimental flaws.

Origin of UHECR

Cosmic rays of energies above tens of EeV fuel special interests as messengers because they are only slightly affected by magnetic fields and do not experience galactic confinement. The study of anisotropy in the arrival direction of these cosmic rays are important to find their origin and to understand magnetic fields. Recently a dipole, most dominant at energy above 8 EeV, was confirmed by the Auger collaboration with more than 6σ significance [22]. Else, no other multipoles have yet been found in the Auger data that deviates from the isotropic hypothesis. If experiments reach an event-per-event mass composition, which is foreseen with the combination of detecting methods, as in the AugerPrime upgrade for example, it would then be possible to fix the atomic number Z. Thus, rendering the rigidity formula for the propagation of cosmic rays only dependant on the inverse of the energy.

Many types of sources are considered as possible candidates for UHECR acceleration. For a source to qualify, two criteria have to be met: the Hillas criterion [11] and the energy budget constraint. The Hillas criterion, similarly as explained before, states that a source must have a magnetic field and a size big enough that particle can be confined long enough to reach a certain energy. Hence, the maximum energy of the cosmic ray reachable by the acceleration power of the source is directly proportional to the size of the source R and the strength of the magnetic field B, i.e., $E_{max} \propto BR$. The energy budget refers to the total energy the source class needs in order to fully explain the diffuse UHECR flux observed. This energy budget is inferred from the infrared, radio, X-ray, and γ-ray observations. This is a soft limit because the exact relation between the visible spectrum luminosity and the CR luminosity is not well established, and also because the diffuse spectrum could be caused by a compilation of different sources.

Figure 2.5 illustrates both constraints. The Hillas condition on the left shows the magnetic field as a function of the size of the source (R = comoving size $\cdot\,\Gamma$). The lines define the minimal size and magnetic field required to accelerate a proton (red) or iron nuclei (blue) to a maximal energy of 10^{20} eV. The energy budget plot is on the right. It illustrates the relation between the density and the luminosity of the sources required to account for the total diffuse flux of UHECR. The lines show the minimal requirements assuming different scenarios of equivalence between electromagnetic spectrum luminosity and cosmic ray luminosity. Sources that lie below the lines in at least one of the plots do not meet the necessary requirements for explaining completely the cosmic ray flux at the highest energies.

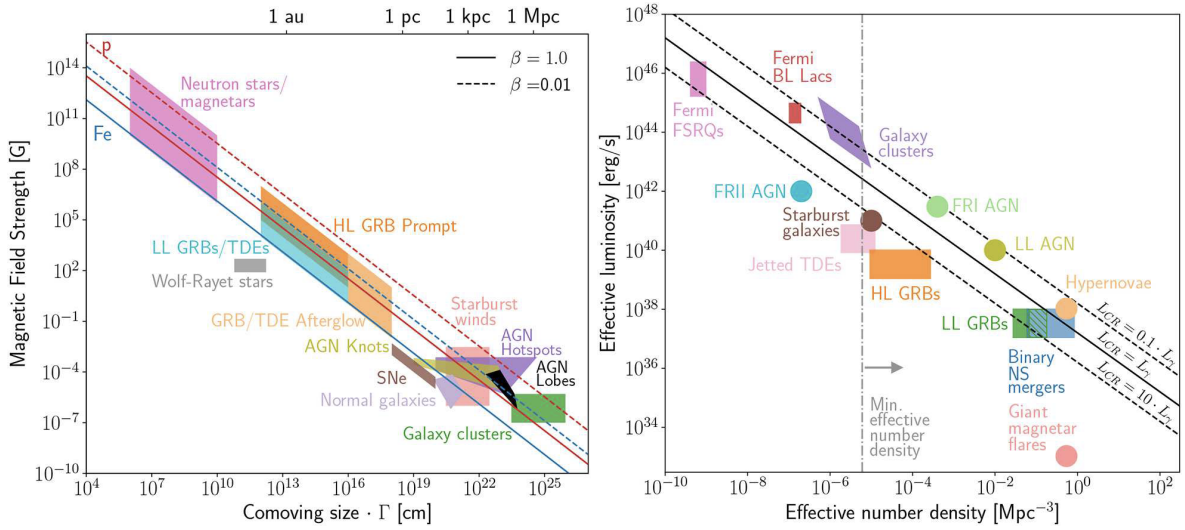


Figure 2.5: On the left, the Hillas condition illustrates the acceleration power of sources. The lines define the minimal size and magnetic field required to accelerate a proton (red) or iron nuclei (blue) to a maximal energy of 10^{20}. On the left, the energy budget condition is shown. The lines define the condition needed for a source class to explain fully the CR flux, using different hypotheses of visible luminosity to CR luminosity conversion. On both plots, sources are required to be on the right of the diagonal lines to fulfill the conditions. Plots taken from [20].

Several source candidates remain: gamma-ray bursts (GRBs) from energetic supernovaes (SNs) have been prime candidates over the years for UHECR. Active galactic nuclei AGNs, whether blazars with their jets towards the Earth or radio galaxies, where the acceleration could occur in the extended lobes are also potential sources. TDEs are in the focus recently, especially after the discovery of a coincident neutrino [2] which was described at the beginning of this chapter. Although from SWIFT data, these would not be energetic enough, theoretical arguments from [23] could make them meet the energy budget. Galaxy clusters are hypothetical sources as well. Although their magnetic fields are in the order of microGauss they are extremely extended in size. At the other end of the Hillas line are located the magnetars which possess an extremely high magnetic field, in the order of gigaGauss to teraGauss. Finally, the starburst galaxies, galaxies that are experiencing a high rate of star formation, were found to correlate best with the Auger data with a significance of 4.0σ, are also a candidate [24]. The acceleration process would involve shocks from the superwind flowing from the star creation engine [25]. Nonetheless, up to this day, no source has been yet identified with certainty as a source of UHECR.

2.3 Electromagnetic spectrum - gamma rays

Keep your eyes on the stars, and your feet on the ground. – Theodore Roosevelt

From the ancestral ocular observation of the stars to the more modern space telescope, the most intuitive way to look at the universe is the observable frequencies of the electromagnetic spectrum. These observations have been extended to cover a much broader range of frequencies, from radio to γ-rays. Depending on the frequency of interest, the tools and detectors differ, using telescopes in space, radio antennas, and Čerenkov telescope on the ground.

γ-rays are of special interest in multimessenger astronomy as their creation processes are believed to be closely related to charged cosmic rays. First via the interaction of the latter with the CMB, similarly to the UHECR, creating high energy (HE) γ–rays. Secondly, the various emission process models predict the creation of γ-rays as secondary particles together with cosmic rays via interaction processes, for example, the production of photons via the p-π_0 decay. High-energy γ-rays are usually detected on the

ground using atmosphere Čerenkov telescope, like the H.E.S.S. [26] or MAGIC experiments [27], taking advantage of the Čerenkov emission of air showers initiated by the γ-rays and effectively using the atmosphere as calorimeter (more details in Section 3.3). Another method uses particle detectors on the ground as the HAWC [28] or LHAASO [29] experiments do. The main downside of γ-rays astronomy is the high absorption of HE photons by the CMB, as predicted in 1965, by Gould and Schréder [30]. As the HE photons interact with the relic photons of the CMB, they create positrons and electrons via the inverse pair production process: $\gamma + \gamma_{CMB} = e^+ + e^-$. As a consequence, the universe is mostly opaque to γ-rays as shown on Figure 2.6. Therefore, the efforts in gamma astronomy are focused mainly on the detection of the most energetic sources in our local galaxy.

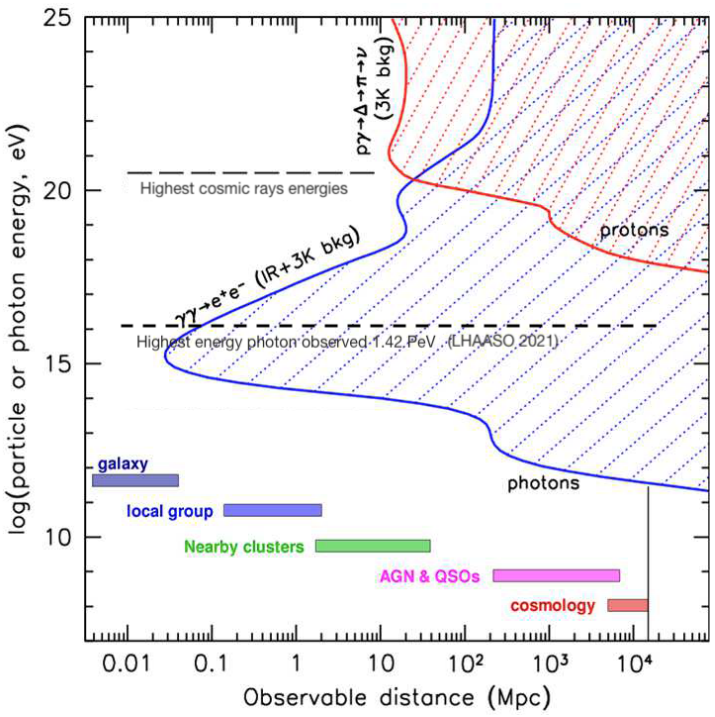


Figure 2.6: Distance at which the universe becomes opaque to certain messengers. The length of the dashed lines is arbitrary. Modified [31].

One exciting recent result in this field is the first detection of PeV photons in the Milky Way by the LHAASO collaboration. In 2021, the LHAASO Collaboration cataloged 12 sources emitting between 0.21 to 1.42 PeV γ-rays [32]. Although the sources creating these energetic γ-rays have not been identified with confidence yet, this is a great leap forward in the search for galactic PeVatrons.

2.4 High Energy Neutrinos

I have done a terrible thing, I have postulated a particle that cannot be detected
- Wolgang Pauli

There exist many types of neutrinos: solar neutrinos, atmospheric neutrinos produced by cosmic rays air showers, reactor neutrinos produced by nuclear plants, all of them studied especially for constraining parameters of the neutrino oscillations; punctual supernovae neutrinos as well as expected cosmological neutrinos of very low energy that would have been created in the early universe, similar to the relic cosmic microwave background. On the high energy end of the neutrinos spectrum are located the astrophysical neutrinos and the cosmogenic neutrinos. The diversity and the large span in the energy of neutrinos are illustrated in Figure 2.7.

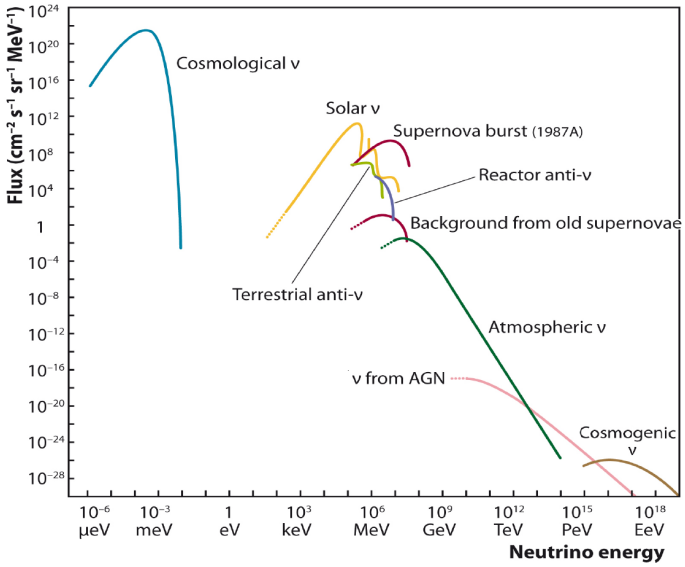


Figure 2.7: Measured and expected flux of neutrinos according to their creation processes. Taken from [33].

The astrophysical neutrinos are coming from extra-galactic sources, and their discovery was the prime design goal for the IceCube Neutrino Observatory, which successfully made the first detection of PeVs neutrinos in August 2011 and in January 2012 [34]. Since then many more astrophysical neutrinos have been recorded, and recently one at the Glashow resonance with a recorded event energy of 6.05 ± 0.72 PeV, where the theoretical Glashow resonance is expected to peak at an anti-neutrino energy of 6.3 PeV [35]. The field of neutrino astronomy opens a new window into the universe as these messengers can travel very large distances without interacting as well as without losing energy other than with the universe expansion. Moreover, due to their neutral nature, they point directly to their sources without being deviated by magnetic fields. IceCube is the leading experiment at the moment in astrophysical neutrino astronomy. There are plans to build more cubic kilometer detectors in the northern hemisphere to increase the resolution in the northern sky as well as plans to increase the volume of IceCube (see Chapter 4) to reach higher energies.

The cosmogenic neutrinos are a predicted ultra-high energy population of neutrinos expected from the interaction of the cosmic rays with the CMB photons, as explained earlier. The expected energy of these neutrinos is above 10 PeV and they have so far not been detected [36]. Nonetheless, their detection would clarify the question of whether the cosmic rays are experiencing the GZK cut-off or not. This is a great example of how messengers can complement each other.

Because of all the reasons stated before, neutrinos are great probes of the universe. However, they are challenging to detect because of their extremely low cross section in the weak interactions, hence necessitating huge detector volumes.

2.5 Gravitational waves

The real reason why general relativity is widely accepted is because it made predictions that were borne out by experimental observations.
– Brian Greene

It is worth spending some words talking about the newest addition to the multimessenger astronomy field: gravitational waves. These ripples of the space-time fabric were already predicted by Einstein in his general relativity theory in 1916. It was calculated that the merging of ultra-massive objects would cause big enough ripples in the fabric to be detected on the Earth with ultra-sensitive detectors. The

working principle of the detector is the observation of tiny differences, in the order of 10^{-20} nm[1], in the interference pattern of two lasers traveling kilometer distances. In 2015, these gravitational waves were detected for the first time by the LIGO Collaboration and thereafter received a Nobel prize for their discovery [37].

It is clear that new gravitational wave detectors will appear in the next decades. Plans are already being proposed for, amongst others, a European triangular-arms interferometer: The Einstein telescope, as well as the space-based interferometer: LISA. The future experiments of gravitational waves will be very interesting in their ability to probe the very early universe, thus giving insights into the creation processes of the universe. Furthermore, since gravitational wave detectors have a constant full view of the sky, they will be useful in delivering alerts to the other experiments when merger events happen. They are caused by very heavy objects and although a coincidence with neutrino or cosmic-ray has not yet been discovered, their discovery would shed new light into acceleration processes. An interesting event concerning gravitational waves, is a coincident detection in space of a binary neutron star merger, GW170817, with a γ-ray burst, GRB 170817A. The γ-ray burst arrived 1.7 s later than the gravitational wave (GW) [38].

Now that all the messengers have been introduced and their complementary explained, and that the open questions related to cosmic rays were exposed, the next chapter will focus on the principle of air showers and their detection via instrumented arrays on the ground: The tools needed to answers the questions lay out in this chapter.

[1] http://gwplotter.com/

CHAPTER 3

Extensive Air Showers and their Radio Emission

As a consequence of the wide energy range and the exponentially decreasing flux, cosmic rays at the highest energies can only be detected indirectly. The indirect detection of air showers, also referred to as extensive air shower (EAS), via coincidences between ground detectors was pioneered by Auger et al. in 1939 [39]. Since then, the usage of sparse detector arrays on the ground has been the primary method of studying the UHECR. Cosmic ray experiments have different scales, detector types, grids, as well as spacing between the ground detectors depending predominantly on the desired energy range of study. The experiment currently covering the largest surface area is the Pierre Auger Observatory located in Malargüe in the southern hemisphere, with an instrumented area reaching about 3,000 km^2 [40]. Utah houses a smaller array of roughly 700 km^2, named Telescope Array, which complements the Pierre Auger Observatory in the northern hemisphere [41]. Both are designed for cosmic-ray investigation at the highest energies. Going to even smaller scales, spanning an area of a square kilometer and probing the presumed galactic to extra-galactic transition region, is the IceCube Neutrino Observatory with its corresponding surface array [42]. The principal strength of this array lies in the combination of its in-ice and surface detectors. More details on the IceCube and IceTop arrays can be found in the next chapter (Chapter 4).

This chapter discusses the basics of air shower physics, the type of detectors used nowadays as well as their functioning while keeping the radio emission of the air showers as the focal point.

3.1 Air shower physics

As explained in the previous chapter, when an extra-terrestrial particle interacts with a nucleus of the Earth's atmosphere, a cascade of secondary particles develops in the atmosphere. The development of a complete air shower is a complex multi-particle system depending on probabilities and cross-sections. Therefore, Monte-Carlo simulations are the best modern technique to describe the development of an EAS accurately. Albeit this complexity, to understand the main characteristics of an air shower, a simplified model can be used to describe the cascade processes fairly accurately.

The first few interactions produce mostly pions, some kaons, and a few other more exotic particles. Let us assume that a proton induces the air shower, creating only charged and neutral pions on its first interaction. Then let us separate the air shower into two parts: The electromagnetic component and the muonic component, as shown in Figure 3.1. The following sections examine first the electromagnetic cascade instigated by the neutral pions and secondly, the muonic cascade initiated by the charged pions.

3.1.1 The electromagnetic cascade

Neutral pions π_0 decay rapidly, with a half-life $c\tau$ of 25 nm [43], into two photons, instigating an electromagnetic shower.

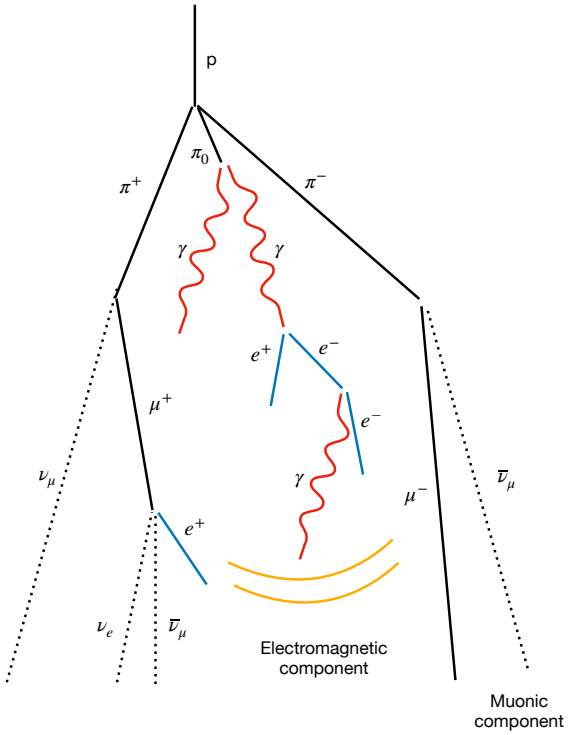


Figure 3.1: Schematics of a simplified extensive air shower (EAS).

$$\pi_0 \longrightarrow \gamma + \gamma \tag{3.1}$$

The electromagnetic (EM) shower can be, in first order, explained by the simplistic Heitler model [44], where all interactions are due to pair productions or Bremsstrahlung. They are approximated with the same interaction length[1] λ ($\lambda/\ln 2$ is around 37 g/cm^2 [43] in air), hence doubling the shower size at each interaction. The number of particles at a certain slant depth X can be expressed by

$$N(X) = 2^{X/\lambda} \tag{3.2}$$

where the slant depth X is defined as the density traversed in the atmosphere along the shower axis expressed and in units of g/cm^2 with $N(X)$ the number of particles at that slant depth. Assuming that the energy is distributed equally at each new branching and that the primary particle had initial energy of E_0, the energy per particle at a certain slant depth $E(X)$ is

$$E(X) = E_0/N(X) \tag{3.3}$$

The increase of particles in the shower continues until it reaches a certain critical energy E_c (E_c=87 MeV in air [45]). At that point, the Bremsstrahlung interactions stop being the dominant mechanism for energy losses. Beyond that, the energy losses from ionization dominate over the creation processes, and the particle number of the EM shower begins to decrease. At that critical point, the shower has reached its maximum development and thus its maximum number of particles. The maximum number of particle N_{max} is defined by

$$N_{max} = N(X_{max}) = E_0/E_c \tag{3.4}$$

[1] For electromagnetic interaction often λ is replaced by the radiation length X_0, where $X_0 = \lambda/\ln 2$. Here we keep λ to not confuse it with the slant depth X.

Substituting Equation (3.2) in the equation above, we obtain

$$X_{\max} = \lambda \frac{\ln(E_0/E_c)}{\ln 2} \tag{3.5}$$

This is an important finding, which means that the physical observable, the depth of the shower maximum, $X_{\max}$, is related to the energy of the primary cosmic ray. In other terms, the higher the energy E_0 of the primary particle, the deeper the shower maximum in the atmosphere is located, with a proportionality described by $X_{\max} \propto \ln E_0$.

Another important aspect emerges from the superposition model of cascades [46]. This model suggests that a nucleus can be considered as a clump of protons, which results in the energy of a nucleus E^{N} of mass A being

$$E^{\mathrm{N}} = E_0/A \tag{3.6}$$

In combination with Equation (3.5), that leads to

$$X^{\mathrm{N}}_{\max} = \lambda \frac{\ln(E_0/A\,E_c)}{\ln 2} \tag{3.7}$$

removing the constants, we obtain

$$X^{\mathrm{N}}_{\max} \propto \ln E_0 - \ln A \tag{3.8}$$

Therefore, for the same energy of the primary particle E_0, the depth of the maximum shower development $X_{\max}$ is higher for a proton than for a heavier nucleus-induced shower.

3.1.2 The muonic component

Charged pions, π^+ and π^-, with a longer mean life of $c\tau = 7.8$ m [43], are responsible for the muon and neutrino content of air showers. They decay into a neutrino and a muon via the weak force

$$\pi^- \longrightarrow \mu^- + \overline{\nu}_\mu \tag{3.9}$$

$$\pi^+ \longrightarrow \mu^+ + \nu_\mu \tag{3.10}$$

Muons, in turn, can further decay into an electron/positron and two neutrinos.

$$\mu^- \longrightarrow e^- + \overline{\nu}_e + \nu_\mu \tag{3.11}$$

$$\mu^+ \longrightarrow e^+ + \nu_e + \overline{\nu}_\mu \tag{3.12}$$

It can be derived from the Heitler-Matthews model [47], which not only encompasses the EM part of the showers but also the hadronic interactions, and the superposition model, that the mean muon number N_μ produced in an air shower is related to the atomic number of the primary cosmic ray (A) and its energy (E) via

$$N_\mu \propto A \cdot (E_0/A)^\beta \tag{3.13}$$

where β is approximately 0.9 in air showers. β is related to the fraction of energy given to the charged pions at each interaction.

From Equation (3.13), we see that the muon number intrinsically carries information about the primary particle that initiated the air cascade. Furthermore, unlike their electromagnetic counterpart, the propagation of muons in the atmosphere is almost without energy loss. Hence, at high zenith angles the only remaining particles from the EAS to reach the ground are muons and neutrinos [46]. To give the reader a feeling of the difference between a proton and an iron nuclei-induced shower, for a primary energy $E_0 = 10^{18}$ eV using $A_{\mathrm{p}} = 1$ and $A_{\mathrm{Fe}} = 56$, the iron nucleus induced EAS contains on average almost 50% more muons than its proton counterpart for the same energy. Regarding $X_{\max}$, the

proton-induced shower will have its $X_{\max}$ on average around $150\,\mathrm{g/cm^2}$ deeper in the atmosphere than its iron nucleus counterpart.

A detector that could accurately measure the number of muons in EAS would be very powerful for mass determination of cosmic rays. It is however not so trivial to disentangle muons, electrons, and gamma signals with ground detectors. There is active research to improve the separation power in already existing detectors (see Section 3.3 for a description of the current detectors). Notably, for example by the usage of machine learning to separate the muon part from the electromagnetic part in the detector signal of water Čerenkov tanks, [48], by carefully choosing the distance where the charges are measured in detectors as in [49], or simply by adding new detectors of a different type to existing ones, as in Auger Prime [19] where additional scintillation detectors are added on the water Čerenkov tanks. An additional complication with muons is that the main hadronic interaction models predict significantly fewer muons at energies above 10^{17} eV than what is measured by most experiments [50]. This discrepancy between simulation and measurement is named the "muon puzzle". One challenge with hadronic interaction models at these high interaction energies is the lack (or the big uncertainties) of cross-section measurements of proton-air collision because the particle accelerators on Earth cannot approach the energies of the cosmic rays. Furthermore, the forward direction of the collision is not probed enough at large particle accelerators. A future detector is proposed to probe further this phase-space [51].

3.2 Radio emission and footprint

With advances in technology and simulations, the radio emission of EAS in air is nowadays well understood and simulations closely resemble measurements [52]. The radiation is produced by the electrons and positrons of the air showers. The muons are too heavy to contribute significantly to the emission and the photons and neutrinos are neutral. Most of the radiation in the EAS occurs close to when the number of electrons and positrons is maximal. An important concept to understand radio emission is coherence. The coherence of the radiation is related to the size of the emission region and the wavelength of the radiation. The size of the emission region can be taken as the thickness of the shower front which is a few meters [53]. This is roughly equivalent to a frequency equal to or smaller than 100 MHz. The highest coherence may be for low frequency but there are other advantages with the usage of higher frequencies. First, the galactic background noise is smaller [54]. This is an obvious reason; a lower noise level allows for the detection of smaller signals, as well as a decrease in the uncertainty of the signal. The second advantage is that the Čerenkov-like ring for the radio footprint is sharper at higher frequencies. The reason for that advantage is related to the determination of $X_{\max}$ and it becomes clearer in Chapter 10. From a few MHz to a few GHz, the emissions processes can be described as follow.

A few experiments have already confirmed the possible detection of air showers with radio, confirmed the similarities with simulations, as well as obtained results about the mass composition of CR. To name a few: one instigator of the modern radio arrays is LOPES [55] located within the KASCADE experiment in Karlsruhe, which is now decommissioned, but paved the way for other experiments. The Tunka-Rex [56] experiment has a radio array located in Siberia as a radio extension within the TAIGA observatory. The LOFAR [57], used for radio astronomy and astroparticle physics, particularly its densest part, located in the Netherlands, is used for cosmic-ray physics. AERA [58], the precursor of the complete radio array planned for the AugerPrime upgrade, is currently in operation in Argentina. Several coming new experiments regarding the detection of cosmic rays are likewise planned to have a radio array, like the future radio array at IceCube, with its prototype station, which is discussed here, or the densely instrumented radio array, SKA [59] in Australia.

3.2.1 The nature of radio emission from air showers

The emission processes from the EM cascade in air with low to medium inclination with a geomagnetic field, are commonly explained via two dominant effects: the geomagnetic effect [61] and the Askaryan

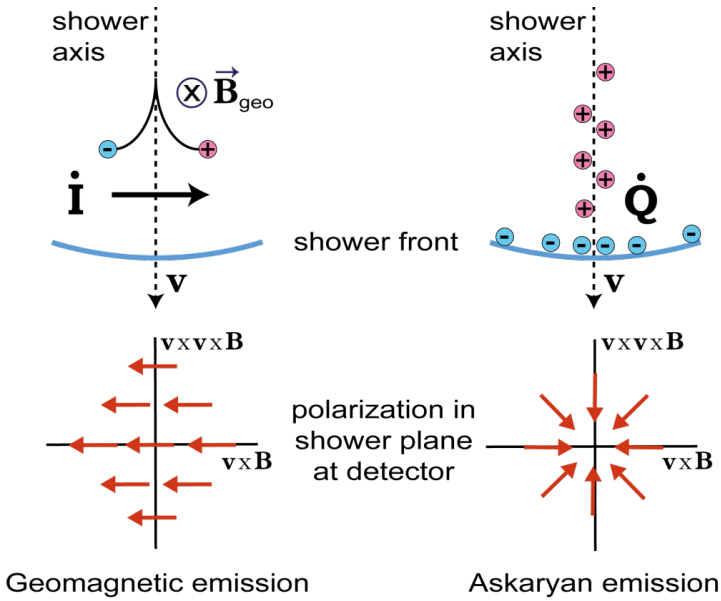


Figure 3.2: Schematic representation of the two main radio emission processes in air showers. On the left, the geomagnetic emission is due to the separation of the electrons and positrons by the Earth's magnetic field. On the right, the Askaryan (or charge excess) effect is due to the time-varying negative charge excess in the shower front. The second row shows the direction of the polarization of each mechanism respectively in the $\vec{v}\times\vec{B}$ and $\vec{v}\times(\vec{v}\times\vec{B})$ coordinate systems. Taken from [60].

(or charge excess) effect [62]–[64].

The **geomagnetic effect** is due to the interaction of the positrons and electrons of the EAS with the Earth's magnetic field ($\vec{B}$). When they travel across the B-field, the Lorenz force deflect them in opposite direction on the $\vec{v}\times\vec{B}$ axis, inducing a drift current. Since the number of electrons and positrons changes as the shower develops, the current varies with time creating an electric field. The amplitude of the emission is proportional to the Lorentz force $\vec{F}_L$ and the Lorentz force is proportional to the cross product of the magnetic field and the shower velocity, as

$$\vec{F}_L = q\vec{E} + q\vec{v}\times\vec{B} \propto (q\,\|B\|\sin\alpha)\,\vec{n} \tag{3.14}$$

where α is the angle between the shower axis and the magnetic field. If we assume the magnetic field constant at a given location, the radio emission due to the geomagnetic effect depends on the azimuth ϕ, the zenith θ, and the number of electrons and positrons created in the EAS.

The **Askaryan effect** emission is created by the ionization of the air during the propagation of the cascade. A time-varying excess of negative charges builds up in the shower front, leaving an ionized plasma along the track. This changing net charge in the shower front generates radio emissions. The Askaryan emission is related to the density of the media in which the cascaded develops and the most sensitive part is at $X_{\max}$ when the shower is at its maximum development.

The geomagnetic emission is not surprisingly maximal when the shower axis is perpendicular to the magnetic field. The Askaryan is, however, invariant to the shower azimuth direction. This can be seen in the left plot of Figure 3.3, where the relative strength of the emission varies with the geomagnetic angle α, which is the angle between the shower axis and the magnetic field vector. The ratio of both emissions is also dependent on the density at the shower maximum $\rho(X_{\max})$ as shown on the right plot of the same figure. This infers a dependency on the zenith direction of the air shower and the weather condition, rendering a global analytical description of the radio emission of an air shower without the usage of Monte Carlo simulations arduous and not as precise.

Interestingly, in the case of low air density showers in a strong magnetic field area, the emission processes become more complex. The synchrotron radiation, where the electrons and positrons start

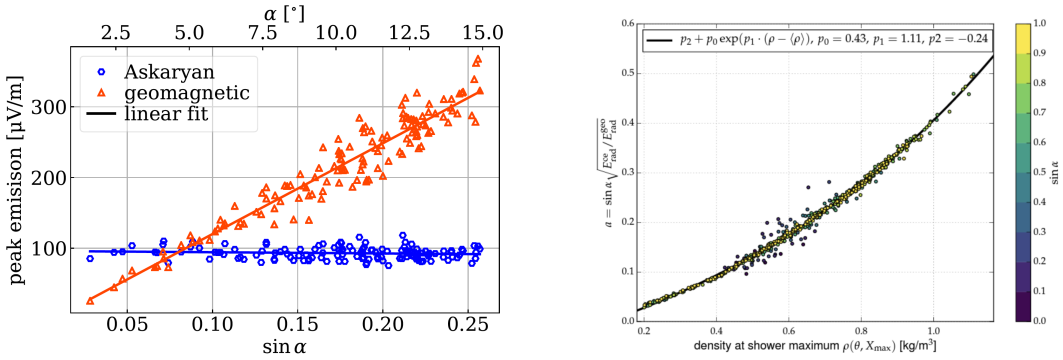


Figure 3.3: Left: Demonstration of the relative strength of the Askaryan emission and the geomagnetic remission as a function of their angle to the magnetic field α at the South Pole. The Askaryan emission is independent of the shower geometry whereas the geomagnetic emission scale linearly with α [65]. Right: dependence of the ratio of the Askaryan radiation to geomagnetic emission on the density at $X_{\max}$ [66].

rotating around the magnetic field lines, becomes dominant. It is also related to the frequency, the synchrotron radiation is dominant at a higher frequency. This was investigated theoretically recently by James in [67]. At the South Pole, this effect presumably plays a role only for very inclined showers (θ>70°) which develop high in the atmosphere [65].

Polarisations

Polarisation is defined as the direction of the electric vector. According to Lenz's law [68], the polarisation of the electric field is opposed to its cause of origin, thus the geomagnetic emission is polarised opposite to the Lorenz force direction, in the $-\vec{v}\times\vec{B}$ direction. For the charges excess, the polarisation of a varying charge moving along a line will point towards the line for negative charges and inversely for positive charges [69]. The shower can be approximated as a negative point charge which varies during its propagation, hence its polarisation will point radially toward the shower axis.

These two emission processes and their respective polarisation are sketched in Figure 3.2. A common way of describing a EAS is by using a plane perpendicular to the shower axis. In radio, a clever choice is to set the $\vec{v}\times\vec{B}$ and $\vec{v}\times(\vec{v}\times\vec{B})$ directions as axes in this coordinate system. This coordinate system is represented in Figure 3.5, compared to the local ground coordinate system. For each individual emission mechanism, we would expect the emission to be rotationally symmetric around the shower axis in a plane perpendicular to the shower axis. However, due to the interference of the polarization of the two emission mechanisms, there is an asymmetry of the footprint in the $\vec{v}\times\vec{B}$ direction.

3.2.2 The Čerenkov-like ring

It is easy by looking at the footprint of the radio emission on the ground, as in Figure 3.5 for example, to see that there is a strong circular emission a few meters away from the shower axis. This is explained via Čerenkov-like effects [70], [71]. Since the electrons/positrons travel at close to the speed of light $v \approx c$ and the radio emission travels slower due to the refractive index of the medium $v = c/n$, by using trigonometric arguments, it can be shown that at a certain angle the particles and the radiation travel at the same speed. This leads to a strong coherence at this position. The angle can be calculated with $\theta_c = \arccos\frac{1}{n}$, which is around 1° in air. The emission at the Čerenkov-like ring is coherent throughout a larger frequency band, up to several GHz [72]. The spectrum of the radio emission at different distances from the ring is illustrated on the right plot of Figure 3.4.

In summary, the geomagnetic and Askaryan emissions explained above lose coherence rapidly with higher frequency (>100 MHz), whereas the coherence at the Čerenkov angle last for higher frequencies.

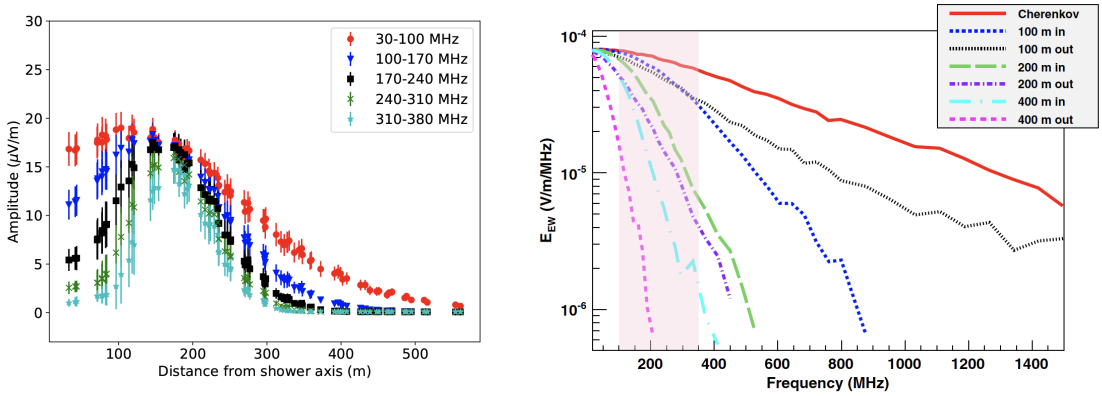


Figure 3.4: We observe the increase of the slope of the Čerenkov ring for different frequencies. 10 PeV proton showers θ=61° α=79° [73]. Left: Spectrum for observers located at different positions relative to the Čerenkov ring, the interesting band is highlighted in red. Slightly modified from [72]

As a consequence, the Čerenkov-like ring is sharper when measuring with higher frequency as shown on the left figure of Figure 3.4.

3.2.3 Frequency

The radio technique is powerful as it contains other information than just the amplitude in the time domain. A Fourier transformation can be done to obtain the total power contained in the time series in each oscillation frequency, i.e. the frequency spectrum. For air shower radio emission, the spectrum differs relative to the position of the observer, as shown in Figure 3.4. The spectrum has maximum power in each frequency at the Čerenkov ring because of the strong coherence in this region as explained above. Then the spectrum becomes steeper the farthest from the ring, with, in general, for the same distance from the ring, a steeper spectrum outside than inside the ring.

3.2.4 Footprint or two-dimensional lateral distribution function

During the shower development a radio emission, whose footprint can range up to a diameter of many kilometers for very inclined showers, is created. The term footprint here refers only to the radiated energy of the air shower at the antenna level, it can also be understood as a two-dimensional lateral distribution function (LDF). The term here can be used for either the emission in the ground coordinate system or in the $\vec{v}\times\vec{B}$ and $\vec{v}\times(\vec{v}\times\vec{B})$ plane. As the measurements happen in the ground coordinate system and usually the analysis is carried out in the $\vec{v}\times\vec{B}$ and $\vec{v}\times(\vec{v}\times\vec{B})$ plane. In Chapter 10, we will review the definition of footprint slightly. Another clarification that must be made here concerns the definitions of energy. There are three types of energy of relevance:

- EM energy (E_{EM})– Total energy carried by the electromagnetic components of the cascade (electrons, positrons and photons).
- Radiation energy (E_{rad})– Energy carried in electromagnetic emission [66],

$$E_{rad} = \int_0^{2\pi} d\phi \int_0^{\pi} dr\, r\, f(r,\phi)$$

where $f(r,\phi)$ is the energy fluence, r, the distance to the shower axis and ϕ the angle from the $\vec{v}\times\vec{B}$ axis. It is the energy deposited by area and is defined as

$$f(\vec{r}) = \epsilon_0 c \Delta t \sum_i E^2(\vec{r}, t_i)$$

where $E(\vec{r}, t)$ is the electric field. E_{rad} is usually calculated in the $\vec{v}\times\vec{B}$ and $\vec{v}\times(\vec{v}\times\vec{B})$ plane.

- Primary energy (E_0) – Initial energy of the cosmic ray before interacting in the atmosphere.

As we saw previously, the geomagnetic emission is dependent on the shower geometry, which means that the radiated energy is different between showers induced by the same primary with the same energy when the shower incoming direction is different. Moreover, the radiation energy is dependent on the density $\rho(X_{\max})$ from the Askaryan effect. This is a challenge because $X_{\max}$ is not known beforehand. However, we can deduce that a vertical shower will have its $X_{\max}$ in a denser area than an inclined shower. It is worth investigating the possible correction of the Askaryan emission relative to the zenith angle θ, which can be known, instead of $\rho(X_{\max})$, but this is outside of the scope here.

Nevertheless, the radiation energy needs to be corrected for these dependencies. Once these corrections are applied, the corrected radiation energy scales quadratically with the primary particle energy [66], [74]. Equivalently, since E_{EM} scales linearly with E_{CR} and E_{rad} scales quadratically with E_{EM}. One issue with E_{rad} is that it is assumed to be calculated after the full development of the shower, once the full energy has been radiated. Questions arise when $X_{\max}$ is below ground, it is unclear at the moment what the meaning of E_{rad} is and how it can be corrected.

We know now that the footprint depends on the shower geometry, on the density of the atmosphere at $X_{\max}$ as well as on the energy of the primary particle. The core position obviously only translates the footprint on the ground, in the shower plane there is no effect. A recent simulation study also found that the refractive index of the atmosphere causes a displacement of the core [75], [76]. This effect is especially concerning for very inclined showers and nonexistent for vertical showers. Finally, as the radio emission is mainly originating from the position of the highest density of particles in the shower electromagnetic cascade development, in other words, closed $X_{\max}$, this makes the footprint sensitive to the distance between $X_{\max}$ and the ground. This aspect is used in Chapter 10 for the reconstruction.

Figure 3.5 summarise a few of the items in this section. In the sketch on the left, the density of the electrons along the shower axis is illustrated, as the radio footprint with the Čerenkov-like ring (which is explained a little later) as well as the $\vec{v}\times\vec{B}$ and $\vec{v}\times(\vec{v}\times\vec{B})$ reference frame.

$X_{\max}$ is indicated in this drawing at the point where the distribution of electrons is at its maximum. This figure also exhibits the difference in the radio footprint for two events with identical parameters except for their $X_{\max}$ which differs. One can see in this figure that the radio footprint differs between the two events, especially the localization of the Čerenkov-like ring.

3.2.5 Symmetrisation of the footprint

There are techniques for collapsing the footprint into a 1D LDF. For that, the objective is to render the footprint radially symmetric. The first correction is universal. It is the disentanglement of the Askaryan effect from the geomagnetic effect. The second is the correction of the early-late effect for very inclined air showers.

Removal of the Askaryan effect

Because of the difference in the polarisation of the two dominant emission processes, it is possible to decompose the measured emission into two components. As we know the expected polarisation of both emission mechanisms, they can be disentangled by clever usage of the $\vec{v}\times\vec{B}$ and $\vec{v}\times(\vec{v}\times\vec{B})$ coordinate system. Two techniques are summarised here. The first method assumes that both emissions add/subtract symmetrically on both sides of the $\vec{v}\times(\vec{v}\times\vec{B})$ mirror plane in the $\vec{v}\times\vec{B}$ direction, then $f^{\dagger}_{\mathrm{vxB}}(x,y) = f_{\mathrm{vxB}}(-x,-y)$ [77]. The geomagnetic fluence f_{geo} and the Askaryan fluence f_{ask} can be calculated with

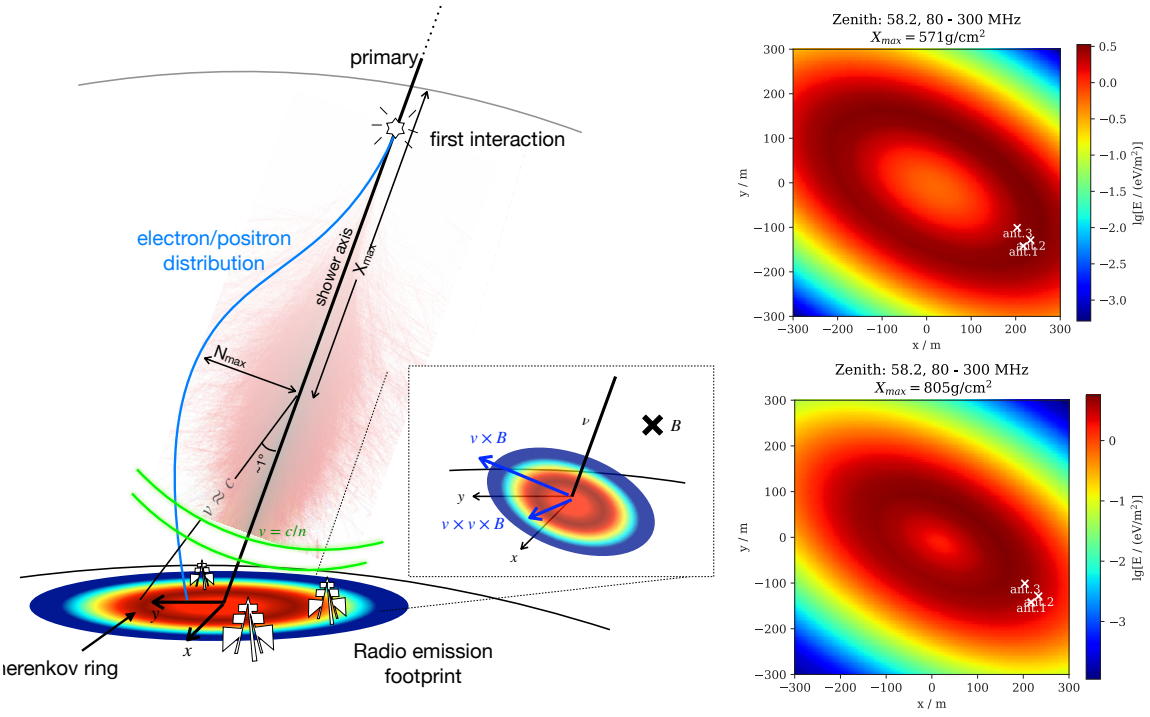


Figure 3.5: On the left, a schematic of a shower development along the shower axis, with the density distribution of the EM cascade illustrated. The definition of N_{max} and X_{max} are pictured as well. The footprint on the ground and in the $\vec{v}\times\vec{B}$ and $\vec{v}\times(\vec{v}\times\vec{B})$ reference system is shown, and the color represent the energy fluence of the signal. On the right, the radio footprint of the same event with a difference in their X_{max} values. On the top, a shower with X_{max} of 571 g/cm^2 , on the bottom, a shower with X_{max} of 805 g/cm^2 .

$$f_{geo} = \frac{1}{4}\left(\sqrt{f_{vxB}} + \sqrt{f^{\dagger}_{vxB}}\right)^2 \tag{3.15}$$

$$f_{ask} = f_{vxvxB} + \frac{1}{4}\left(\sqrt{f_{vxB}} - \sqrt{f^{\dagger}_{vxB}}\right)^2 \tag{3.16}$$

where f_{vxvxB} and f_{vxB} are the portion of the fluence that has its polarization in the $\vec{v}\times(\vec{v}\times\vec{B})$ and $\vec{v}\times\vec{B}$ directions respectively. The limitation of this technique is that if the core is displaced, the concentricity of the footprint is broken and $f^{\dagger}_{vxB}(x,y) = f_{vxB}(-x,-y)$ is not valid anymore. Corrections must then be applied to the footprint to retrieve the concentricity.

The second method describes the emission as a function of polar position in the $\vec{v}\times\vec{B}$ and $\vec{v}\times(\vec{v}\times\vec{B})$ plane. It assumes that the electric field polarised in the $\vec{v}\times(\vec{v}\times\vec{B})$ direction contains only the Askaryan emission. Then this contribution is removed from the geomagnetic fluence. This results in the following relations

$$f_{geo}(r) = \left(\sqrt{f_{vxB}(r)} - \frac{\cos\phi}{|\sin\phi|}\sqrt{f_{vxvxB}(r)}\right)^2 \tag{3.17}$$

$$f_{ask}(r) = \left(\frac{f_{vxvxB}(r)}{\sin^2\phi}\right) \tag{3.18}$$

where ϕ is the angle from the positive $\vec{v}\times\vec{B}$ axis. For the complete derivation please refer to [78]. The limitation of this technique is for points exactly on the $\vec{v}\times\vec{B}$ axis where $\phi = 0$ which makes the term $1/\sin^2\phi$ diverging.

Early-Late effects

Another asymmetry arises from an early-late effect for very inclined showers. This is due to the important time difference between the first antenna hits and the last is substantial. The distance traveled by the radio waves is substantially larger for "late antennas" than for "early antennas". This asymmetry is easily corrected by projecting the emission at the shower maximum assuming a point source and applying a geometrical correction [79], [80].

Multiple radio experiments have used a one-dimensional lateral distribution function with their measurements as in [81], [82], for more vertical showers, and [83] for horizontal air showers. Two-dimensional LDFs has also been used for example in [84], for an absolute uncertainty on the radiation energy of the order of 28%, and primary CR energy of 14%. In 2013, the LOFAR experiment developed a 2D LDF shower-per-shower based fit using simulations [85]. This is widely accepted as being the most accurate reconstruction technique for energy and $X_{\max}$ available for radio experiments so far. Chapter 10 will review this technique in greater detail.

3.3 Detection methods and instrumentation

The standard in experimental cosmic-ray physics is the utilization of arrays of particle detectors on the ground. Over the years, these were complemented by other detector types like telescopes and radio antennas. For high-accuracy measurements aiming at mass composition, it is necessary for experiments to combine detector types together. Especially, because one way of having a good handle on the primary mass is to separate the muonic from the electromagnetic part of the air shower. The currently used instrumentation of the main cosmic rays experiments still operating today are summarized in this section.

3.3.1 Radio antennas

Antennas are not a new technology, Hertz built the first antennas in 1888 and Marconi started developing them for practical usage in 1895. Nowadays, they are widely used for communication, radio channels, and multiple other reasons (including creating a lot of radio background noise). Antennas detect the electric field of propagating electromagnetic waves via the difference of potential experiences by a conductor. This induces an alternating current that can be recorded by an electronic device. They are the physical interface between electromagnetic waves and electricity. There are two types of simple antennas: monopole and dipole. The former is connected at the end of the conductor, referencing the voltage to the ground and the latter has two symmetrical conductors around its terminals. They can receive or transmit, but in cosmic rays experiments, they only need to receive.

It is interesting here to remark that we closed the loop; the electrons and positrons are separated by a magnetic field, this creates a difference of potential $\mathrm{d}V/\mathrm{d}t$, and the difference of potential induces a current $\mathrm{d}I/\mathrm{d}t$, the time-varying current emits an electric field, the electric field when meeting the conductor creates a difference of potential $\mathrm{d}V/\mathrm{d}t$ in it creating a current $\mathrm{d}I/\mathrm{d}t$ that can be read-out by simple electronics.

There is many type of antennas on the market from horn antennas to the simple dipole antenna. There is no one fits-for-all antenna for air shower experiments. For example, Tunka-Rex and AugerPrime use loaded loop antennas, AERA has tested many different types, LOPES and LOFAR operate inverted V-shaped dipole antennas, the ANITA balloon experiment even uses horn antennas positioned in a circle, and the surface enhance will use log-periodic dipole array (LPDA) antennas, which are explained in more details in Chapter 5. Some of the general criteria that need to be considered are omni-

directionality, gain uniformity, desired frequency range, robustness, phase delay, cost, sensitivity to environmental and ground conditions, etc.

Radio antennas are a great tool for cosmic rays studies because, as we saw, the emission from the air showers contains a lot of information. They are also economical and can operate almost 100% of the time. However, there are two sides to every coin, challenges are the absolute calibration of the antennas, as well as the background management, which implicitly encloses self-triggering. The absolute calibration of antennas is not trivial. First, because the calibration of the antennas can be only as accurate as the calibration source, and second because anechoic chamber tests are expensive and the chambers are usually too small for the frequency used in experiments. Nevertheless, the community found innovative ways of overcoming these issues. For the calibration of antennas, sometimes a drone with a reference source is used. This gave an absolute gain from all directions, it is however long to accomplish and still relies on the accuracy of the source [86]. The galactic background radiation is also often used as a calibration source because it is freely available at any location and can be done passively. It is not possible, however, to obtain the directionality of the gain with the method. The background can be reduced by placing the array in a radio-quiet environment. Nonetheless, not all the background noise can be removed as the galaxy is continuously radiating toward the Earth. The galactic noise can be reduced by using a higher frequency [54], because the galactic noise decreases exponentially with increasing frequency. The detection threshold of radio antennas for cosmic rays is naturally dependent on the background at the site, usually, the order of magnitude is tens of PeVs of primary CR energy.

3.3.2 Particle detectors

Ice/Water Čerenkov tanks

The tanks, commonly referred to as surface detector (SD), are arranged in grids on the ground to gather data points on the lateral distribution of air shower particles. The tanks are filled with a transparent dielectric material, such as water or ice, to allow for the creation and propagation of Čerenkov light. They function by measuring the energy deposited by the electrons, muons, and photons produced during the development of the air shower. The Čerenkov emission is created directly by electrons and muons when they cross the tanks, while photons require conversion to electrons within the tank. Although the tanks could in principle detect atmospheric neutrinos, their contribution can be ignored due to the small volume of the tanks relative to the cross-section of the neutrinos. The tank signal is primarily dominated by photons close to the shower axis and by muons farther from the axis. In Auger and IceCube, the tanks are designed to reflect light on the walls to maximize the signal, which helps lower the threshold at the cost of reduced timing resolution.

Scintillation detector

Scintillation detectors are deployed in a grid pattern on the ground, similar to the tanks, enabling the lateral distribution of particles to be probed as they pass through the detectors. They are often considered an alternative to tanks due to their ability to measure particle energy with high precision and their cost-effectiveness. Most scintillation detectors utilize plastic scintillation material, which emits light upon de-excitation of its atoms after being excited by a charged particle. In contrast to the tanks, the scintillation detectors used in AugerPrime and IceCube have a primarily two-dimensional volume, with their depth being significantly smaller than the other dimensions. To read out the scintillation light, plastic scintillation bars are coupled to wavelength-shifting fibers and connected to a light sensor. The combination of tanks and scintillation detectors at the same location can aid in disentangling the muonic signal from the electromagnetic signal, as the detector responses to different particle species vary.

3.3.3 Čerenkov telescopes

The non-imaging Cherenkov telescope (NIAC) utilizes the atmosphere as a detection volume to probe the lateral distribution of air showers. By measuring the Čerenkov emission produced in the atmosphere,

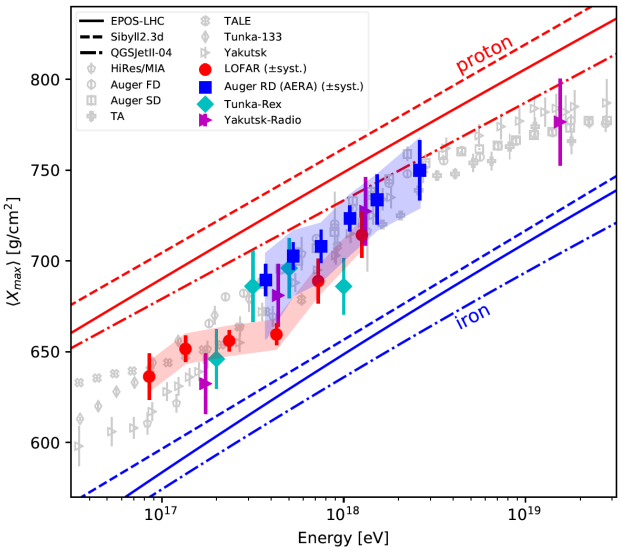


Figure 3.6: Composition of the cosmic ray flux according to the average $X_{\max}$ values. The red lines represent the values according to simulations that the data would have for a 100% proton composition and in blue the same but for 100% iron nuclei. The different line types represent the different hadronic interaction models used. The reconstructions using radio arrays are shown in colors and the reconstruction using other techniques in gray. For the radio measurements, the band represents systematic uncertainties and the error bars the statistical uncertainties. Taken from [4]

a ring on the ground can be recorded and used to reconstruct the position of the shower maximum. Compared to radio antennas, Čerenkov telescopes can record showers with lower energies. While they have traditionally been used for gamma-induced air shower detection, they can also be used for cosmic ray detection. An exciting current project is the IceAct telescope, which is a low-cost and cold weather resistant Čerenkov telescope designed to operate at the IceCube location [87]. However, Čerenkov telescopes are limited by the requirement for clear moonless nights, which reduces their uptime to approximately 10%.

3.3.4 Fluorescence detector

The fluorescence detector (FD) is the only detector able to observe the longitudinal development of the air shower directly, enabling shower-per-shower measurements of the energy and $X_{\max}$. FDs measures the light from the de-excitation of the nitrogen in the atmosphere that was excited during the propagation of the charged particles of the EAS. They deliver high-quality data of energy and $X_{\max}$ because they can probe directly the longitudinal distribution of the air shower. They are also used for cross-calibration of the ground detectors. As for the Čerenkov telescope, the main problem with the fluorescence telescope is that they can only operate in very specific conditions, with a similar up-time of approximately 10%.

3.4 Shower maximum $X_{\max}$ and mass composition

Now that the physical nature of cosmic rays and the importance of mass composition are introduced, this section continues with the definition of the shower maximum $X_{\max}$, why it is a powerful parameter for mass composition, and how to detect it.

Figure 3.6 illustrates the status of mass composition of the cosmic ray flux according to $X_{\max}$ measurements. These measurements agree within the uncertainties at energies higher than ∼2 EeV. There is however a wide discrepancy at the lower energies. Nevertheless, a trend can be seen. The

flux transitions from heavier to lighter nuclei around 200-300 PeV, and reverts back to heavier around 10 EeV. Results of the three radio experiments are shown in this figure: AERA tends to predict a lighter composition than LOFAR, and Tunka-Rex shows a different trend. Inter-collaboration efforts are ongoing to bring the results in better agreement. The results from radio experiments still bear important systematic and statistical uncertainties. The statistical uncertainties are reduced with larger experiments and longer exposure. The systematic uncertainties, in turn, are mostly arising from the calibration uncertainty of the antennas as explained in Section 3.3.1. For example, the systematics of the LOFAR measurements are dominated by the gain calibration and system response at the level of 13% in the absolute uncertainty scale [88].

As discussed before, besides X_{max} , the muons density, or the number of muons in an air shower is also an indicator for mass composition of the initiating particle of the air shower. However, there is a discrepancy between the experimental data and the simulations (the muon puzzle) which leads to large systematic uncertainties. Fortunately, several new radio arrays are either in R&D, in commissioning, or in construction at the moment. The multi-detector layout of the IceCube array and the future IceCube-Gen2 extension will enable cross-calibration through a wide variety of detector types, thus reducing the uncertainties on energies for all the detectors. In particular, the new radio array of the surface array enhancement, which is at the heart of this work, will add data points in the LOFAR energy range and also at lower energy.

CHAPTER 4

The IceCube Neutrino Observatory and its Surface Array Enhancement (SAE)

The IceCube Neutrino Observatory is a unique detector with its main component buried deep in the Antarctic ice. The observatory is complemented by a surface detector called IceTop. In the coming years, a surface array enhancement will be deployed in order to mitigate the limitations of IceTop, limitations that are detailed in this chapter, and contribute to the understanding of cosmic ray physics. This chapter will describe the in-ice detector and the surface cosmic ray detector, explain the motivation and design of the surface array enhancement (SAE), and introduce the prototype station deployed in January 2020.

4.1 IceCube - the in-ice detector

The in-ice detector of the IceCube Neutrino Observatory [42] is a 1 km^3 instrumented volume detector located at the geographic South Pole. As shown on Figure 4.1, the neutrino detector is situated from

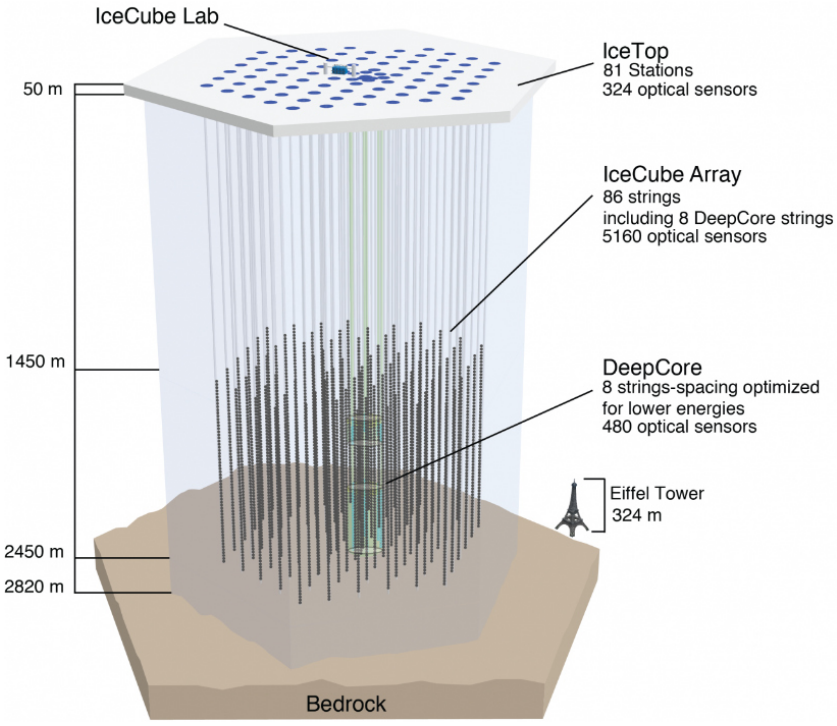


Figure 4.1: The IceCube Neutrinos Observatory with its sub-array DeepCore, and its surface component IceTop. Taken from [42]

around 1450 m to around 2450 m below the icy surface. The final configuration of the standard array comprises 78 cables (strings), where each string has 60 digital optical module (DOM) deployed on a triangular grid with 125 m spacing, effectively forming a hexagonal shape. The DOMs serve as photon sensors and are depicted on the left of Figure 4.3. They are made of a 13" glass vessel carrying and protecting the necessary electronics to operate the 10" diameter photomultiplier tube (PMT) that looks downwards towards the bedrock. The observatory is designed to detect astrophysical neutrinos in the energy range of $\mathcal{O}$(TeV) to $\mathcal{O}$(PeV).

In the center, a more densely instrumented area, called DeepCore [89], increases the sensitivity to low-energy neutrinos. DeepCore complements the standard array with 8 additional strings. It has different horizontal spacing of the strings as well as different spacing of the DOMs. More importantly, the design of DeepCore was optimized for the detection of neutrinos from $\mathcal{O}(10)$ GeV to $\mathcal{O}(100)$ GeV, therefore giving IceCube the ability to study atmospheric neutrino oscillations, test WIMP dark matter annihilation, and measure galactic supernovae neutrinos. To further improve the low energy capabilities of IceCube, an extension with an even more densely instrumented array is planned in the coming years. The so-called IceCube-Upgrade [90] will be composed of 7 more strings, with smaller spacing between the DOMs. This extension will lower the energy to $\mathcal{O}(1)$ GeV, which will enable high-precision detection of tau neutrino appearances and will set the world's best limits on the unitarity of the Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix, which describes the neutrino oscillations. Deviation from unitarity would hint toward new physics beyond the Standard Model.

4.1.1 Neutrino detection

The working principle of IceCube is to detect the leptons produced from the interaction of the neutrinos in the ice via their Čerenkov light. The Čerenkov light is recorded by the in-ice DOMs and converted into charge and timing. It is then possible to reconstruct properties like the arrival direction, the energy, and the type of neutrino that instigated the interaction.

The events measured by IceCube, which are shown in Figure 4.2, can be classified by three signatures: cascades, tracks, and double bangs. The former two compose almost the total recorded events of IceCube, although in 2020, the first two double bang candidate events were identified [91]. Cascades are caused by a neutral current (NC) interaction with a neutrino of any flavor or by a charged current (CC) interaction with an electron neutrino. Tracks are caused by a CC interaction with a muon neutrino and double bangs by a CC interaction with a tau neutrino, resulting in a tau lepton that produces another cascade. In terms of contained events, cascade events provide a good energy estimate as all their energy is deposited within the volume of the detector, the incoming direction of the neutrino is, however, challenging to reconstruct. Inversely, track events deliver a good directional reconstruction because of their long lever arm, yet the energy reconstruction is more difficult than for cascade events.

Neutrinos can also, in principle, be detected via radio emission. Similarly to radio emissions of EAS, discussed in the previous chapter, the interaction of the neutrino in the ice induces an electromagnetic cascade that emits radio emission dominantly via the Askaryan effect (explained in Section 3.2). Although neutrinos have not yet been detected via radio emission, this field received a lot of interest lately and is a site of current ongoing efforts. The detection of neutrinos via radio emission would enable the study of very high energy neutrinos around $\mathcal{O}(10)$ PeV [92], which could reach the expected cosmogenic neutrinos flux (see Section 2.4 for more details about cosmogenic neutrinos). These very high-energy neutrinos could not be surveyed with the traditional photon sensors due to the prohibitive costs of instrumenting an area big enough that would entail.

4.2 IceTop – the surface component

IceTop[1] is the surface array detector of the IceCube Neutrino Observatory. The detector is composed of 162 tanks filled with clear ice, each containing two downward-looking DOMs that are configured with

[1] This section is highly inspired by the design report of IceTop, for more information please refer to [93]

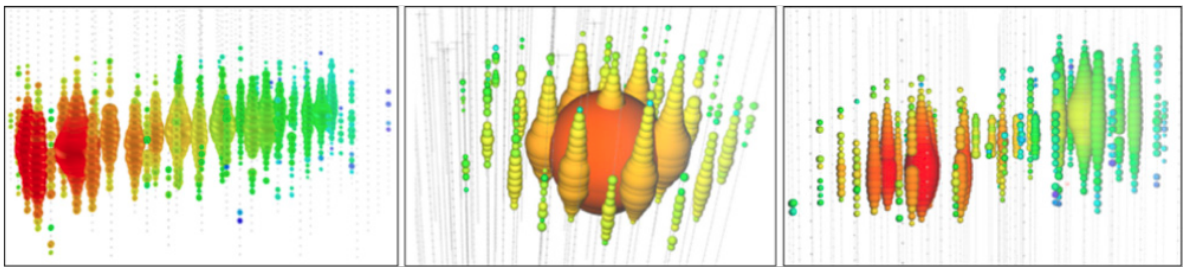


Figure 4.2: This figure shows different signatures of neutrino flavors in the IceCube detector. The left panel displays a track event, the middle panel displays a cascade event, and the right panel displays a double bang event. The size of the sphere indicates the charge deposited in the DOM, and the colors represent the timing, where red is early and blue is late. The left and middle events are from measured events, while the right one is a simulated event. The image is taken from [92].

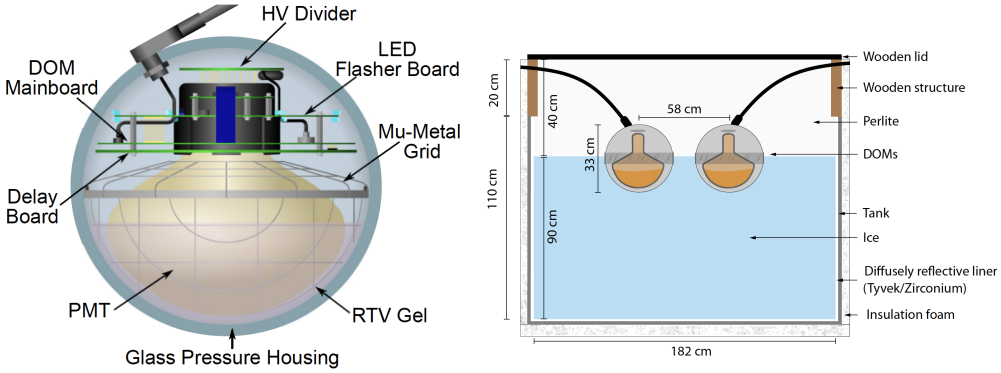


Figure 4.3: On the left a close-up view of a DOM. On the right, a cut-view of an IceTop ice Čerenkov tank with the two DOMs in optical contact with the ice. Not to scale. Taken from [93].

overlapping low and high gain. The tanks are grouped in 81 pairs which are distributed approximately above the strings of the in-ice array. The two individual tanks of the pair are separated by roughly 10 m. The tanks measure the photons, electrons, and muons created in the air showers. Electrons and muons produce Čerenkov light on their own when they traverse the tanks while the photons interact in the snow above the tanks or inside the detector volume and expel an electron which will generate Čerenkov light in the tanks. The tanks are lined with a reflective material to maximize the charge deposited in the DOMs. From the signal of the tanks alone, it is currently not possible to distinguish which type of particle deposited energy in the tank. Figure 4.3 shows a close-up of one DOM on the left and a schematic view of the ice Čerenkov tank with the two DOMs in optical contact with the ice on the right. The DOMs of IceTop are the same as the in-ice DOMs.

The purpose of the IceTop detector is two-fold. One is to conduct cosmic ray physics and the other is to reject air shower events for neutrino searches. The latter has a straightforward rationale that if the in-ice detector identifies a neutrino signal with a direction coming from the surface, the measurements of IceTop are used to see if simultaneous hits occur. A recent implementation of a more sensitive IceTop veto in the real-time alert has been developed [94]. From the six alerts of VHE neutrinos from IceCube between June 19, 2019, and December 31, 2020, three were vetoed by IceTop activities, and one within these by the new real-time tool. The principle is simple; assuming a shower plane front with the in-ice direction reconstruction, the expected arrival time of coincident hits on IceTop can be retrieved. A certain time window is allowed to take the front thickness and the curvature into account.

For the cosmic ray studies, the first analysis step is to reconstruct the basic properties of the primary particle and the air shower, which are:

- the incoming direction (θ, ϕ),
- the shower core (x, y, z),
- and the energy of the primary particle (E_{CR}).

Standard IceTop reconstruction of air showers requires five or more coincident surface station triggers, this confers an energy range of the IceTop array from a few PeV to a few EeV. Using special triggering schemes and reconstruction, the energy range of IceTop can be further decreased down to 250 TeV, for more details on the low energy please refer to [95]. The energy is reconstructed via a lateral distribution function fit where a reference signal S_{ref} measured at a reference position from the shower core R_{ref} is found. After calibration with the zenith angle, this S_{ref} signal is proportional to the energy E_{CR}. The lateral distribution of the signal follows:

$$S(r) = S_{\mathrm{ref}} \left(\frac{r}{r_{\mathrm{ref}}} \right)^{-\beta - \kappa \log_{10}\left(\frac{r}{r_{\mathrm{ref}}}\right)} \tag{4.1}$$

where β and κ correspond to the slope and curvature of the LDF respectively. The value found to minimize the influence of the primary type on the LDF, for the specific IceTop layout, used between tens of PeV to Eev energies, is at a r_{ref} of 125 m [96], [97]. κ is fixed at 0.303 from simulation studies whereas β and S_{ref} are a free parameters of the fit. The energy is linearly proportional to $\log_{10}(S_{125})$ where the linear function changes depending on the zenith angle. However, this linear relation with energy and S_{125} breaks outside the calibrated energies and zenith angles of IceTop, which are for vertical shower with cos(zen) between 0.8 and 1. The reconstruction of the shower direction is done by describing the shower front by a sum of a parabola and a Gaussian function, described in [93].

The standard reconstruction uses a module from the IceTray framework [98] called Laputop, which has several configurable parameters. The fit procedure of Laputop is done in multiple steps. First, the LLH minimization is seeded with the center of gravity (COG) of the signal, as well as the arrival time of a plane shower front to estimate the core position and the shower direction respectively. Then a 3-step log-likelihood fit is applied, varying the allowed parameter ranges and fixing/unfixing some parameters. One challenge of the IceTop reconstruction is the constant snow accumulation on the tanks which needs to be taken into account in the EAS reconstruction.

Currently, the snow correction is incorporated into a corrected LDF fit. However, this technique has a caveat as it relies on snow measurements that are only taken twice a year. Processing of the data is on hold between measurements, and the height of the snow accumulation itself is subject to uncertainty due to the considerable delay between surveys. Additionally, the snow correction parameter is challenging to calculate as the density of snow is not uniform over the tank surface and depends on the height of the accumulation. Furthermore, attenuation also varies depending on the type of particle, with muons experiencing significantly less attenuation than electromagnetic particles.

4.2.1 IceCube Laboratory and infrastructure

At the surface of the ice, in the center of the array, the operations building, named IceCube lab (ICL), is located. The signals from the in-ice detector and the surface detector are routed via cables and "online" processed in the server room. The floor of the server room is shielded against electromagnetic interference [42].

As shown on Figure 4.4, the data acquisition system and the online filtering are done in the ICL. The data is archived on a hard disk at the site as well as sent via satellite to the north for further processing. The planned SAE and IceCube-Gen2 that are discussed in the following section will be treated similarly in the ICL. The main difference for IceCube-Gen2 is that the surface detectors and the in-ice strings will share a common elevated fieldhub that will house electronics such as the data acquisition (DAQ) system of the surface array. This allows easy access for repairs or maintenance in the operational years of the detector.

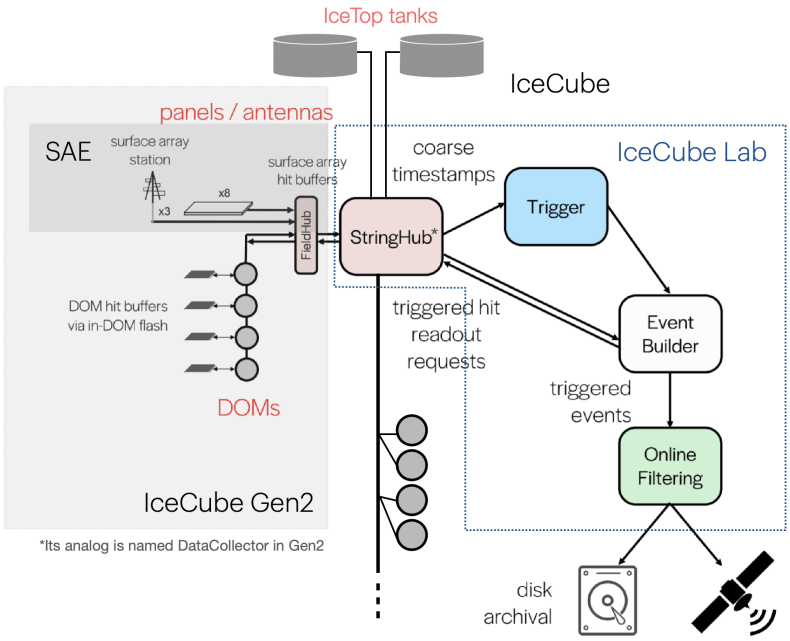


Figure 4.4: Infrastructure for the communication of the different types of detector for the IceCube, as well as triggering schemes, with the addition of the surface array enhancement and the IceCube-Gen2 surface on the left.

4.3 IceCube-Gen2

To increase the number of observed astrophysical neutrinos as well as to possibly detect cosmological neutrinos, which both have a low flux, a larger array is necessary. The implementation of that ambition is presented in the plan for IceCube-Gen2 [92]. The design of the detector aims for the observation of neutrinos from $\mathcal{O}$(TeV) to $\mathcal{O}$(EeV) with a sensitivity to point sources at least five times larger than IceCube, which will be achieved by increasing the statistics by one order of magnitude. IceCube-Gen2 will be the largest neutrino detector in the world and will be composed of three parts: an in-ice optical array, an in-ice radio array, and a surface array. The in-ice optical array of IceCube-Gen2 is designed based on the experience gained from IceCube, and the IceCube-Upgrade. The instrumented volume with its optical modules will be 8 km^3, arranged in a grid pattern with a so-called "sunflower geometry", due to its resemblance to the seed pattern of a sunflower. This particular geometry is chosen to avoid corridors that would be created with a regular pattern and could allow muons to travel undetected until a certain point. The spacing between the strings is planned to be ~240,m, which is approximately twice the distance of the strings in IceCube.

IceCube-Gen2 will be more than only an optical array as it will be complemented by an extensive in-ice radio array. The in-ice radio array is designed for the highest neutrino energies above 10 PeV, and aims for the first detection and the eventual characterization of those neutrinos. The design and analysis techniques will be based on the experience from the radio detector, RNO-G, under construction in Greenland [99].

On top of that, a complementary surface extension, carrying a similar purpose as IceTop for IceCube, is planned for IceCube-Gen2. The surface array will veto the atmospheric muons in the southern sky, which will in turn increase the coverage and statistic of the in-ice neutrino detectors. This array will also advance cosmic-ray science by its combination of multiple detector types and will also be, with the first generation IceCube, the only cosmic-ray array that has an in-ice component able to characterize the high energy muons (from $\mathcal{O}$(TeV) to $\mathcal{O}$(PeV)). This set of detectors comprises scintillation detectors measuring electromagnetic particles and low energy muons ($\mathcal{O}$(GeV)), radio antennas measuring the electromagnetic component of the air shower and its longitudinal development, and the in-ice array

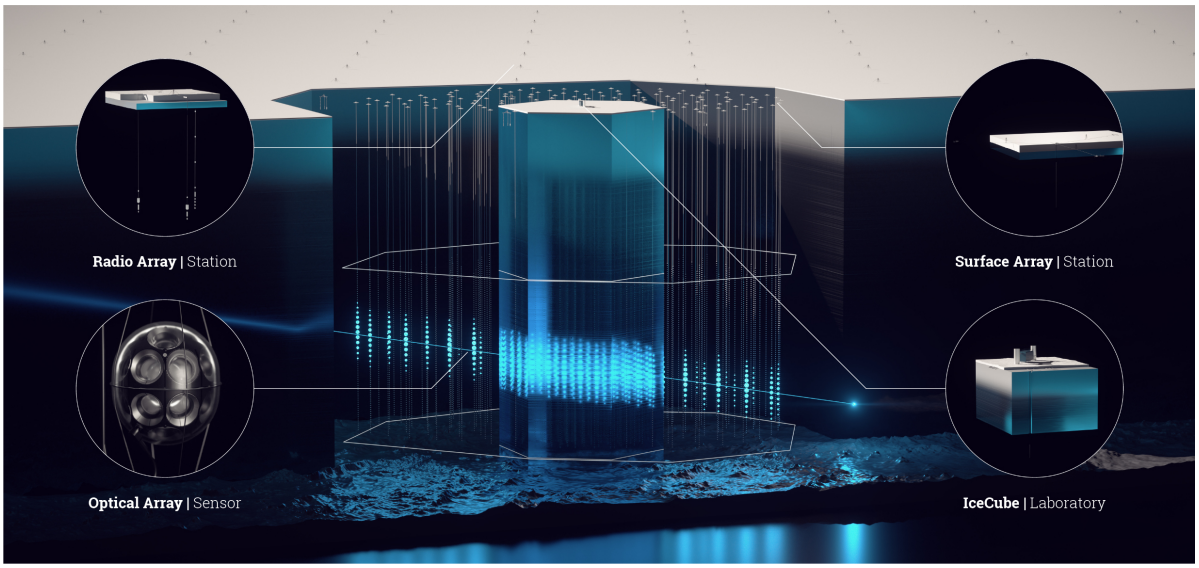


Figure 4.5: Artist view of IceCube-Gen2 with the three components: the in-ice optical array, the in-ice radio array, and the surface array. Taken from the IceCube-Gen2 Collaboration library.

probing the higher energy muons from cosmic ray induced air showers will enable an unprecedented resolution for mass composition in the region where the transition from galactic to extra-galactic cosmic ray is expected to occur. Furthermore, with the surface-to-height ratio of IceCube-Gen2, the aperture for coincidence between surface and in-ice will be roughly 30 times larger than with IceCube only. The surface array of IceCube-Gen2 will be strongly based on the design of the SAE explained in the following sections.

4.4 The Surface Array Enhancement

In order to mitigate the snow accumulation over the IceTop detectors as well as further increase the accuracy of cosmic-ray measurements, an surface array enhancement (SAE) is planned in the next few years. It consists of an array of scintillation panels and radio antennas that will be deployed within the whole IceTop footprint. The upgrade of IceTop with radio antennas will provide X_{max} measurements, a variable widely used to reconstruct the cosmic-ray mass composition as seen in Chapter 3. The scintillation detectors, in turn, will reduce the detection threshold of cosmic rays down to hundreds of TeVs [100] and allow for cross-calibration of the tanks, improving the energy calibration of the data already taken in the last 10 years. Combined, the scintillation detectors, the antennas, the ice-Čerenkov tanks, and the in-ice detector will provide a unique tool for understanding air-shower particle physics and the composition of cosmic rays in the energy range around 10^{14} eV to 10^{18} eV.

The planned layout of the SAE is shown on the left plot of Figure 4.6. It will comprise 32 stations, spread over the hexagonal footprint of Icetop. This design was optimized for ease of deployment considering the trenches and other obstacles as well as reaching a minimum energy threshold for cosmic ray detection [100]. Each of these stations comprises 4 pairs of scintillation detectors, positioned at 90º to each other, 3 antennas, and a central fieldhub protecting the electronics. The stations have a 3-arm star geometry around the central fieldhub as can be seen on the right plot of Figure 4.6. All the devices that are installed on the surface of the ice, i.e. the scintillation detectors, antennas, and the fieldhub are elevated and raisable to avoid snow accumulation in the future.

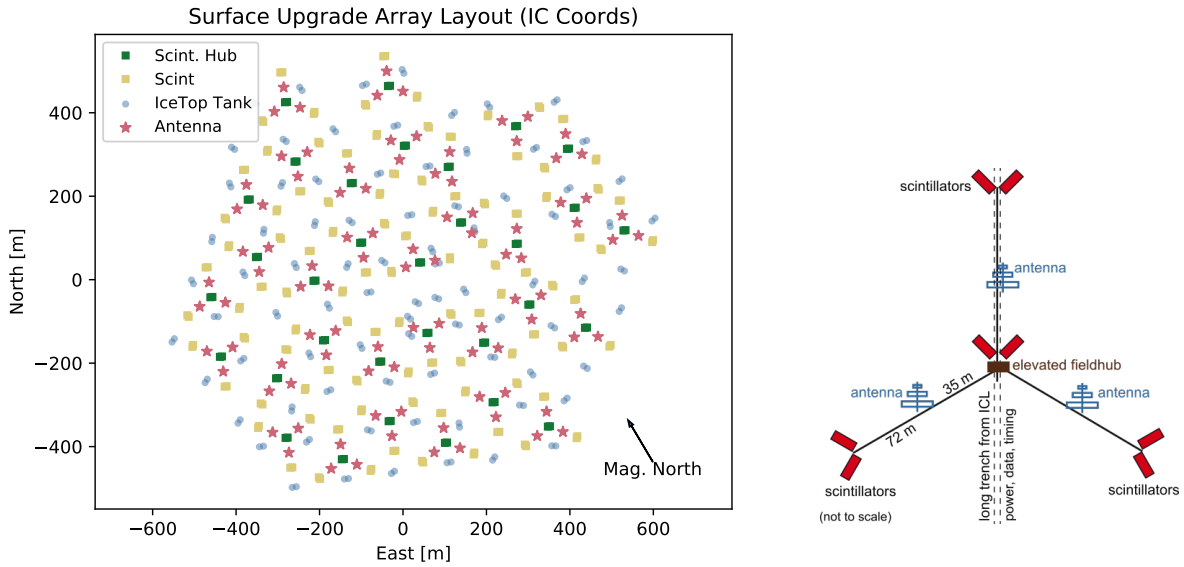


Figure 4.6: On top: layout of the surface array enhancement. On bottom: layout of one station. Bottom illustration produced by F. Schröder.

4.5 Prototype stations over the years

The history of cosmic-ray R&D detectors at the South Pole dates back earlier than this thesis. Several persons were involved in this extensive work of building a new detector. The first complete installation of scintillation panels as a proof-of-concept at the IceCube site was successfully carried out in the southern summer of 2017/18, which used an older version of the DAQ system and scintillation detectors [101]. The year after, further work complemented the array by modifications on the DAQ system to accommodate the addition of two antennas [102]. The first complete station using the almost final design with three antennas, eight scintillation detectors, and one fieldhub was installed in January 2020. This station is the main theme of this work. The antennas and mounts are detailed in Chapter 5, and the whole system and the calibration of the radio electronics are described in Chapter 6

4.5.1 Deployment season 2019/20

This prototype station is recording data since its deployment in January 2020. The layout of the prototype station is shown in Figure 4.7 with the ICL sketched as reference. The scintillation detectors are placed in the standard 3-arm star-shaped station layout whereas the three antennas are slightly shifted from to the station layout. Two of the antennas have a wooden mounting structure (detailed in Section 5.2.2) and one has the fiberglass mount used in the 2018/19 deployment. Pictures of the deployed fieldhub, the antenna, and the scintillation detector are shown in Figure 4.8. All the detectors and the fieldhub were elevated at approximately 1 m above the ground at the moment of the deployment.

4.5.2 Deployment season 2021/22

The electronics of the deployments 2019/20 and 2021/22 and their modifications are discussed in Chapter 6. To summarize it here, after the deployment of 2019/20, it was observed from measurements done with a spectrum analyzer at the site [103] that the electronics are emitting radio frequency interference (RFI) in the frequency band of 70 to 100 MHz, which is also seen in the background spectrum of the prototype station, as discussed in Section 8.3. Other emissions occurring every 10 MHz in the frequency domain were also noticed, with their emitting strength increasing closer to the server

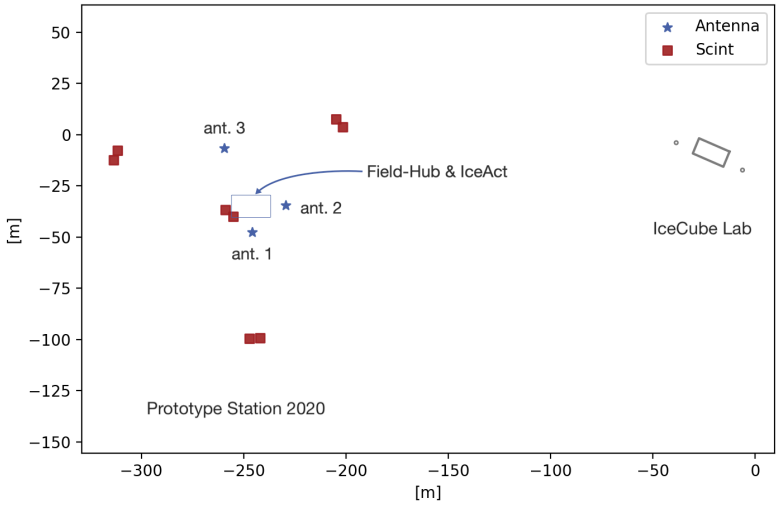


Figure 4.7: Layout of the prototype station deployed in January 2020. The red square represents the eight scintillation detectors, the three blue stars the radio antennas, and in gray the ICL as reference. The square area represents the area in which the fieldhub and the IceAct telescope are located.

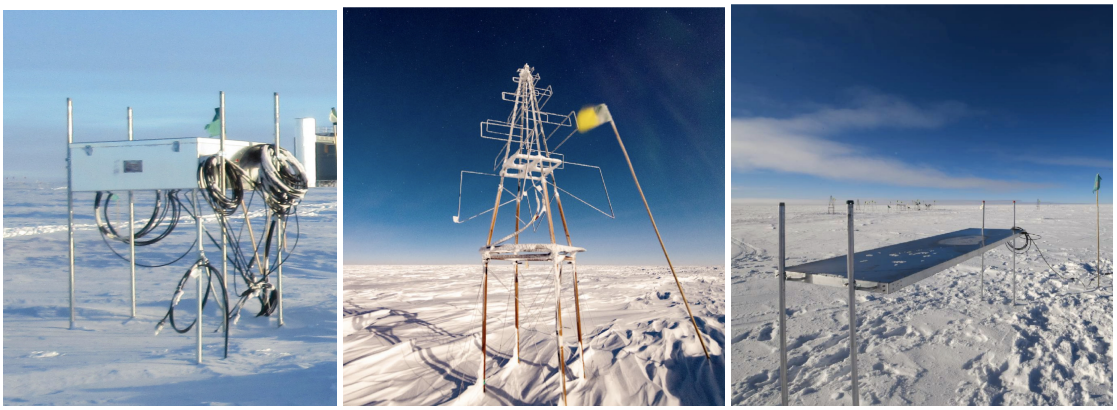


Figure 4.8: Pictures of deployed fieldhub, antenna, and scintillation detector from left to right (2020).

room of the ICL. These emissions are relatively narrow in frequency and can be easily removed in the post-processing of the data by notch filters or other more sophisticated techniques. The emitting noise from the data acquisition system, however, is more complicated to mitigate because it covers a larger band. Modifications on the DAQ system in order to reduce these emissions were conducted. The improved version of the DAQ, TAXI v3.2, replaced the previous TAXI v3.0, in the fieldhub in January 2022 by the winter-over.

4.5.3 Future deployments

In the pole season of 2022/2023, the eight scintillation detectors as well as one radio antenna will be exchanged for their final version. Furthermore, the production of the seven first final stations is completed, and those stations, including all antennas and scintillation detectors hardware, will transit to the South Pole probably via vessel and the material will be stored for one year at the McMurdo station and would be installed during the next open season. The next chapter describes the antennas as well as their mounting structure, which are custom-made for the harsh South Pole environment.

CHAPTER 5

Antennas, Mounting Structure and Field Tests

This chapter covers the mechanical components of radio detectors, while the next chapter will focus on the electrical components. It starts with a description of the antenna used in the prototype station and planned for the complete array. The chapter then goes on to discuss the mounting structure required for deployment at the South Pole and the improvements made based on lessons learned from the prototype deployment.

5.1 The SKALA-v2 Antenna

The SKALA-v2 antenna is of the log-periodic dipole array (LPDA) type and the chosen antenna for the surface array enhancement (SAE). It was developed as a prototype for the Square Kilometer Array (SKA) for its low-frequency array [59]. It is also the antenna foreseen for the IceCube-Gen2 surface array. The LPDA antennas can be understood as a phased array of multiple dipole antennas, this type of aperiodic antenna is renowned for delivering uniform gain over a large frequency band. This specific SKALA-v2 antenna [104] has a wide frequency bandwidth between 50 to 650 MHz and provides a good omnidirectional coverage from the zenith down to 60°. Since the number of showers detected by the scintillation panels, used for triggering, decreases with $\cos\theta$, very few showers are detected with the antennas below a zenith angle of 60° [105].

Figure 5.1 shows the simulated zenith gain for three different cuts in azimuth at 4 different frequencies[1]. The data of the gain of antennas is composed of the gain in ϑ (G_ϑ), the phase in ϑ, the gain in φ (G_φ), and the phase is φ. The data is given by steps of 1° in zenith and azimuth, and 1 Mhz in frequency. The total gain shown in the figure is calculated by converting the gain to amplitude, adding them in quadrature, and transforming them back into gain.

A simulation study showed that the inclusion of the 100-190 MHz band maximizes the signal-to-noise ratio of air shower pulses [106]. Two factors come into play, first the lower galactic noise at higher frequencies, and second the spectral shape of comic ray radio emission. Combined, these factors point to an ideal bandwidth to reduce the detection threshold of cosmic-ray-induced air showers.

Although the antenna can receive frequencies of up to 650 MHz, a low-pass filter is installed directly in the radio front-end board radioTad (see Section 6.5) to filter out frequencies above 350 MHz. The reason behind this design choice is the suppression of emissions from a communication channel at the deployment site which uses the 360.2 MHz band [107]. This would significantly increase the noise level in the signal traces and possibly saturate the waveform with noise. For all the above-stated reasons, the covered bandwidth, the wide and flat gain, the omnidirectionality as well as the sensitivity make this antenna ideal for this radio array, without the extensive work and hurdles of developing a new antenna.

The receiving elements of the SKALA-v2 antenna are soldered onto a common metal pole; this constitutes one arm, see fig.Figure 5.2. Those elements are sized to be in resonance with a defined

[1] Personal communication with E. de Lera Acedo.

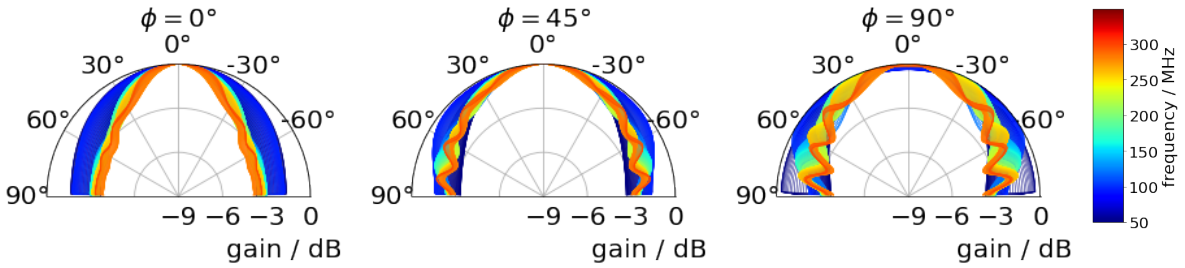


Figure 5.1: Normalized simulated total gain pattern of the SKALA-v2 antenna for three fixed azimuth cuts (0°, 45°, 90°). From data provided by E. de Lera Acedo.

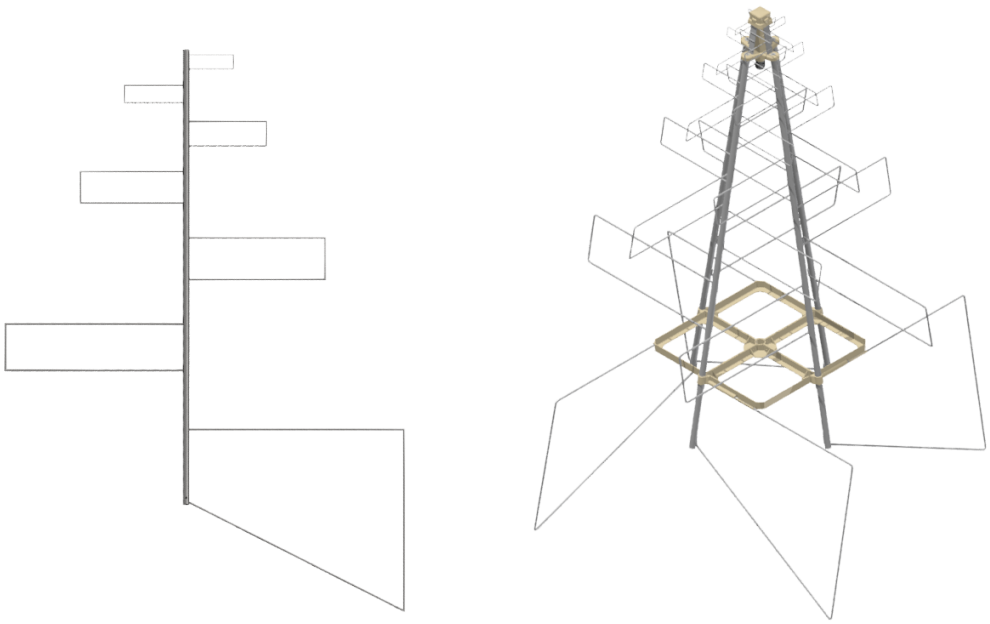


Figure 5.2: On the left, one antenna arms with its log-periodic elements. On the right, the antenna once build with its four arms and two polarisation channels.

frequency, namely half of the electromagnetic wavelength to detect. The spacing between the elements follows a logarithmic law in function of the frequency, hence the name. One antenna is composed of four arms spread symmetrically around the center of the antenna. Two arms face-to-face are a polarisation channel, and the two other arms make up the second independent, perpendicular polarization. In order to be deployed at the IceCube site, the simple supporting structure provided with the antennas and developed for the SKA site had to be changed to conform with specific environmental and functional requirements of the South Pole.

5.2 Mounting structure of the antenna – the mount v1

As mentioned in Chapter 4, the snow height at the site increases by about 20 cm per year. Therefore, for the deployment at the South Pole, in order to avoid the instrumentation being buried by the snow accumulation the detectors need to be on a structure that elevates them above the ground and is raisable over the years. The height of the structure (and the height of its extensions) dictates the number of years before maintenance is required and the maximum number of extensions influences the lifetime of the detector.

Table 5.1: Requirements for the mounting structure

Elevated	> 1 m
Resist temperature	-70 to 0 °C
Easy assembly at the site	2 people/station/day
Metal parts $\ll \lambda_{\text{min}}$	< 125 mm
No extra snow accumulation around detector	
Leveling accuracy	$< 1°$
Raisable	every 5 years
Resist wind	max 100 km/h [1]
Weight and packaging for shipping	< 50 kg
No metal in contact with the antenna	
Affordable	< 200 Euro
Producible at KIT in large quantity	96 mounts total

5.2.1 Requirements

In addition to being elevated and raisable, one of the challenges of installing and operating an array of detectors at the surface of the Antarctic continent is the harsh weather to which it will be exposed over the years. For radio detectors, in addition, considerations concerning the suppression of noise collected from the environment is essential. The detailed requirements affecting the design choices for the mounting structure of the antenna are enumerated in Table 5.1. The general requirement of 1° uncertainty on the leveling comes from the experience of other experiments. In the last section of this chapter, the effect of misalignment for the SKALA-v2 antenna will be quantified, supporting the 1° requirement.

5.2.2 Mount v1

The first two antennas deployed in 2019 were mounted on straight fiberglass legs directly inserted into the antenna arms. The legs were following the angle of the arms and were held together by a thick rope [102]. The mounting structure was fixed to the ground with 1 m long snow spikes which were hammered into the snow. This structure was on the one hand not raisable because of the angle of the legs and on the other hand forbiddingly expensive, with around 1200 €[2] per structure. The mounting structure used in the prototype station of 2020 is named mount v1 and the one that will be installed in the next deployments is called mount v2.

In order to fulfill all requirements, wood is the most suitable option for the structural skeleton as it is affordable and sustains cold temperatures. This is especially true considering that the relative humidity at the South Pole is very low and the temperature is always below the freezing point, hence the wood will barely expand or contract once installed.

Three types of wood were tested in the climate chamber: maple, oak, and sapele mahogany. The dowels, which have a diameter of 30 mm and were cut to a length of 40 cm, are installed in the climate chamber at -70°C and stayed there overnight to test the structural integrity after the exposure of sustained subzero temperatures. Once removed from the climate chamber, they are immediately tested by applying, in the center, the highest possible force without the usage of special instruments and tools. The overall quality of the surface after the freezing is observed as well. All three types of wood passed a breaking test. The surface of the mahogany wood remained sleek after the freezing whereas the maple wood and the oak wood became rougher. Furthermore, the diameter of the dowels and then length is taken at different temperatures, and changes in the dowel size were too small to be noted. The change in the dowel size due to the temperature is thus less than 0.1 cm. Dowels of Sapele mahogany (*Entandrophragma cylindricum*) are chosen as the main material for the mount due to its low water

[2] private conversation with F. Schroeder

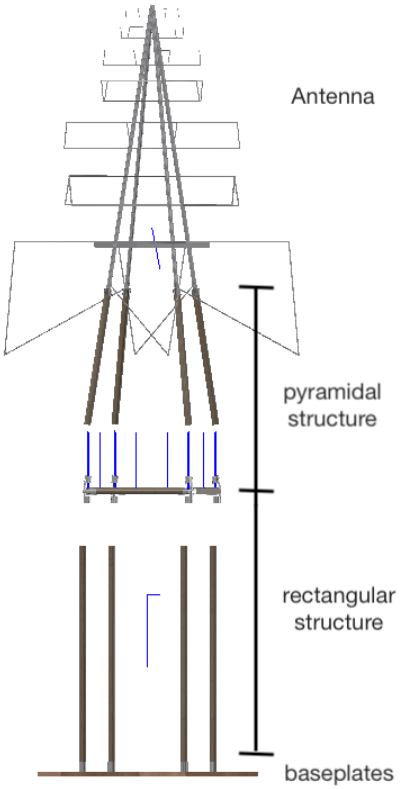


Figure 5.3: CAD drawing of the mount v1 with description of the main structures.

content, its hardness, and the overall good properties it exhibits after the climate chamber tests. Using wood instead of fiberglass successfully reduced the cost of the mount to around 120 €.

In addition to the average yearly snow accumulation at the South Pole, extra snow accumulation around the antenna mounting structure due to the structure has to be minimized. The first design choice is the use of round dowels instead of rectangular parts to ensure a better flow of the wind and the snow around it. The drag coefficient, which defines the resistance of an object to a flow, is around 1.0 for a cylinder compared to about 2.0 for a long square section with the flow on the face [108]. The second design choice is the use of a thin rope for cross-stability instead of more massive wooden parts.

In order to create a mounting structure that can be raised, the angle of the legs must be bridged to straight wooden posts that are perpendicular to the ground. This creates a constant cross-sectional area that does not change as the mount is raised. The antenna mount design includes a pyramidal structure at the base of the antenna that follows the angle of the antenna arms, as well as a rectangular structure beneath that bounds the baseplate to the pyramidal structure. The mount can be raised at the transition point between these two sections. These features are illustrated in a computer-aided design (CAD) drawing of the mount v1, which can be seen in Figure 5.3.

The baseplate used in 2019 [102] consists of two paddle-like wooden plates intersecting in the middle to form a cross shape, and this same design is used for the current baseplate. The dowels of the rectangular structure are 1 meter long, which means that the mounting structure can support 1 meter of snow accumulation before requiring raising. This is equivalent to around 5 years of average snow accumulation before maintenance is needed. Additionally, a 2mm-thick rope is used on the rectangular structure to increase its cross-stability.

This version of the mount features steel tube fittings and metal sheets welded together to connect the different wooden parts. The idea behind it is the reduction of the workload for fabrication by using commercially available tubes. Figure 5.4 depicts the connectors used in the first version of the mount.

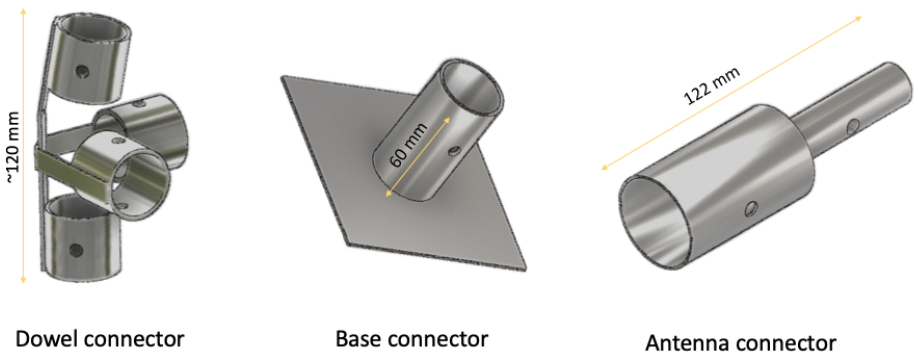


Figure 5.4: Three-dimensional technical drawing of the connectors used in the mount v1

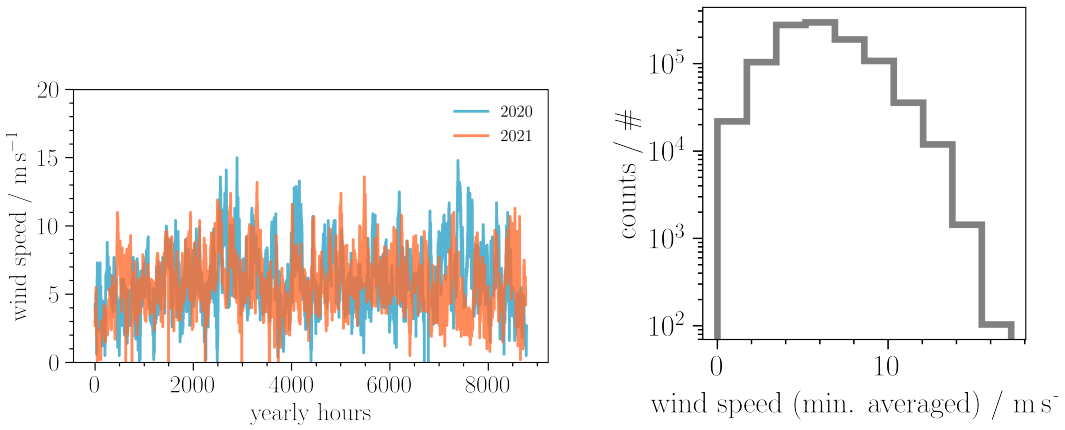


Figure 5.5: Wind speed at the South Pole for two calendar years 2020 and 2021. On the right, averaged for 1 hours plotted versus the hours elapsed in the year and, on the left, an distribution of the wind speed averaged for 1 min.

The dowel connector, on the left, joins the rectangular structure to the pyramidal structure, the base connector, in the middle, fixes the rectangular structure into the baseplate, and the antenna connectors, on the right, are inserted in the antenna and fix the pyramidal structure to the antenna.

5.2.3 Field test at the South Pole

The first version of the mount is installed at the South Pole since January 2020, hence underwent more than two complete years of harsh weather.

The left plot of Figure 1 shows hourly averages of wind speed data for the years 2020 and 2021 from the National Oceanic and Atmospheric Administration (NOAA) station at the South Pole (SPO) in Antarctica, while the right plot displays a wind speed histogram with minute averages to capture possible gusts of wind with high velocity. In both cases, the maximum wind speed recorded was roughly 15,m/s, equivalent to resisting winds up to 54,km/h.

Similarly, Figure 5.6 shows the hourly-averaged temperatures to which the antennas and their associated low noise amplifiers (LNAs), which will be discussed in more detail in the following chapters, were exposed during the two years under consideration. The lines going straight down in the plots are merely artifacts where data is missing, and the value is set at -999.0. The mount v1 thus withstood temperatures as low as -70°C.

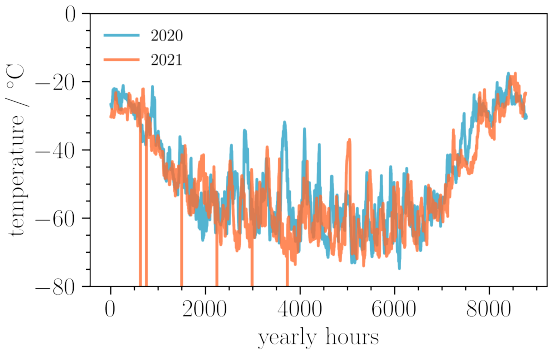


Figure 5.6: Hourly temperature at the South Pole for two calendar years 2020 and 2021 as a function of the time in hours elapsed in the year.

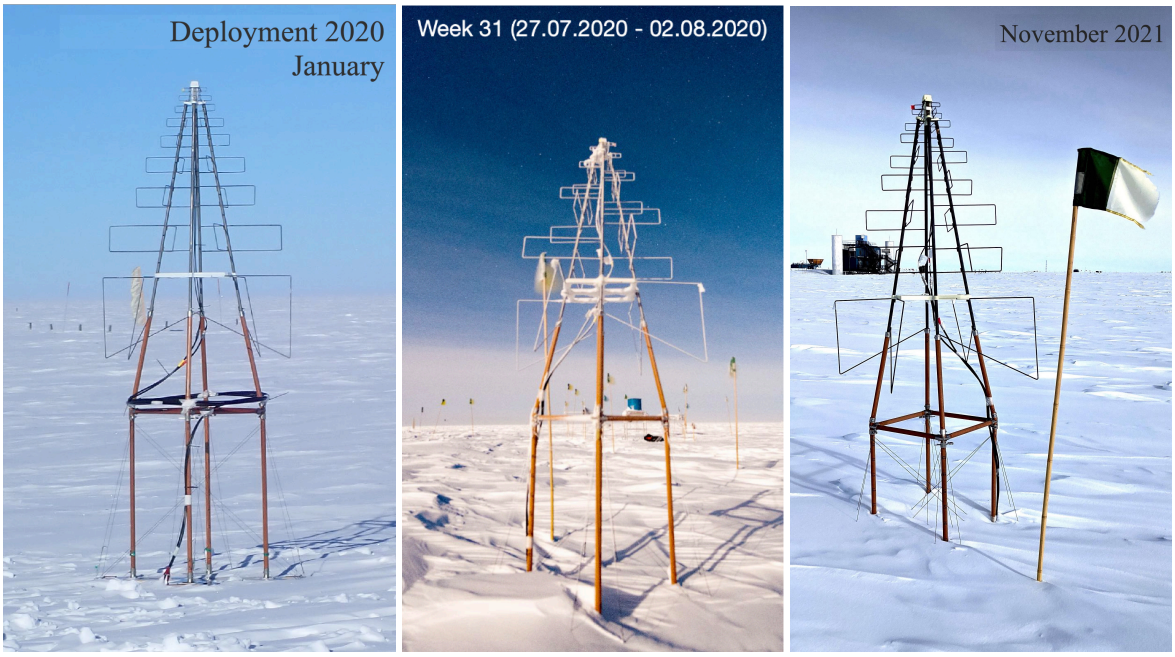


Figure 5.7: Snow accumulation around the antenna and the mount v1 at different dates over the operational time. Note that the mount v1 has 1 m of dowel below the square brace, whereas the mount v2 has 1.20 m.

Figure 5.7 shows pictures of mount v1 with the antenna of the prototype station 2020 at three different times. During the three years of operation of the prototype station, no noticeable extra snow accumulation was observed around the antenna. Instead, the snow accumulation around the antenna corresponds to the average in that area.

In conclusion, the mount v1 is well suited for the harsh condition of the deployment site.

5.2.4 Lessons learned

Despite the success of the mount v1 in withstanding the harsh conditions of the South Pole for almost three years, there are still some areas that could benefit from improvement. Various aspects of the mount's production and deployment were identified as potential areas for enhancement, including improving the tolerance on the tube compared to the wood, simplifying the production of parts that require welding, and making the installation of the rope less complex. Additionally, the connectors

using metal sheet have been noted to bend and the weight of the baseplate could be reduced to facilitate easier shipment. Lastly, the challenge of leveling the antenna by the baseplate on the snow has been identified as an area for improvement.

5.3 Improved design – towards the mount v2

The improved design v2 of the mount was developed based on the lessons learned from the mount v1. While remaining overall similar, the main focus of the improvements was on enhancing the ease of deployment and production at Karlsruhe Institute of Technology (KIT). The issues identified in the first version, such as the tolerance on the wood, welding, rope, and bending of connectors, were addressed with the implementation of new connectors. Additionally, the baseplate weight was optimized through a finite element analysis (FEA), and a system of shims was introduced to enable fine leveling of the antenna during deployment with an accuracy of 0.5°.

5.3.1 Finite elements analysis for the baseplate

The mount v1 weight is below the maximum suggested by the IceCube maintenance and operation leads. Nevertheless, its principal weight factor is the baseplate. It is made of two paddles screwed together in a cross shape in the middle. In order to reduce its weight while ensuring its strength, a finite element simulation is done to optimize the strength/weight ratio. The software Inventor 2019 from Autodesk is used for that purpose. A force is applied symmetrically on the edges of one paddle and the center is fixed in place. This scenario is worse than reality, because the force on the paddle comes from the place where the dowel is connected to the paddle and the snow spikes hammed in the paddle will counteract this force. In the simulations figures (Figure 5.8 and Figure 5.9, the holes closest to the center are where the dowel connects to the paddle and the hole closest to the edges is where the snow spikes are positioned. These figures show the safety factor of the paddle under specified stress, defined below.

Safety factor. The safety factor shown in these plots is commonly used in engineering. It defines the ratio between the expected load (here the simulated load of 50 N) and the maximum load before failure [109]. For example a safety factor or 1 would indicate that the designed load is equal to the maximum load the structure can take before breaking. Usually, the safety factor is between 1.5 and 2.0. The safety factor is there to compensate for five major sources of failure [110], [111]: higher loads than expected, worse properties of the material, imperfect theory on the failure mechanisms, unknown failure mechanism, and human error.

For the bending simulations, a force of 50 N is applied at both ends on one paddle. The drag force, $F_D = 1/2\, C_D\, \rho_{air}\, A\, v^2$, on the mount-antenna combination is in the order of 200 N. However, the baseplate is composed of two paddles assembled together and fixed to the ground with 1.20 m snow spikes at each extremity. Thus, the FEA is using half the wind force to account for these two aspects. As an illustration, a 50 N force, would lift both sides of the paddle by 1 cm. For the torsion simulations two forces of 25 N are applied in opposite directions at the corners of the edges. On Figure 5.8 and Figure 5.9, the yellow arrows represent the force applied. The baseplate of mount v1 is on the left and of mount v2 is on the right in both figures.

For one paddle of the baseplate of version 1, the safety factor in the bending, as well as torsion stress simulations, is over 15 for the whole area of the paddle. The baseplate-paddles were hence overdesigned for the needs and added unnecessary weight. For the mount version 2, the safety factor is above 15 in the torsion simulation and has a minimum of 2.75 in the bending test. By design, the paddle dimensions were reduced until the minimum value of the safety factor lies between 2 and 3. The critical point of the paddle for the bending stress is located where the paddle thickness reduces from 15 mm to 7.5 mm. This groove is there so the two paddles of the baseplate can cross each other while retaining a uniform thickness over the whole area. It is expected that the maximum stress of the paddle is located there. The addition of the other paddle will decrease the load at that point by countering the force of the other paddle. Overall, this baseplate optimization reduced the weight to

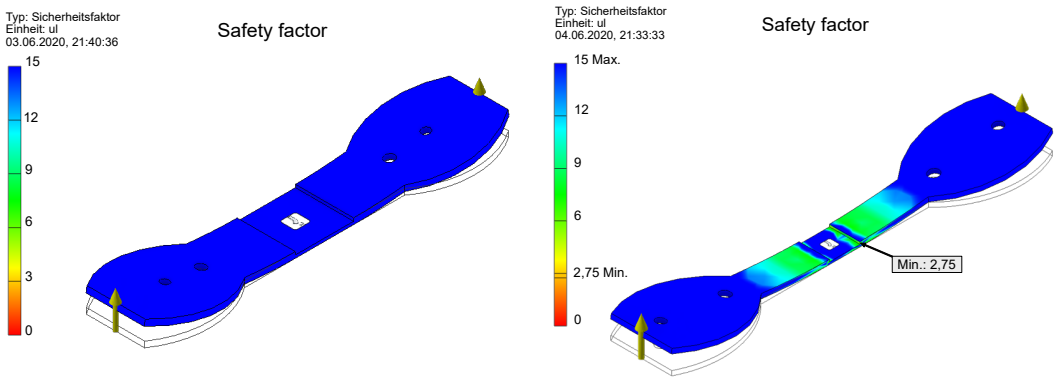


Figure 5.8: Safety factor of one paddle of the baseplate under bending stress. The yellow arrows represent a force of 50 N. On the left, the baseplate of the mount v1 and, on the right, the baseplate of the mount v2 is shown.

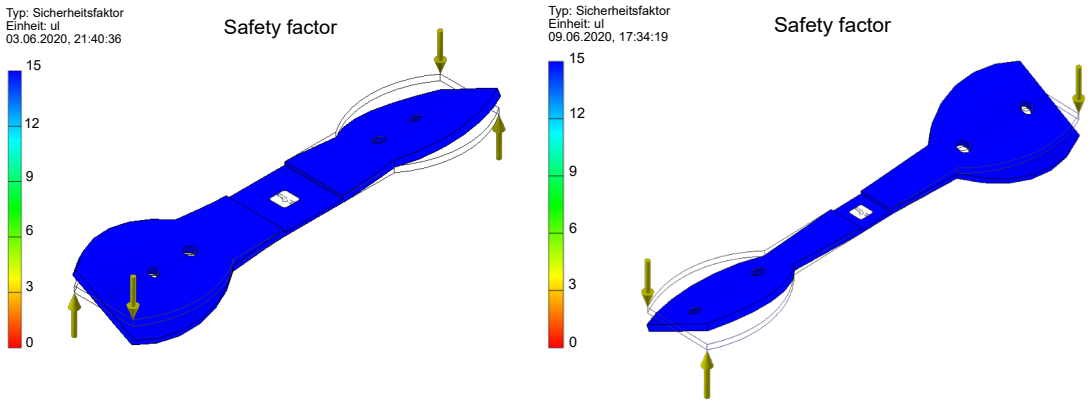


Figure 5.9: Safety factor of one paddle of the baseplate under torsion stress. The yellow arrows represent a force of 25 N. On the left, the baseplate of the mount v1 and, on the right, the baseplate of the mount v2 is shown.

Table 5.2: Baseplates characteristics

	Baseplate v1	**Baseplate v2**
thickness	20 mm	15 mm
width body	150 mm	100 mm
width extremity	200 mm	200 mm
weight per paddle	3.759 kg	2.0 kg
safety factor (bending)	> 15	min. 2.75
safety factor (torsion)	> 15	>15

almost half of the initial value while keeping the structural integrity well over the maximum expected load. The dimensions and weight of a paddle of the baseplate, as well as the important results from the finite element simulation, are summarized in Table 5.2.

5.3.2 General updated for mount v2

Based on the lesson learned from mount v1, an improved mount v2 was designed. As mentioned earlier, the changes from mount v1 to mount v2 are mainly aimed at easing the assembly at the site and

reducing the fabrication load. Since the welding was removed from all the parts except the baseplate connectors, the square connectors, and the angle connectors are made of aluminum to reduce the weight. The square connectors are also hollowed out to reduce further the weight. The antenna connectors are made of *Vetronite® EGS 102* [112], a fiberglass material that is rated for temperatures down to 0 Kelvin. The dowels connected to the baseplate are changed from 1.00 m to 1.20 m in length, in order to gain one additional year of maintenance. The joining point between the pyramidal structure and the rectangular structure is completely redesigned. It is separated into two parts: the square connector and the angle connector (see Figure 5.11 for the nomenclature). A system of threaded rods inserted inside the 1.20 m dowels of the legs enable the square brace to sit properly on these dowels, then the angle connectors of the antenna legs slide into the rods. By this, the whole pyramidal structure is thus joined to the rectangular structure only by the tightening of one nut on the top of each of the threaded rods. The longest metal part in the design is the threaded rod with a length of 100 mm. This redesign decreases the assembly time and increases the cross-stability. Pre-assembly of some parts of the mount is done at KIT. This pre-assembly, which is visualized in Figure 5.10, involves:

- building the square braces,
- screwing the baseplate connectors to the baseplates,
- gluing the antenna legs onto the antenna arms and fixing the angled connector to it,
- fixing the threaded insert and the threaded rod in the 1.20 m dowel,
- and screwing a wooden plate over the snow spikes.

Everything together significantly reduces the assembly time and the work necessary during deployment, reaching approximately 30 minutes of assembly with only eight nuts and bolts to be tightened on the field.

Figure 5.10 shows the mount v2 with the antenna, all the sub-assembly parts are annotated. All the connectors are designed with two grooves on the side to adapt to the low tolerances on the wood dowel's diameters as well as to allow for different temperature-dependent expansion and contraction of the metal and the wood. Those grooves permit the connector to slightly bend around the wood. Once the bolts tighten, the connector and the wood are secured firmly together. Finally, a system of shims that can be inserted under the square parts is developed to obtain an accuracy on the leveling of $\sim 0.5°$, these are shown in Figure 5.12. This translates to an maximal uncertainty of approximately 0.2 dB for air showers with a maximal inclination of 60º. More details about the calculation are in Section 5.4. Those values are much lower than the total 10% (0.8 dB) targeted uncertainty for the radio amplitude to allow for accurate measurements of air showers (see Section 6.1).

The antenna connectors are glued inside the antenna's arm. For this, a two-components epoxy bonding agent, EP21ND [113] from Master Bond, is used. This glue was recommended by the company for the specific operation of the mount and was tested in the freezer at -70ºC. All the technical drawings of the parts used in the mount v2 are in Appendix D.

5.4 Evaluating the effect of a misalignment

The initial requirement is 1º uncertainties on the leveling, which is common in the field. This section proves that 1º is indeed enough to stay below the target uncertainty for the radio amplitude and this is the reason for the design of the shims. All antennas have a specific gain pattern which is never completely isotropic and also differs between frequencies. A misalignment of the antenna translates directly into an uncertainty for the gain. To calculate the influence of a misalignment in the leveling or in the rotation of the antenna, a derivative of the gain pattern is calculated by

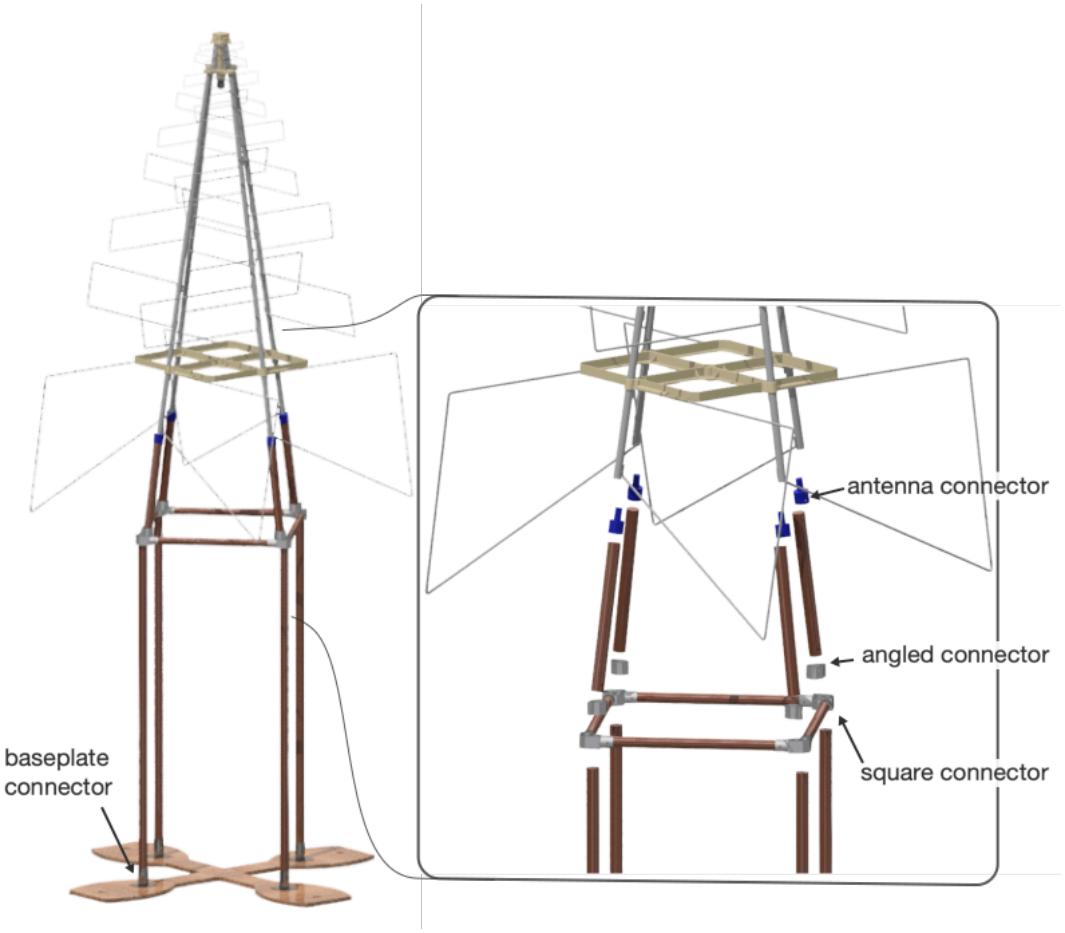


Figure 5.10: CAD drawing of the assembly of the mount v2 with a close-up on the pyramidal part.

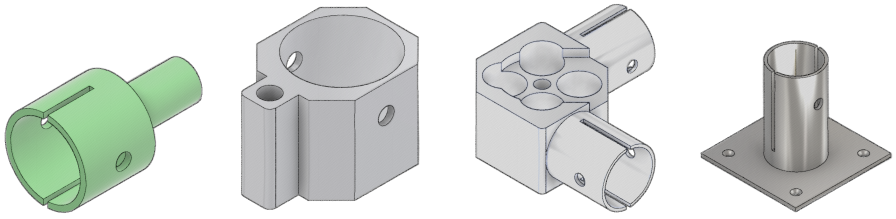


Figure 5.11: Three-dimensional technical drawing of all the connectors used in the mount v2. From left to right: antenna connector, angled connector, square connector, and baseplate connector.

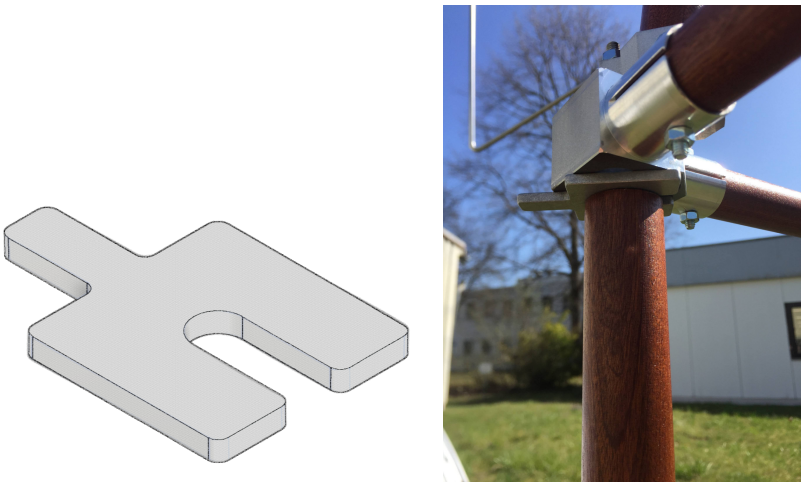


Figure 5.12: On the left, a CAD drawing of a shim for a leveling of 0.5° and, on the right, a picture of a shim inserted below the square connector of the mount v2.

$$\begin{aligned} &\frac{\mathrm{d}A_{f,\phi,\theta}}{\mathrm{d}\theta} = A_{f,\phi,\theta+1} - A_{f,\phi,\theta}, \\ &\forall \phi \in \{0,1,...,360\}, \\ &\forall \theta \in \{0,1,...,90\}, \\ &\forall f \in \{50,51,...,350\} \end{aligned} \tag{5.1}$$

and

$$\begin{aligned} &\frac{\mathrm{d}A_{f,\phi,\theta}}{\mathrm{d}\phi} = A_{f,\phi+1,\theta} - A_{f,\phi,\theta}, \\ &\forall \phi \in \{0,1,...,360\}, \\ &\forall \theta \in \{0,1,...,90\}, \\ &\forall f \in \{50,51,...,350\}. \end{aligned} \tag{5.2}$$

respectively. Where A represents the amplitude of the radio signal, θ the zenith, influenced by the leveling, ϕ the azimuth, influenced by a rotation of the antenna, and f the frequency. To compute these equations the gain is first transformed into an amplitude. Then the derivative of the amplitude is calculated with the previous equation and transformed back into a gain for clarity. For the gain-to-amplitude and vice-versa conversion equations, please refer to Chapter 6.

The Figure 5.13 illustrates the variation of the gain for every frequency between 50 and 350 MHz, in steps of 1 MHz. Different zenith angles are plotted circularly and the radius is equivalent to the variation of the gain for a misalignment of 1° in leveling. The three plots have three different azimuth cuts (0°, 45°, 90°). Comparable, Figure 5.14 depicts the variation of the gain for every frequency between 50 and 350 MHz. Different azimuth angles are plotted circularly and the radius is equivalent to the variation of the gain for a misalignment of 1° in rotation. The four plots have four different zenith cuts (90°, 60°, 30°, 1°).

From the results above, one can see at first glance that the leveling alignment is more critical than the rotational alignment. Nevertheless, both are below the 0.8 dB gain uncertainty requirement for a 1° misalignment. The largest uncertainty, for the two situations, arises from zenith angles larger than 60°, i.e. for very inclined air showers. However, as explained before, those showers are suppressed due to the lower efficiency of the scintillation detectors at those angles. The maximum uncertainty on the leveling decreases from 0.56 dB/° for all zenith angles to 0.42 dB/° for zenith angles smaller than 60°. Similarly, the maximum rotational uncertainty goes from 0.066 dB/° for all showers to 0.050 dB/° for showers with zenith angles smaller than 60°. The Table 5.3 summarizes the maximum values and the average values of the uncertainties for different zenith angles.

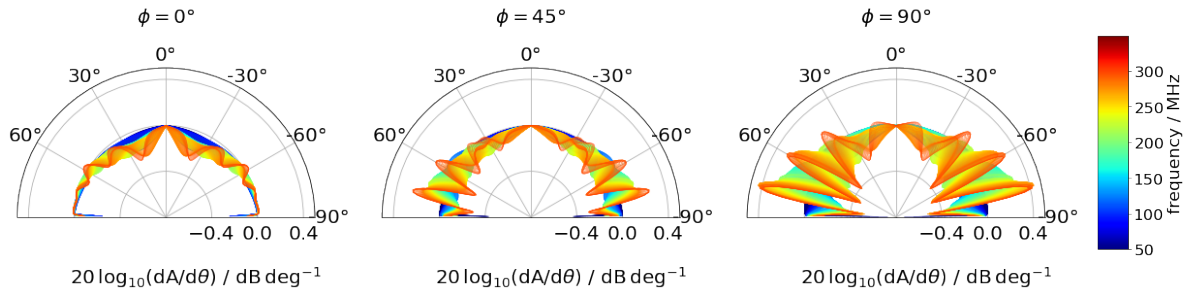


Figure 5.13: Effect on the total gain of the antenna for a 1° displacement in leveling for three azimuth cuts (0°, 45°, 90°)

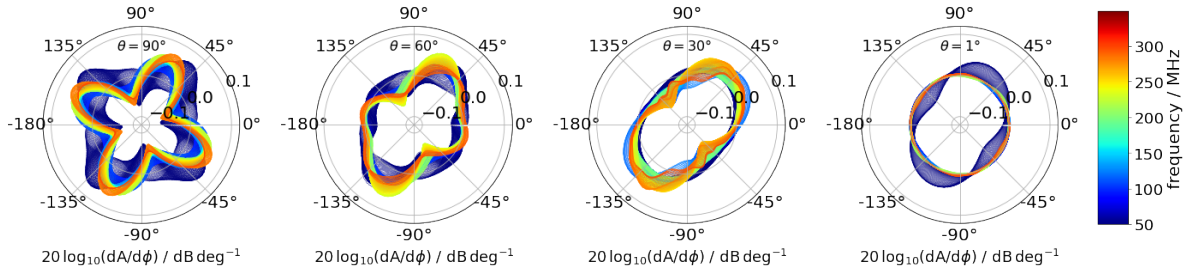


Figure 5.14: Effect on the total gain of the antenna for a 1° displacement in rotation for four zenith cuts (1°, 30°, 60°, 90°)

Table 5.3: Uncertainties on the total gain of the antenna due to misalignment

zenith angles [°]	max. $\mathrm{d}\lvert A\rvert/\mathrm{d}\theta$ [dB/°]	average $\mathrm{d}\lvert A\rvert/\mathrm{d}\theta$ [dB/°]	max. $\mathrm{d}\lvert A\rvert/\mathrm{d}\phi$ [dB/°]	average $\mathrm{d}\lvert A\rvert/\mathrm{d}\phi$ [dB/°]
all	0.56	0.069 ± 0.066	0.079	0.022 ± 0.016
< 60	0.42	0.063 ± 0.050	0.074	0.024 ± 0.016
< 30	0.26	0.051 ± 0.044	0.074	0.025 ± 0.017

The uncertainty of the orientation of the antenna, i.e., on the azimuth, is evaluated from a differential GPS survey in the field. Each extremity of the antenna arms is taken with a differential GPS position device on the site, effectively gathering four positioning points for each antenna. A small study is done to evaluate the uncertainty of the GPS survey from the prototype station antennas deployed. The angle between the two polarisations, α, of each of the three antennas is calculated. This angle is shown in Figure 5.15,

Table 5.4: Uncertainties due to misalignment

	α
antenna 1	89.4°
antenna 2	89.5°
antenna 3	90.7°
standard deviation	0.59°

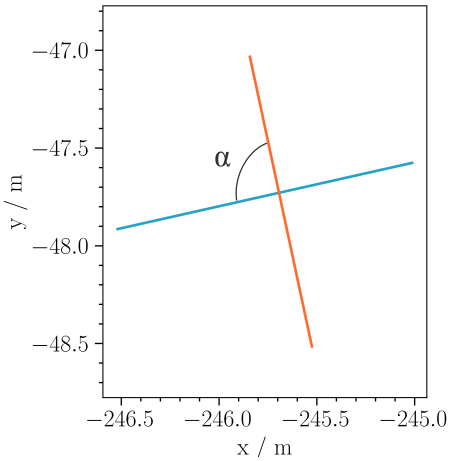


Figure 5.15: Illustration of the two polarisation channels of antenna 1 of the prototype station, and the angle α between them. The illustration is created from the GPS survey done after the deployment, in January 2020.

The Table 5.4 synthesizes the difference between the polarisation channels of the antenna's arms. With a standard deviation of 0.59° for the three antennas, it is reasonable to assume that the uncertainty of the azimuth positioning is in the order of 1°, which corresponds to a maximum gain uncertainty of 0.079 dB. This is negligible compared to the total targeted uncertainty of 0.8 dB. On the other hand, the gain is more sensitive to the leveling. It was thus decided to design the shim system for steps of 0.5°, reaching a maximum uncertainty of 0.28 dB.

5.5 Discussion

Overall, the mount v1 met the necessary requirements and withstood well the harsh climate of the South Pole. It also reduced the cost by an order of magnitude compared to the predecessor deployed in 2018/19. Nevertheless, some shortcomings, mostly touching the assembly time and the production, were highlighted and resolved with the mount v2. The design of the mount v1 as well as the mount v2 allows the raisability of the structure over the years, with mostly infinite extension possibility, only limited by the maximal cable length, needing maintenance every 5 years.

The study found that the zenith uncertainty, which is primarily affected by leveling, is more critical for gain than azimuth uncertainty. To address this, shims were implemented to allow for leveling accuracy of 0.5°, which corresponds to an uncertainty of approximately 0.28 dB on the gain for all showers. Differential GPS surveys were found to be sufficient to achieve an azimuth uncertainty of less than 1°, which corresponds to an uncertainty of approximately 0.08 dB.

Although the mount v2 has not yet been tested at the South Pole, its design is very similar to the mount v1 regarding the bulkiness, and other factors that could affect the snow accumulation or the resistance to the weather. Three exemplars of the mount v2 with an antenna are installed at KIT for around two years and resisted the Karlsruhe weather thus far.

CHAPTER 6

Electronics of the Surface Array Enhancement

The design and calibration of the whole system for the prototype station as well as for the future array involves the work of many people. For coherence, all of the sub-systems needed for the surface array enhancement (SAE) are discussed in this chapter, with a strong emphasis on the radio component of the surface array. For the deployment of the prototype station of January 2020, the system is composed of the TAXI v3.0 and the radioTad v1. Improvements have been achieved for both of these systems and the prototype station has been updated in 2021/22 with the TAXI v3.2 and the radioTad v2. In this chapter, the design and the functioning of the data acquisition (TAXI) as well as the radio front-end board (radioTad) are explained. Additionally the workings of the scintillation detectors and the timing system will be briefly summarized. For more information about the scintillation panels of the prototype station 19/20 and their functioning please refer to [114].

6.1 Requirements of the electronics

For cosmic ray analysis, an accurate measurement of the on radio emission of air showers is desired. To this end, a requirement of 10% total uncertainty on the total radio amplitude is required. The length of the radio waveforms for the completed array must be at least 4 μs long to record the pulses from a horizontal air shower in all the antennas. The scintillation panels must be able to trigger the radio recording, with a flexible triggering scheme controllable via the firmware.

The timing precision between the different antennas should be below 1 ns to enable radio-interferometric techniques [115], [116], and of the order of ms between the instruments to allow for coincident event searches. The nominal bandwidth should include the 100-190 MHz band, which was shown to increase the signal-to-noise ratio of air showers compare to the commonly used 30-80 MHz band [73]. The low limit of the frequency band is around 60 MHz to reduce the exponential rise of the galactic noise in the lower frequency range, while the high limit is at 360 MHz in order to attenuate a communication channel located in the 360.2 MHz region at the deployment site [107]. The sampling rate must be higher than two times the highest frequency, due to the Nyquist-Shannon sampling theorem, which states that to reconstruct a continuous signal reliably from its discrete sampling points, the sampling frequency must be a least twice the highest frequency of the continuous signal. On top of these requirements, low cost and low power consumption are targeted, the second statement is especially applicable for experiments located at the South Pole where the power capacity and generation are limited.

6.2 Data acquisition and related systems

The DAQ of the prototype station is based on a project called Transportable Array for extremely large area Instrumentation studies (TAXI) [117]. The TAXI board underwent multiple iterations as the plans for the surface array developed over the years. The prototype station of 2019/20 uses a version called

TAXI v3.0, which was replaced by the TAXI v3.2 in January 2022. The idea behind the TAXI project was the creation of a universal data acquisition system that could be used as a plug-and-play device for diverse cosmic ray or cosmic-ray-related experiments. In terms of the SAE, that development leads toward a SAE-specific hardware, designed to meet the above-mentioned requirements. The TAXI system v3.2 is foreseen as the deployment-ready version of the surface enhancement array. With IceCube-Gen2, an upgrade of the DAQ is planned, implementing additional requirements. These requirements entail, for example, a longer buffer for the radio data to enable flexible triggering options throughout the different detectors types while keeping the output data low, an accommodation of the radio read-out for a larger array, or the increase of the analog-to-digital converters (ADCs) from 14 bits to 16 bit for a larger dynamic range.

The SAE electronics, located on the field, contain a central processing board, TAXI[1], a timing system, white rabbit (WR) node (WR-LEN) [118], a communication board for the scintillation detector, a fanout board to connect the scintillation panels, and radio pre-processing boards, called radio front-end boards (radioTads). All are housed in a shared metal container with connectors for detectors, power, communications, and timing to the IceCube Lab. That metal housing is also referred to as the TAXI box. The data recorded by the radio and scintillation detectors are first processed within the TAXI box before being sent through a WR layer to the ICL via a glass fiber cable with 1 Gbit bandwidth [119]. The following sections describe the design and the interactions of the SAE electronics with the TAXI as its central processing unit and, furthermore, show the developmental changes of the TAXI from v3.0 to v3.2

6.2.1 The TAXI board

A detailed diagram of the data acquisition system and the related components is shown in Figure 6.1. The TAXI board has a Linux operating system running on an embedded ARM-based microprocessor *Stamp9G45* by TasKit GmbH and a field programmable gate array (FPGA) *SPARTAN-6* by XILINX for signal processing. It is powered by a redundant input of 24 V, which is converted via three DC-to-DC voltage converter to 3.3 V, 4.0 V (4.5 V in V3.2) and 5 V. The FPGA, microprocessor and scintillator uDAQ v4.1 (uDAQ) in the scintillation detectors (Section 6.3) require 3.3 V, the 5 V layer is used by the WR-LEN, the radioTad and antennas have their independent 4.0 V layer. The WR system is by design able to provide nanosecond timing synchronization between the individual scintillation panels, the radio read-out, and the already existing IceCube infrastructure. Although the radio read-out is synchronized with the WR, it is regulated by the FPGA clock which has a ticking of approximately 8.4 ns, which limits the timing accuracy between the scintillation panels and the radio antennas. Connected to the TAXI board, the WR-LEN acts as a slave and synchronizes itself with the master WR switch located inside the ICL, the latter receiving its time from GPS satellites [118]. The scintillation detectors are connected to the TAXI board via a fanout board which serves as an interface for communication, power, trigger signal, and data.

Each antenna outputs an analog signal which, after passing through the LNA (Section 6.4) and radioTad (Section 6.5), is supplied to one domino ring sampler (DRS)4 chip. This sampling chip is a ring buffer chip featuring 8+1 channels of switched capacitor arrays with 1024 sampling cells each. Each channel takes a maximal 1 Vpp differential signal as input. The purpose of this chip is to quantify the time intervals of the analog signal. This is regulated by the sampling rate which can be tuned between 700 MSPS to 5 GS/s. In the prototype station, the sampling rate is configured at 1 GS/s, satisfying the Nyquist limit, which translates to a time quantification of 1 ns bins for the waveforms. The ring buffer of this chip is continuously recording unless stopped by a trigger and operates at a nominal 0.35 mV noise level [120]. Each DRS4 output is then digitized by a 8-channels ADC (*LTM9007IY-14*) with a dynamic range of 14 bits. The read-out is operated via the FPGA on the TAXI.

In order to estimate the maximum energy that the system can record before reaching saturation, one can make some assumptions: first, the system behaves linearly for the complete dynamic range,

[1] The word TAXI here refers to both TAXI v3.0 and TAXI v3.2 because the general functions of the system are equivalent for both versions.

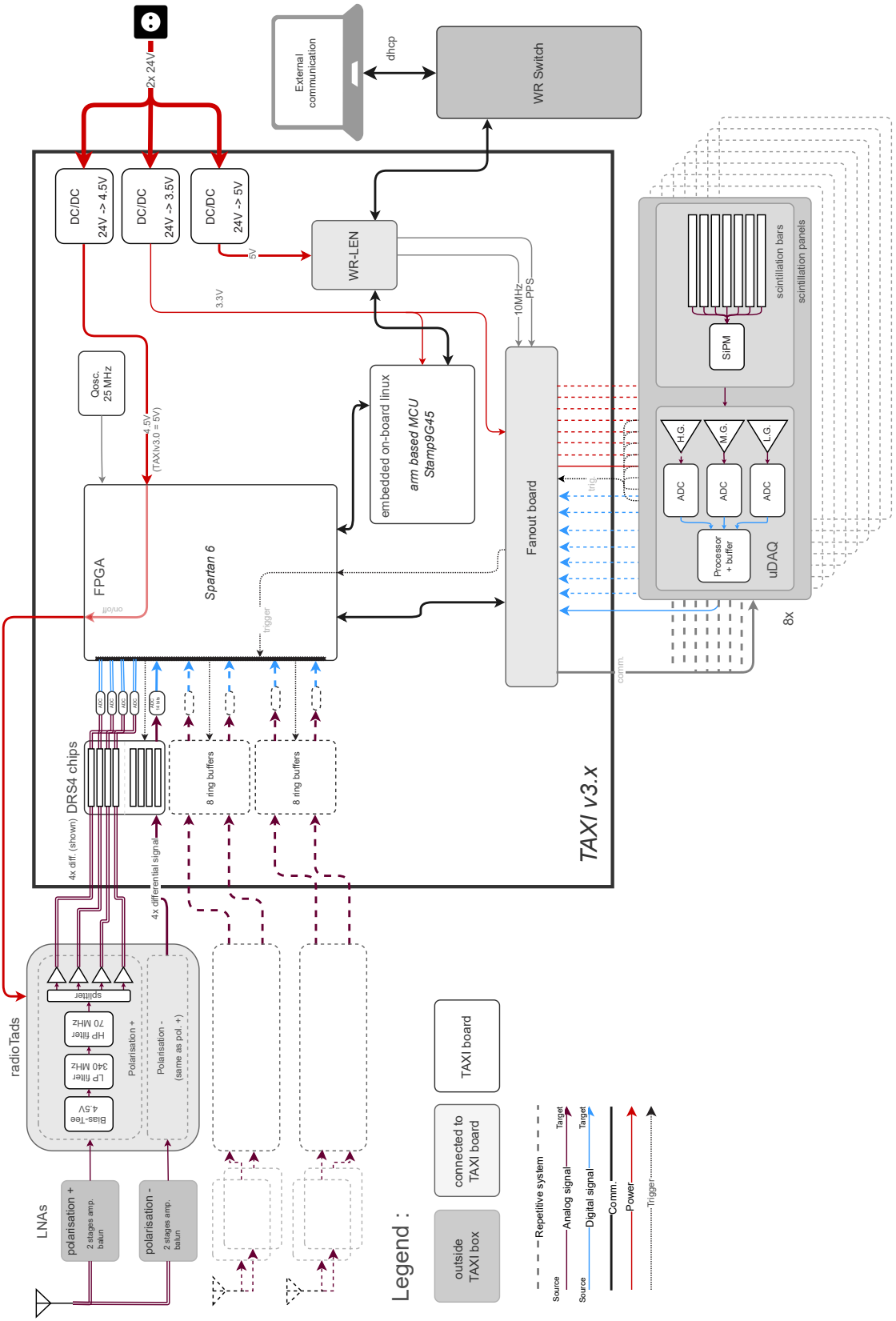


Figure 6.1: Block diagram of the electronic components of the surface enhancement.

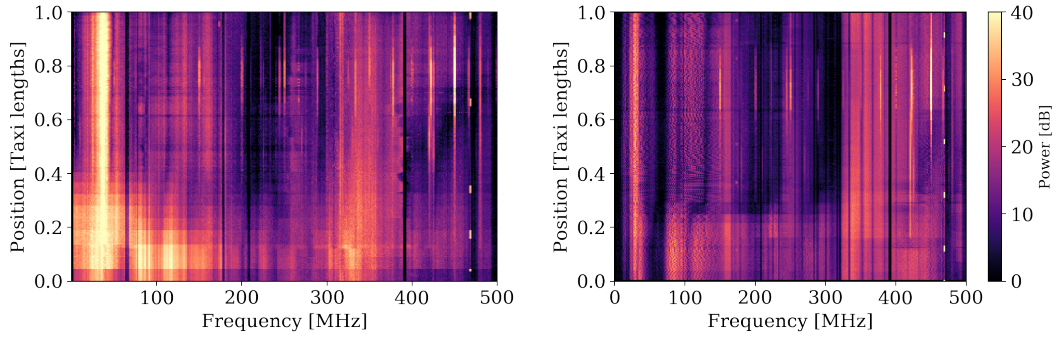


Figure 6.2: Waterfall spectrum between 0 to 500 MHz of the electromagnetic emission from the TAXI v3.0 (left) and the TAXI v3.2 (right). Plots by H. Dujmovic

second, a shower can be detected as soon as it reaches above the noise level, and third the amplitude of the radio signal scales linearly with the energy of the air shower. And the following approximations: the noise level is at 1 mV (which is close to the root mean square (RMS) of the modeled noise discussed in the next chapters), the minimum air shower energy detectable with the radio array is 10 PeV (it is more in the few tens of PeVs [121]), and the system amplification after the LNA is 0 dB. In that case, the maximum energy that the air shower can reach before saturation is 10 EeV on a system with a dynamic range of 1 V. In regard to the aimed amplification for the radio array of 2-3 dB, this is about 70-80% that value.

6.2.2 Parasite radio emissions

It was noticed on the field and in the commissioning data (see Chapter 8) that there are strong emissions collected by the antennas from approximately 70 to 100 MHz. Based on the results of the field measurements in 2020, it is believed to be coming from the TAXI v3.0 system [103]. To understand where exactly the emissions originate from on the TAXI v3.0 board a test is set up.

A spectrum analyser (SA) N9340B by Agilent Technologies with a telescopic antenna is slowly and constantly swept across the TAXI board along the longest side at approximately 1 cm above the board. The SA is put in dynamic spectrum (waterfall) mode with high sensitivity on. The schematics of the experimental setup can be found in Appendix C

Figure 6.2 presents the results of this test for the TAXI v3.0 board on the left and the TAXI v3.2 board on the right. The waterfall mode records the spectrum continuously, the color represents the strength of the emission in the Fourier domain, the x-axis the frequency, and the y-axis the elapsed time. However, since the antenna of the SA is moved at a constant rate across the board, the relative position of the antenna above the TAXI board replaces the time on the y-axis. The relative position 0 is where the DC-to-DC converters are positioned on the TAXI board.

Caution must be exercised when analyzing these plots because the omnipresent background noise also contributes to the spectrum seen by the antenna of the SA. To reduce the effect of the background, the measurement of TAXI v3.0 and TAXI v3.2 are done in the same room. The major observation from these plots is that a strong emission around 100 MHz and a lesser one around 320 MHz is emitted by the TAXI v3.0 at the position of the DC-to-DC converters whereas for TAXI v3.2 that emission is reduced. The constant strong emission in the TAXI v3.0 at approximately 50 MHz is caused by the power supply of the SA.

6.2.3 TAXI v3.2

Improvement of the TAXI board in collaboration with the electronic institute of KIT ensued where the principal objective was the reduction of the emitted noise. To achieve that, all the DC-to-DC converters are shielded, unused components are removed, the whole board is re-routed, some components are

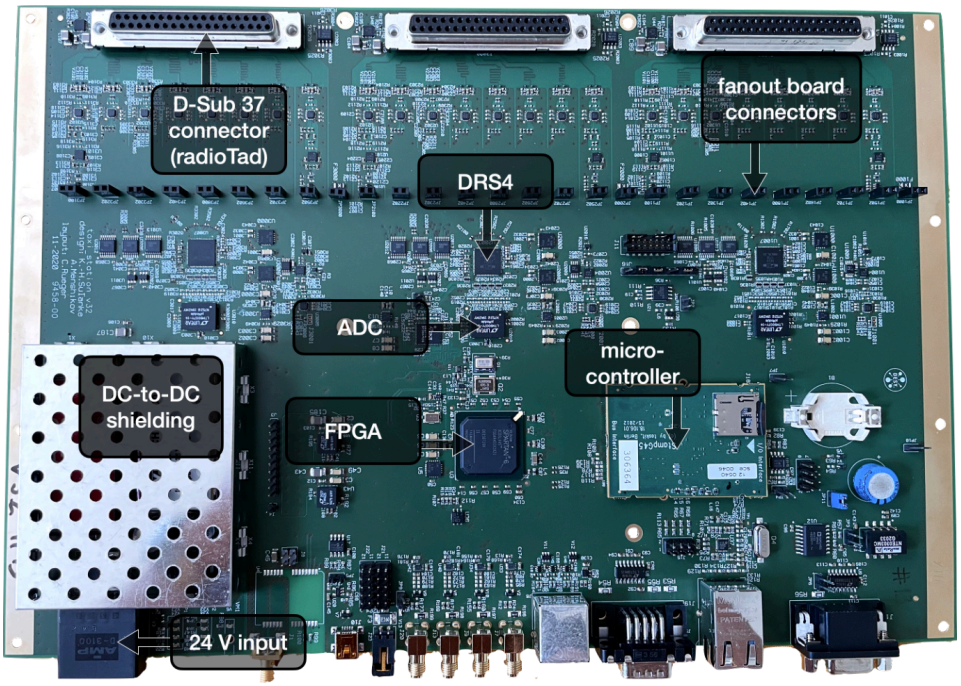


Figure 6.3: Photo of the TAXI board v3.2.

changed, the voltage dividers of offset voltages of the DRS4 chips are modified, the impedance of the routing to the DRS4 is reduced to approximately 30 ohm and other smaller changes[2] were done. Moreover, the voltage layer of the radioTad and by extension the power supply for the LNAs of the antennas is modified from 4.0 V to 4.5 V. This change results from tests of the radioTad, which are explained in Section 6.5. Figure 6.3 shows the TAXI board v3.2 with its principal components labeled.

6.2.4 Configurations of radio data taking

The primary configurable options for the radio data taking with the TAXI system are the cascading mode, the triggering mode, and the Serializer Deserializer (SerDes) delay. The understanding of the Region-of-Interest (ROI) and the `ROI_value` is also important at this point. Each concept necessary for the configuration of the radio-reading mode, and for the treatment of the raw data is explained in the following paragraphs.

Cascading modes: Each DRS4 has 8 channels, and both polarisations of one antenna are treated by one chip. The signal of the first polarisation of the antenna is duplicated four times in the radioTad before being fed in the first four channels of the DRS4. Similarly, the second polarisation signal is duplicated and fed in the next four channels of the chip. This duplication of the signal enables three types of cascading strategies: cascaded, semi-cascaded, and non-cascaded. In the first case, the four channels record sequentially, this results in one waveform of 4096 ns per polarisation. In the second case, semi-cascaded, channels 1-2 and 3-4 record sequentially, and the pairs record concurrently. Effectively, this gives two copies of a 2048 ns waveform. In the last case, non-cascaded, all four channels record concurrently, delivering four redundant waveforms of 1024 ns each. Figure 6.4 illustrate the four different modes of cascading. Note that for the cascaded and semi-cascaded mode, the ROI's cursor can be anywhere in the four or two channels respectively.

Serializer Deserializer (SerDes) delay: This command sets the relative time between the trigger signal and the read-out of the radio waveform, hereby, positioning the pulse of the air shower signal in the waveform. The SerDes delay unit equals one tick of the FPGA clock, which runs at 119 MHz, i.e. one SerDes delay unit equals 8.403 ns. The SerDes delay count starts at the end of the waveform, as

[2] Private communication with A. Menshikov on 16.07.2020

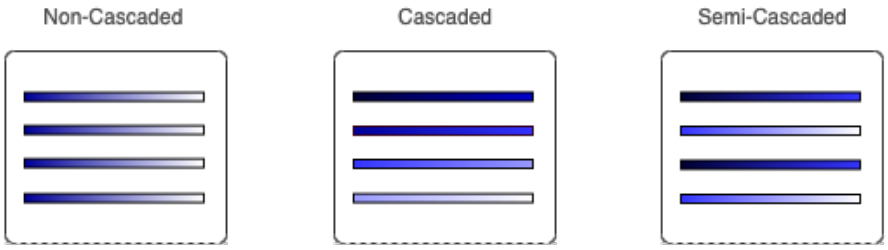


Figure 6.4: Visual representation of the cascading modes inside the DRS4 chip for one polarisation of one antenna.

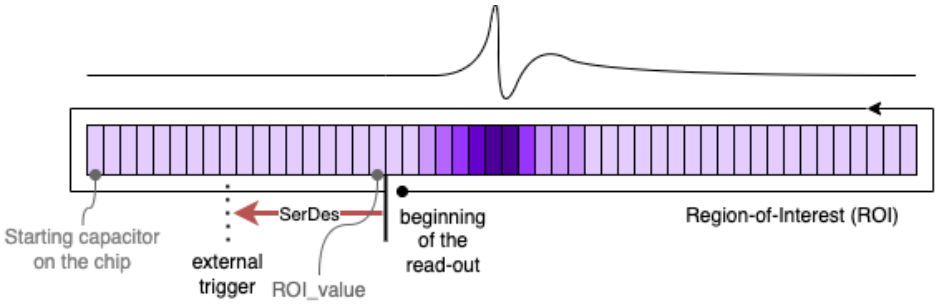


Figure 6.5: Visual representation of the ring buffer of one DRS4 chip with the description of the ROI, the ROI_value, and SerDes delay.

shown in Figure 6.5, meaning that for a high SerDes delay value, the pulse will be close to the beginning of the waveform and vice-versa.

Region-of-Interest (ROI). The ROI is defined as the Region-of-Interest read-out mode in the DRS4 datasheet [120]. The length of the ROI is programmed in the TAXI firmware as being equal to the total length of the waveform, regardless of the cascading mode, e.g., for the cascaded mode the ROI is 4096 and for the non-cascaded mode it is 1024. In the data, a `ROI_value` expressing the end of the readout is sent out by the DAQ. The FPGA, however, already rearranges the waveform for each channel according to the `ROI_value` which necessitates post-processing for the cascaded and semi-cascaded waveforms, as detailed in the following section. Each DRS4 chip has its own `ROI_value`, thus for each station, there are three `ROI_value` associated with each recorded event.

Triggering modes: The radio read-out can be requested in two different ways:

- **Soft trigger**: It is software initiated triggering mode and it is used for background measurements and most of the laboratory measurements. The FPGA requests a read-out every Δt_{soft} seconds. The most used setting for the soft trigger is every ~37 s.

- **Scint. trigger**: The read-out of the radio data is requested when a certain number of scintillation detectors reach a certain threshold within a coincidence interval of time. The threshold, the number of panels, as well as the time interval are modifiable via the TAXI software. The data triggered in that fashion potentially contain air-shower signals, but the signal in the antennas is not necessarily large enough to be usable. For the prototype station the default trigger is 6 panels reaching 4095 ADC counts within 1 µs

6.2.5 Low-level processing of the radio data

The TAXI board communicates to the user via a dynamic host configuration protocol (DHCP) interface. An `event_sender` on the TAXI side sends the data as a binary file, while an `event_receiver` on the computer or server side receives the packets. A conversion from binary to .npz (a compressed file

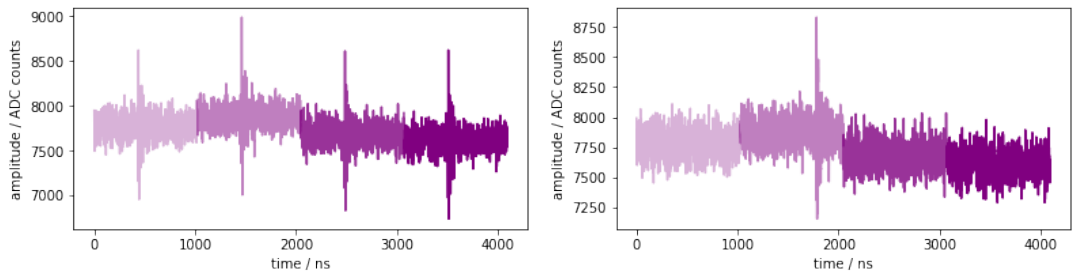


Figure 6.6: Examples of raw waveforms from the TAXI: on the left, the waveform was taken in non-cascading mode with a SerDes delay of 50 and, on the right, in cascading mode with a SerDes delay of 256. The different intensities of purple color represent the four physical channels on the DRS4 chip.

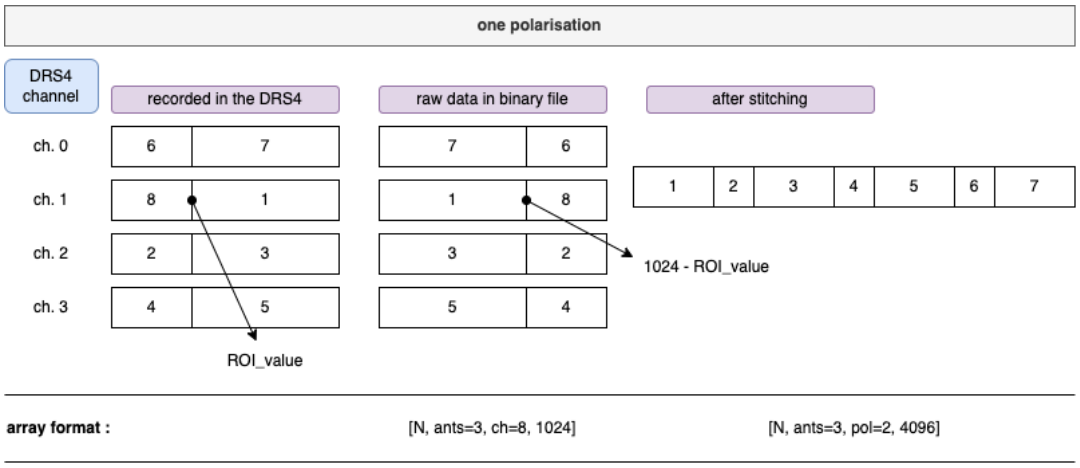


Figure 6.7: schematic of the filling of the DRS4 channels in cascading mode and the stitching process.

format from *NumPy* [122]) or to .i3 (the specific file format of IceTray) is possible with the `taxi_reader` module in the IceTray framework [98]. The basic lab measurements are done with a simple conversion to .npz, instead of the more processed .i3 files, to capture most of the behavior of the raw data. An example of a raw waveform from the TAXI only converted from binary to .npz with a pulse is shown in Figure 6.6. In this figure, on the left, the data is taken in non-cascading mode and on the right in cascading mode. The difference of color saturation in the waveform represents the four physical channels on the DRS4 chip.

As mentioned earlier, the FPGA already rearranges internally the data in each of the DRS4 channels, creating waveform, recorded in the non-cascading mode, in the correct order. However, this poses a problem in cascading mode, as represented in Figure 6.7, because the waveform is scrambled if the channels are simply stitched one after the other to create the 4096 ns long waveform. While processing the data to create the .i3 file, each of the physical channel is separated in two according to `1024 - ROI_value` and then stitched back together in the right order.

In order to calibrate the SerDes delay and to verify the stitching method, an experimental test was set up. A function generator SDG6052X by Siglent sends a pulse simultaneously to one radioTad and one uDAQ, both connected to the TAXI board. The radio read-out request is sent by the uDAQ when it receives the pulse, mimicking a scint. trigger from an air shower. All the cables are the same length in order to remove artifact delays from the cables. Once more, the detailed setup can be found in D.

The results of the test, shown in Figure 6.8, are in accordance with the expectations. The slope of the pulse time position as a function of the SerDes delay is, within the uncertainty, of one FPGA clock tick. The intercept represents the delay caused by the travel of the signal through the DAQ. As anticipated, the non-cascaded data has the same results whether the waveform is stitched or not. For the cascaded

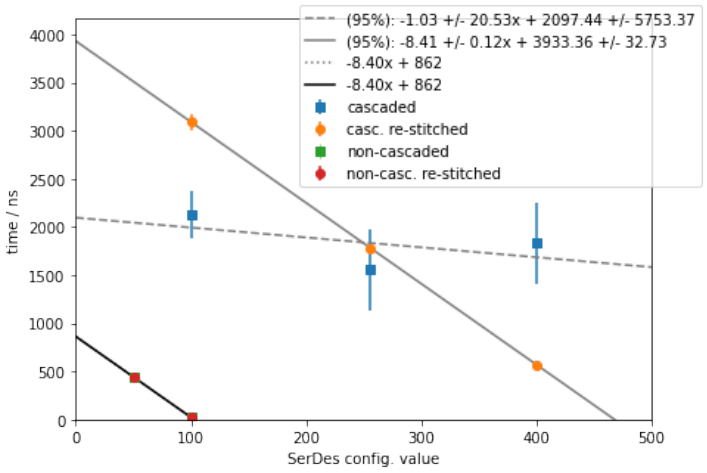


Figure 6.8: Calibration of the position of the pulse in the waveform according to the configuration of the SerDes delay option. Cascaded and non-cascaded data represent the raw data from the TAXI and the re-stitched data show the data after the stitching process. Hardly visible on this plot, the green non-cascaded points are exactly below the red non-cascaded restitched points.

data the story is different, the raw data lead to inconsistent results and large uncertainties. This is due to the pulse position moving in the waveform because of the scrambling of the channels as previously explained. When the data is re-stitched properly the data points follow a linear function with a slope of 8.4 ns per SerDes unit as expected.

These functions are used for controlling the read-out configuration for the prototype station as well as, in combination with the cable delays and the expected incoming direction, for the creation of a search window for air-shower pulses. This search of coincident air showers is, however, not covered in this work, and more details can be found in [123].

6.3 Scintillation detectors

The scintillation detectors are made of plastic scintillation bar connected through wavelength-shifting fibers to a silicon photomultiplier (SiPM) for the conversion of photons to an electrical signal. The fibers are routed in such a way that a certain particle crossing a certain length at any location on the panel will deliver approximately the same response. Each of the panels has a 1.5 m^2 total sensitive area. When a certain number of scintillation detectors in one station have a signal over a certain threshold value within a certain time window, a trigger signal is sent to the DAQ for the read-out of the DRS4 buffers. The most common operating mode for the prototype station is the requirement of 6 panels over 4096 ADC within 1 µs. The scintillation detectors have a custom electronic board with a microprocessor incorporated, located directly in the panel. This board, named uDAQ, receives the signal from the SiPM and amplifies it through three different amplification channels: low gain, medium gain, and high gain. This allows a wide dynamic range of about 1150 MIP The factor between the gain channels is approximately 6.5 between the high gain and medium gain channels, and approximately 77 between the high gain and the low gain channel. The signal is shaped and manipulated to obtain a measure of the charge deposited in the detector and then digitized before being sent to the TAXI board. A temperature sensor is located close to the SiPM records the temperature and sends the information to the uDAQ, which then transfers it via the TAXI.

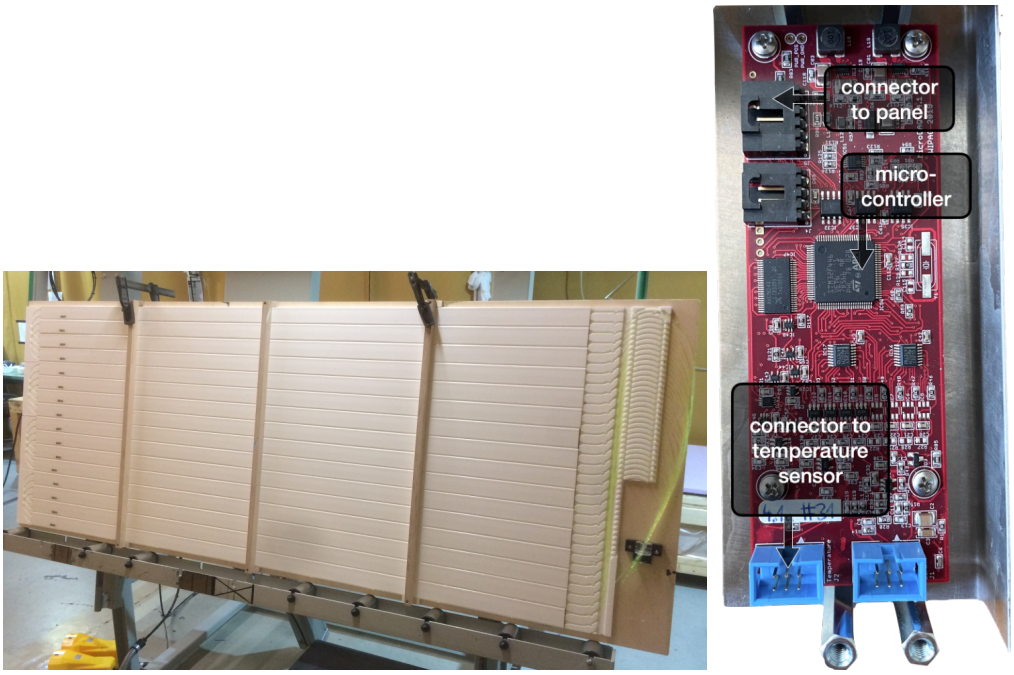


Figure 6.9: Pictures of a scintillation detector. On the left, the inside of the panel with the scintillation bars and the wavelength-shifting fibers and on the right the uDAQ.

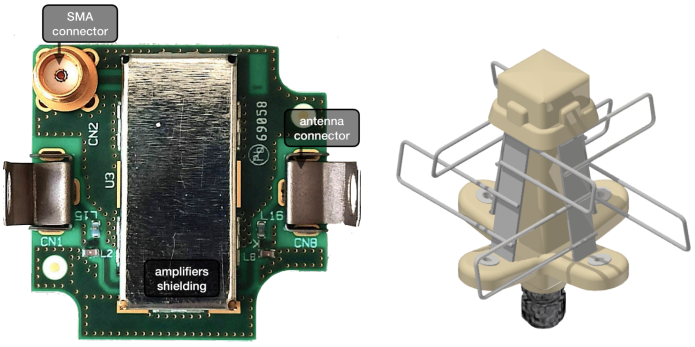


Figure 6.10: On the left, a picture of one LNA. On the left, a CAD drawing of the top part of the antennas, called trumpet. The two LNAs of the antennas are located inside the plastic hat of the trumpet.

6.4 Low Noise Amplifier (LNA)

Each of the antenna polarisation channels comes with a low noise amplifier located at the top of the antenna. The LNA is located as close as possible to the antenna to reduce picking-up noise. For the SKALA-v2, they are directly connected to the two arms belonging to the polarisation channel. The LNA pre-amplifies the radio signal of about 40 dB and converts the balanced signal from the antenna to a single-ended signal. Its noise figure is as low as 40 K, which lies below the expected galactic and thermal noise [124]. Unlike the functioning of the antenna in SKA, where a converter to optical signal is used at the output of the LNA, here a coaxial cable transmits the signal from the LNA to the TAXI box. Thereby, enabling the powering of the LNA together with receiving the signal in a single cable. A photo of an LNA and the top antenna part, the trumpet, where it is located, is shown in Figure 6.10.

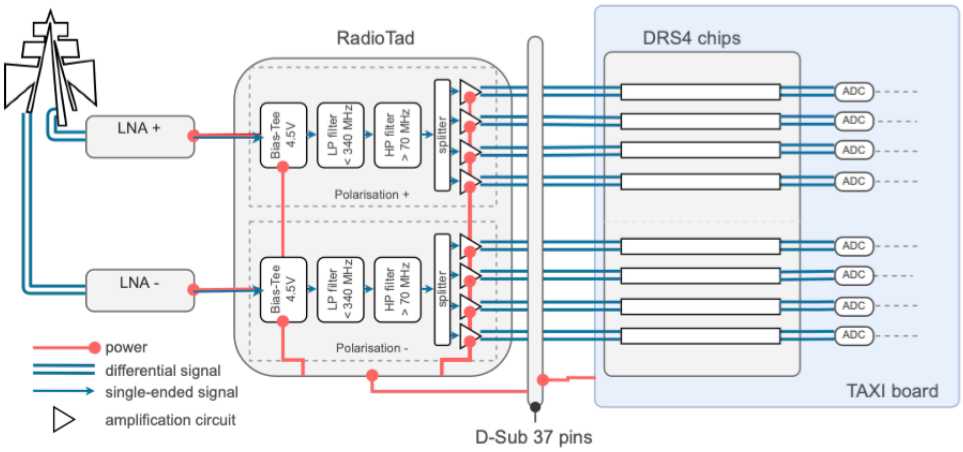


Figure 6.11: General schematics of one radioTad connected to the DRS4 chip. The splitter component was added in the radioTad v2, in the first version the signal was directly fanned-out in the PCB.

6.5 Front-end radio board – The radioTad

The radioTads are based on the first radio front-end presented in [125]. That first front end was composed of two boards, one for the DC-to-DC converters and one for the processing of the radio signal. The DC-to-DC converters are positioned on a different board to reduce the possible contamination of the emitted noise in the processing of the signal. The radioTads, on the other hand, are designed only to process the signal, as the DC-to-DC converters are incorporated in the TAXI v3.0.

In addition to the general requirements stated at the beginning of this chapter, the specific requirements of the radioTad include the filtering of the single-ended signal coming from the LNAs to the desired frequency band, powering the LNAs, and delivering four times per polarisation an identical differential signal to the DRS4 chips.

To accomplish that, the board is made of two identical parts, one for each polarisation channel of one antenna. There is one board per antenna to facilitate the eventual replacement if a component is faulty. Each polarisation channel features a low-pass filter ULP-340+ and a high-pass filter SXHP-48+ to obtain a nominal frequency band of 70 to 350 MHz. A bias-tee JEBT-4R2G+ (in v1) is used to relay the power to the LNA while receiving the signal through the same cable. The signal is then fanned-out and fed into four parallel amplifiers THS4508 from Texas Instrument, which amplifies the final signal, after the losses on the radioTad due mostly to the splitting, by roughly 2-3 dB and converts it to differential pairs. Each antenna has its own radioTad to facilitate eventual replacements. The printed circuit board (PCB) of the radioTad is composed of 8 layers, and to reduce the possible cross-talk, a ground layer is located between each layer with a signal.

The output connection between each radioTad and the TAXI board is achieved via a *D-Sub 37 pins* connector. The inputs to the radioTads are two *SMA* connectors, one for each polarisation. However, once the radioTads are installed inside the TAXI box, the connection to the TAXI box is made via two *N-type* connectors. To avoid RFI contamination in the radioTad from the other electronics in the TAXI box, they are covered with a custom made aluminum shielding, see Appendix D for the technical drawing. The schematic of this board is shown in Figure 6.11.

The radioTad outputs 8 differential signals, each differential signal has two polarities: N and P, that are identical except being out-phased by 180°: four differential signals belong to one polarisation and the next four to the other polarisation. The naming of the channels is tabulated in Table 6.1.

Table 6.1: Polarisations, channels, and polarity of the radioTad

polarisation	channel	polarity	polarisation	channel	polarity
minus (-)	11	N P	plus (+)	21	N P
	12	N P		22	N P
	13	N P		23	N P
	14	N P		24	N P

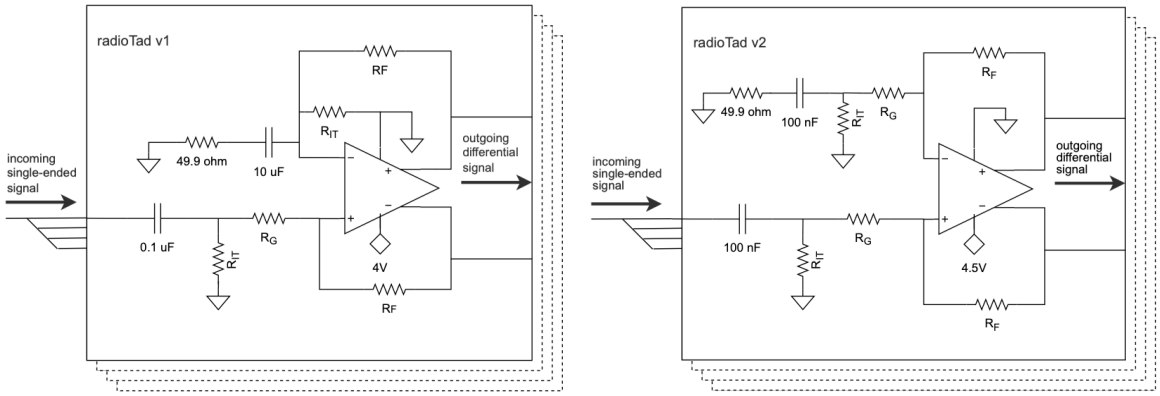


Figure 6.12: Amplification circuit on the radioTad v1 on the left and on the radioTad v2 on the right.

6.5.1 Modifications radioTad v2

On the radioTad v1, the gain of the channels after the calibration tests, discussed in Chapter 7, showed a variation between the channels above the requirement of 10%. The timing between the channels and the flatness of the gain had also room for improvement. Furthermore, the radioTad v1 calibration with the TAXI v3.0 presented strong attenuation in the high frequencies from around 230 MHz and above. All these measurements are explained in the next chapter, but they instigated the redesign of the radioTad that is described here.

To reduce the impedance mismatch at the fan-out point where the lines are split and the signal is diverted to the four amplifiers, the radioTad v2 includes a splitter SCA-4-10+ from Mini-Circuits at the front of the amplifiers in order to obtain lines at 50 Ω of impedance. The bias-tee is changed from JEBT-4R2G+ to TCBT-2R5G+ both from Mini-Circuits to lower the cost and to obtain space on the board. As mentioned earlier, the input power of the radioTad is changed from 4.0 V to 4.5 V to be closer to the nominal power of the amplifiers and to avoid clipping in the gain curve. The amplification circuit is slightly modified to correspond to the test circuits in the amplifier data-sheet. The amplifier circuit of v1 is shown on the left side of Figure 6.12 and the circuit of v2 on the right side. In the radioTad v1, the R_g resistance in the negative loop of the amplification circuit is missing, which led to some last-minute modifications of the radioTad v1 before deployment to obtain a gain in the region of interest. the resistance R_G was changed from 100 Ohm to 33 Ohm. The other resistance R_{IT} and R_F are equal to 68.9 ohm and 348 ohm respectively.

The gain circuit of the radioTad v2 is designed conformally to the datasheet to obtain an impedance of 50 Ω and a gain of 14 dB. The high gain is there to compensate for the ∼7 dB loss at the splitter, and the attenuation from the load at the output of the differential amplifiers. All the striplines of

Table 6.2: Resistors of the amplification circuit for the radioTad v2

R_{IT}	R_G	R_F
88.7 (± 0.1) Ω	56.2 (± 0.1) Ω	348 (± 0.1) Ω

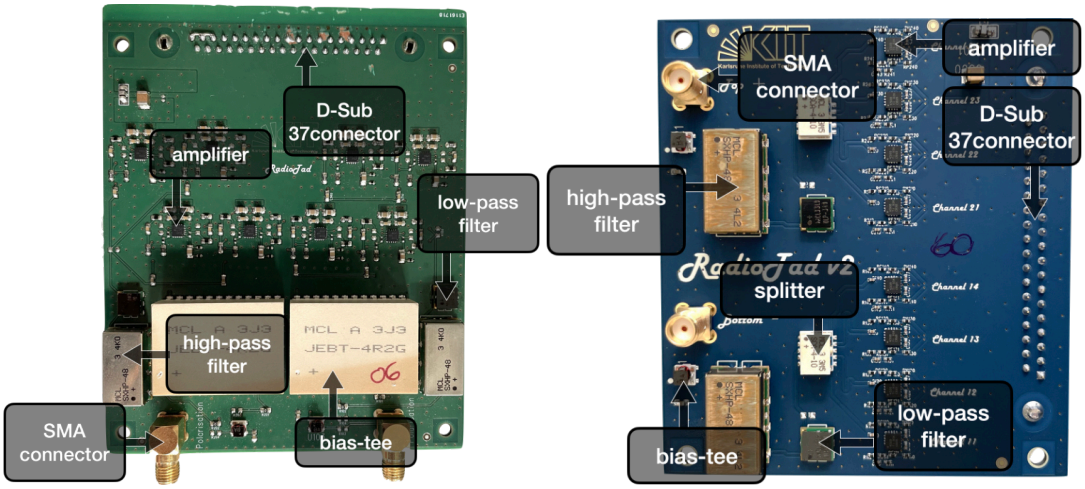


Figure 6.13: Picture of the populated radioTads. On the left in green is the first version v1 and on the right in blue is the second version v1.

the PCB after the splitter are routed to be the same length in order to deliver the same signal in the four channels of the DRS4 without phase mismatch. The resistors of the radioTad v2 have the values tabulated in Table 6.2

Moreover, the material and the size of the striplines are changed to improve the impedance of the board. The PCBs of the radioTad v2 are made of material designed for radio frequencies, I-Tera [126], which is rated to have a stable dielectric constant and dissipation from -55°C to 125°C, and for frequencies up to 20 GHz. Pictures of the physical radioTad v1 (left) and v2 (right) are shown in Figure 6.13. The layout of the PCB of the two versions of the radioTad can be found in Appendix E.

6.6 Cables

The LNAs are powered and transmit signals over coaxial copper cables. A 1.5 m LMR-240 cable [127] is connected via an SMA connector to the LNA and hangs from the antenna center. A 50 m LMR-400 cable [128] is attached via N-type connectors to the LMR-240 cable at one end and to the central DAQ at the other end. The 50 m cable includes 5 m slack on each end, sufficient for four extensions of the mount, for a total lifetime of around 25 years (approximately 5 years initially, and 5-6 years per extension) with the antennas cable slack. Eventually, to increase the lifetime, extension cables could be added to the surface detectors. The cables were carefully chosen to survive the cold weather as well as deliver only small attenuation of the signal over the 51.5 m of propagation length.

CHAPTER 7

Calibration of the Radio Electronics Signal Chain

As the aim in accuracy for the energy measurements of cosmic rays is 10%, and it scales approximately linearly with the radio signal amplitude, the radio signal chain calibration must be at least 10%. Thus, the calibration of the different components is important. Now that the electronics of the data acquisition system have been previously explained, this chapter describes the calibration and the characterization of those electronics related to the radio signal chain. First, the calibration of the LNAs is explained, then the one of the radioTads. Finally, The chapter concludes with the combined calibration of the TAXI and radioTad.

7.1 Gain calculation

Before detailing the experimental setups and results, some basic equations must be introduced. The gain of a device under test (DUT) is defined in voltage G_V as

$$G_V = 20 \cdot \log\left(\frac{V_{out}}{V_{in}}\right) \tag{7.1}$$

which in power G_P gives

$$G_P = 10 \cdot \log\left(\frac{P_{out}}{P_{in}}\right) \tag{7.2}$$

Where V_{out} (P_{out}) and V_{in} (P_{in}) are the voltage (power) send and received, respectively. The gain is in units of [dB] in both cases. To compare measurements of different channels or at different temperatures, the ratio of amplitudes are calculated by converting the gain obtained by the **VA!** into an amplitude with $A = 10^{(G_V/20)}$. It is also possible with these equations to calculate the equivalent requirement in gain for a 10% uncertainty in amplitude. This is equal to a maximum uncertainty of 0.8 dB independent of the total amplification.

7.2 Calibration of the LNAs

7.2.1 Gain measurements

One reference LNA with the serial number w10718, is tested with the vector network analyser (VNA) FieldFox N9923A from Keysight in the climate chamber VT7021 from Vötsch at KIT from -70ºC to 30ºC. The usage of a 2-way 180º splitter ZFSCJ-2-1-S+ from Mini-Circuits is used to simulate the balanced input signal that the LNAs receives from the antennas. A custom board is designed and produced to accommodate the LNA. Its only purpose is to transform the bracket connectors of the LNA to SMA connectors where the VNA can be easily connected. In general terms, the VNA sends a single-ended signal to the 2-way 180º splitter which converts it into a balanced signal. The signal is fed to the LNA through the connector board, the LNA then delivers a single-ended signal back to the antenna. The

Table 7.1: Summary of LNA w10718 temperature measurements

freq.	amplitude ratio		gain difference	
[MHz]	min – max	std dev.	min – max [dB]	std dev. [dB]
	all measurements			
100	0.92 – 1.04	4.0%	-0.73 – 0.31	0.35
300	0.91 – 1.07	5.0%	-0.78 – 0.58	0.44
	low temperature measurements			
100	0.99 – 1.01	0.40%	-0.06 – 0.04	0.034
300	0.96 – 1.03	1.8%	-0.29 – 0.23	0.16

transmission (S_{21}) and the reflection (S_{11}) coefficients are recorded by the VNA. The schematics of the setup and a picture of the connector board can be found in Appendix C, the PCB layout of the LNA connector board in Appendix E.

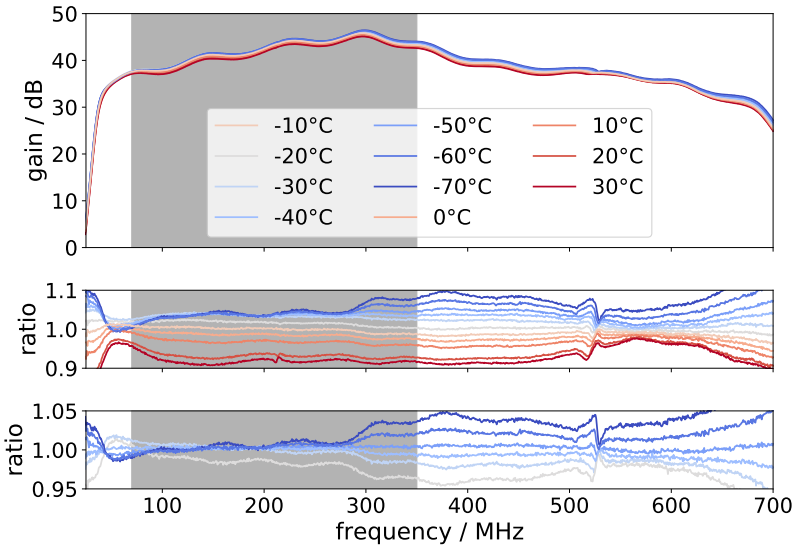


Figure 7.1: Top: Gain of the LNA at different temperatures in the climate chamber. Middle: amplitude ratio of all the measurements over their total average. Bottom: amplitude ratio for measurement between -70°C and -20°C over their average.

The results of these measurements are shown in Figure 7.1. The top plot shows the total gain of the LNA W10718 at different temperatures, the middle plot shows the amplitude ratio of the measurements at each temperature to the average of all the measurements, and the bottom plot shows the amplitude ratio only for the measurements between -70°C and -20°C. At the deployment site, the outside temperature varies between -70°C and -20°C (see Figure 5.6), therefore the bottom plot is more representative of the variation of the LNA gain during a year.

For the measurements including all the temperatures, the maximum influence of the temperature is slightly below ±10%, and ±5% in standard deviation, in the frequency range of interest. For the measurements showing only the temperature expected at the South Pole, the variation is even smaller, with a maximum of less than ±5%, and a standard deviation of approximately ±2% with the highest variation happening at the highest frequencies of the band. The gain variation at specific frequency, 100 MHz and 300 MHz, are summarized in Table 7.1

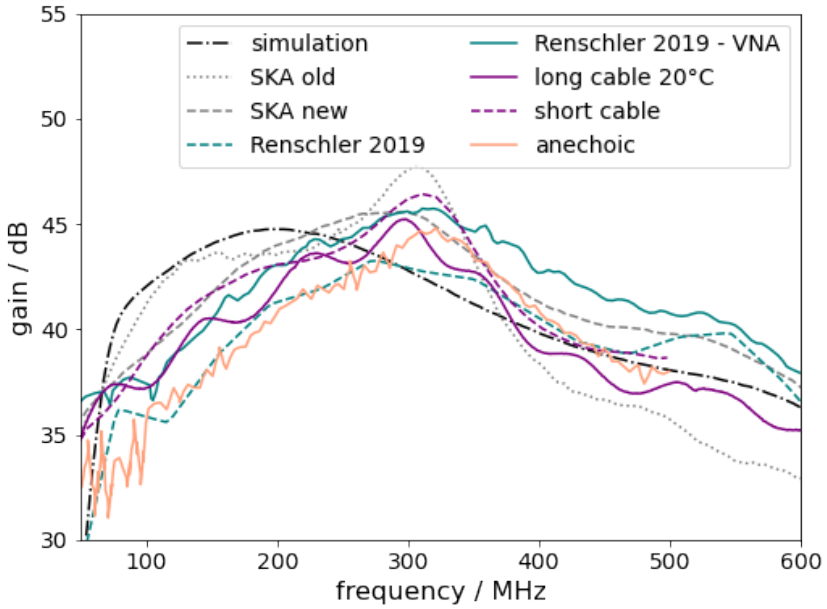


Figure 7.2: Compilation of the LNA measurements and one simulation. The measurements in purple represent the measurements done with the VNA and the connector board described in this section, the measurements in cyan are from M. Renschler's thesis [102], where the dashed line is obtained from a pulse and the solid line from the same VNA. The simulation and the SKA measurements are supplied by E. de Lera Acedo and obtained using a similar setup as the one in this work.

The top plot of Figure 7.1 shows the absolute gain, yet, measuring it for the LNA is hard. The measurements are extremely sensitive to the setup and most likely also to the background. The setup is imprecise because of mainly two aspects: one, the connector board has some custom-made metal connectors which makes it impossible without big uncertainties to evaluate the transmission and reflection of the board and remove it from the measurements; two, because the impedance of the antenna changes with frequency, which is not taken into account in the VNA measurements.

The challenge of calculating the absolute gain is visible in Figure 7.2. The results of measurements done by different people using different setups, and a gain simulation are summarized in that figure. There is a large deviation between them, at the level of ~5 dB. Furthermore, the experimental measurements differ, in shape and in intensity, from the simulation, especially in the frequency band of interest.

Nonetheless, all the measurements agree on a resonance at roughly 310 MHz, whereas the simulation shows no sign of that increase in gain. The oscillations in the measurements are most probably coming from an impedance mismatch in the experimental setup.

7.2.2 LNA-to-LNA comparison

The total gain of eight LNAs is measured on the same day using the same setup to estimate the variation between them. The setup comprises a VNA, and the connector board, as described for the temperature setup. Only the LNAs are at room temperature during this test.

The resulting gain of the eight LNAs is shown in the top plot of Figure 7.3. The bottom plot is the difference between the gain of each LNA to the average of all. The total standard deviation between 50 and 370 MHz is 0.24 dB or 2.8%, with a larger deviation of around 0.5 dB around 300 MHz. This larger deviation is located around the same frequencies as the resonance peak of the gain that was observed in Figure 7.2.

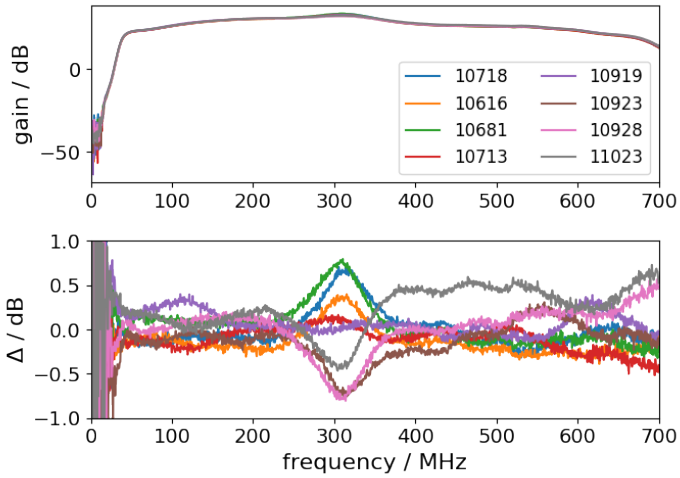


Figure 7.3: Gain of eight different LNAs. The top plot shows the total gain of each LNA and the bottom plot, the difference in gain between each LNA and the average measurement. The serial number of each LNA is indicated in the legend.

7.2.3 Intrinsic Noise measurements

TAXI v3.0 feeds the LNAs of the antennas with 4.0 V via the radioTads, simultaneously powering the amplifiers of the radioTads. The nominal voltage of the radioTad amplifiers is 5 V according to the datasheet [129]. During the calibration of the radioTads, it was noted that clipping happens above 0 dBm at 4.0 V, as demonstrated later. Ideally, the radioTad would be powered with 5 V, however, a higher power might increase the noise in the LNA. Therefore, the following study was conducted to see if the increase in power to the LNA increases its noise.

To do so, the LNA is connected to an oscilloscope and powered with different voltages. The LNA is shielded with two Faraday cages to reduce the environmental noise and its antenna connectors are left floating. The experiment is conducted with several LNAs at different times and locations. The details of these measurements are explained in detail in D. Rech's bachelor thesis [130].

Figure 7.4 shows the results of the noise floor measurements. It is unclear from those results if the noise floor of the LNA is reached or if there is still background emission disturbing the measurements, leaning for the second case due to an observation of decreasing noise when measured in a different room (the blue point on the figure). The conclusion is limited at: for an environmental (or) intrinsic noise floor of the LNA at approximately 2 mV, the input power to the LNA has no significant contribution. In view of these results, it was decided to choose a conservative value of 4.5 V for the TAXI v3.2 system, which is the power level at which the radioTad does not exhibit clippings.

7.3 Calibration of radioTads

7.3.1 Channel-wise calibration

One radioTad v1 is tested alongside one radioTad v2 with the same experimental methodology. The results are also shown for radioTad v1 and radioTad v2 side-by-side for direct comparison. As references, for the radioTad v1, one of the spare boards from the 2020 deployment is used. For the radioTad v2, the board numbered 06 from the production batch is taken.

The first measurement, done at room temperature, is to quantify the similarity between the 8 differential channels of one radioTad. The radioTads are tested with the VNA FieldFox N9923A from Keysight and a custom-made test board that is designed to convert the *D-Sub 37* output of the radio board to a *SMA* output for each polarity of each channel described earlier in Table 7.2. The measurement

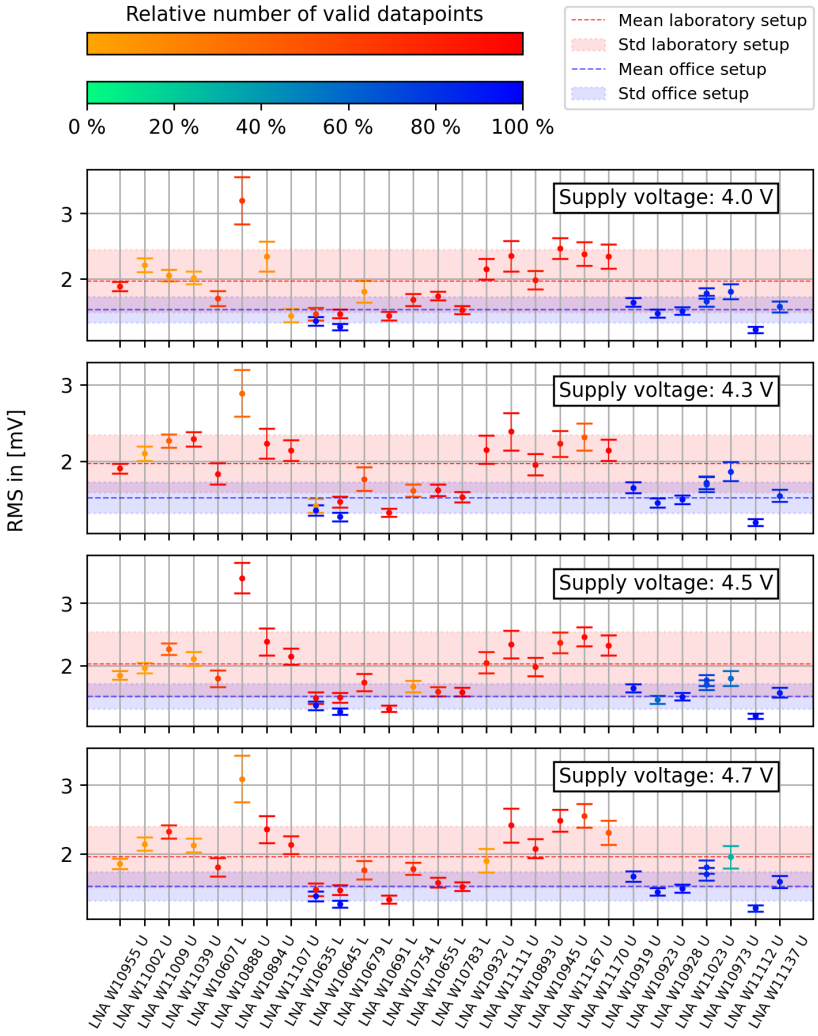


Figure 7.4: Noise level of several LNAs of the same type identified by their serial number. The points represent the average value for one set of measurements and the bars are the standard deviation. The red and orange points were sampled in the laboratory and the blue point in an office [130].

is done on one channel at a time and the not-connected output SMA connectors are terminated with 50 Ω . The test board is connected to a power supply and powers the radio board via some pins in the *D-Sub 37* connectors. A DC blocker BLK-89-S+ from Mini-Circuits is inserted at the input of the radioTad to avoid the transfer of the 4 V power by the bias-tee to the VNA. For these measurements, the power supply is set to 4.5 V, and the dBm power from the VNA is at 0 dBm. The test board and the detailed schematics of the test setup can be seen in Appendix C and the electronic layout of the test board in Appendix E. The gain and phase of the transmission parameter, S_{21}, are taken individually for each polarity of each channel for version 1 and version 2 of the radioTad.

The results of the channel-gain measurements are shown in Figure 7.5 with the radioTad v1 on the left and the radioTad v2 on the right. For both figures, the top plot represents the total gain of each polarity of each channel as a function of the frequency and the bottom plot shows the amplitude ratio of these compared to the average over all the measurements. Similarly, the channel-phase measurements are illustrated in Figure 7.6 with the radioTad v1 on the left and the radioTad v2 on the right. The top plot is the phase over frequency in each polarisation of each channel, the middle plot is the phase unwrapped starting after the high-pass filter region, and the bottom plot is the difference of the phase to the average phase in degree. For the bottom plots, 180° is added to the N-polarities. To calculate the

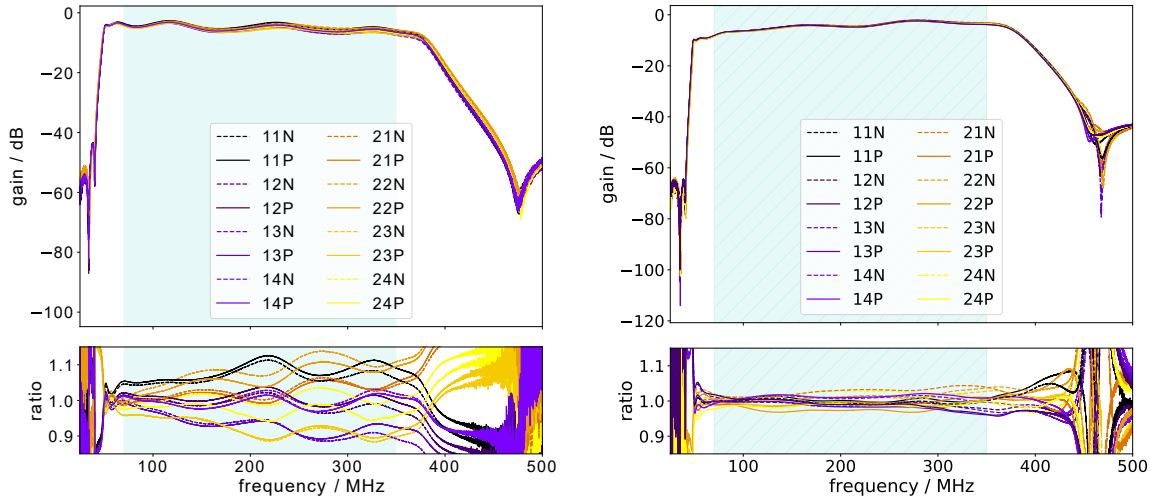


Figure 7.5: Gain of each polarity of each channel of the reference radioTad v1 on the left and of the reference radioTad v2 on the right. The upper plots represent the gain of each polarity-channels, and the lower plots is the ratio of each of those polarity-channels over their average. The blue colors would represent all the channels of one polarisation of one antenna, and the orange colors the other polarisation. The pale blue band is a visual help highlighting the nominal frequency band.

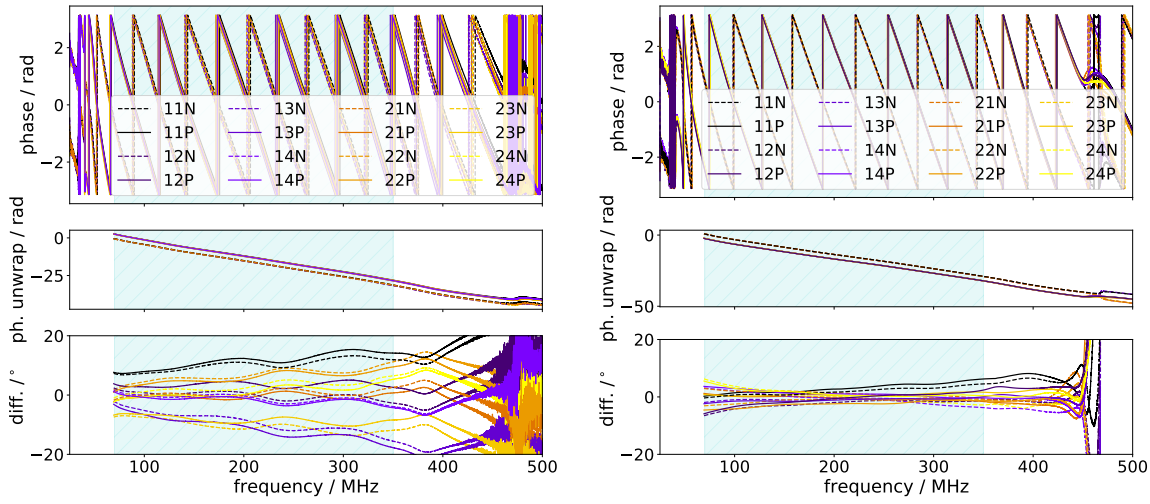


Figure 7.6: Phase of each polarity of each channel of the reference radioTad v1 on the left and for the reference radioTad v2 on the right. The top plot represents the phase of each polarity-channels, the middle plot is the phase unwrapped to be continuous and the bottom plot is the difference of each of those polarity-channels over their average. For the bottom plot, 180° is added to the N polarity. The blue colors would represent all the channels of one polarisation of one antenna, and the orange colors the other polarisation. The pale blue band is a visual help highlighting the nominal frequency band.

time delay in Table 7.2, the formula:

$$\Delta t/\text{ns} = \frac{\Delta\varphi/^{\text{o}}}{360}\,\frac{1}{f/\text{GHz}} \tag{7.3}$$

is used, where Δt is the time delay, $\Delta\varphi$ is the phase difference in degree and f is the frequency.

Channels radioTad v1. The maximum difference between the channels in the channel-gain plot is above ±10%, the total difference at 100 MHz in gain between the two extreme values is 0.86 dB, and worsens with higher frequency and reaches a total difference of 1.99 dB at 300 MHz. The polarisation of the antenna, containing the channels starting with 1, oscillates inversely with frequency than the other polarisation, containing the channels starting with 2, of the antenna. This behavior can be best observed in the ratio plot. The standard deviation at 300 MHz, is around ±7% for the radioTad v1.

On the channel-phase plot, on the top plot, one can observe an offset of approximately 180º between the phases of the two polarities, which is what is expected from differential signals. The same behavior can be noticed as well on the unwrapped phase, where the polarity N is constantly offset from the polarisation P, as well as on the bottom plot where 180º is removed from the N-polarities, and both signals are within the same region. In fact, from the bottom plot, if one observes the two polarities P and N of one single channel, let's say channel 21, the N polarity (dashed) closely follows the behavior of its different pair P (solid).

A similar inverse oscillation between the two antenna polarisation channels, in blue colors and orange colors in the figure, can be seen on the channel-phase plot as well as for the channel-gain plot. The difference between the two most far phase curves at 100 MHz is equal to 22º, which translates to a shift of 0.62 ns between the identical channels. The phase difference gets worst with increasing frequency and reaches a total of 33º at 300 MHz which causes a difference between the channels of 0.31 ns. The time difference decreases, or the phase difference increases, at higher frequencies due to the larger sensitivity of higher frequencies to a time difference. In other words, the phase shift for a constant *Deltat* is proportional to the frequency.

Channels radioTad v2. One can see quickly by eye that the differences between the channels of the radioTad v2 for the gain as well as for the phase are substantially reduced compared to the first version. The difference between channels from the gain is much smaller than ±10% at every frequency in the nominal bandwidth. At 100 MHz the total difference of the two extreme channels in gain is 0.25 dB, which also increases with the frequency to reach, at 300 MHz, a value of 0.52 dB. The standard deviation at 300 MHz, is around ±1.7% for the radioTad v2. This corresponds to a 400% improvement within the one radioTad channels.

For the channel-phase measurements, the maximum difference in time between the channels is reduced to 0.13 ns at 100 MHz and to 0.10 ns at 300 MHz. No obvious difference between the two polarisations of the radioTad v2 can be seen in the channel-phase as well as the channels-gain plots, except for the expected 180º phase offset between the N and P polarities. Table 7.2 summarizes the differences in the gain and phase for the radioTads v1 and v2 at 100 MHz and 300 MHz.

From the same measurements, the total gain of the radioTads is obtained by numerically adding the N and P polarities together for each channels individually. In order to add the gains of the polarities together, they are first transformed into amplitudes and summed up. The resulting amplitude is subsequently transformed back into a gain. The total gain is shown in Figure 7.7, where the lines represent the average from all the 8 channels and the band represents the maximum and minimum from all these channels. Once more one can see that the uncertainty due to channel-to-channel variations is reduced from the first version to the second version. Both versions have a comparable amplification of around 2-3 dB, where the second version amplifies more at the higher frequencies and vice-versa for the first version. Because of the oscillations, the standard definition of cutoffs at -3 dB from the total gain is hardly applicable. The cutoff here will then be defined as when the curve crosses a gain of -5 dB, which is illustrated by a gray band on the figure. In both radioTads, the cut-offs from the high-pass filter and the low-pass filter can be seen. For the radioTad v2 this corresponds to a frequency band between 48 MHz and 394 MHz.

Table 7.2: Summary of the channels differences of the references radioTad v1 and radioTad v2

radioTad v1				
freq.	amplitude ratio		gain difference	
[MHz]	min – max	std deviation [%]	min – max [dB]	std deviation [dB]
100	0.95 – 1.05	2.87	-0.41 – 0.46	0.25
300	0.89 – 1.12	7.04	-0.98 – 1.00	0.61
	phase difference		time delay	
	min – max [°]	std deviation [°]	min – max [ns]	std deviation [ns]
100	-10.7 – 11.3	6.3	-0.30 – 0.31	0.18
300	-19.4 – 13.7	10.1	-0.18 – 0.13	0.094
radioTad v2				
freq.	amplitude ratio		gain difference	
[MHz]	min – max	std deviation [%]	min – max [dB]	std deviation [dB]
100	0.98 – 1.01	0.82	-0.15 – 0.097	0.071
300	0.98 – 1.04	1.7	-0.21 – 0.32	0.15
	phase difference		time delay	
	min – max [°]	std deviation [°]	min – max [ns]	std deviation [ns]
100	-2.70 – 1.87	1.50	-0.075 – 0.052	0.042
300	-4.11 – 6.70	2.82	-0.038 – 0.062	0.026

It is interesting to note that both boards exhibit a gain curve oscillation with frequency, which can hint at an impedance mismatch. However, the oscillation pattern differs from v1 to v2. The impedance mismatched is present in all version of the radioTads as well as their predecessor from P. Steinmüller and M. Renschler [102], [131]. The three boards have different oscillation patterns in their gain curve. It is unclear at this point where the mismatch is happening, but the differences in pattern indicate that the mismatch occurs for different components in each version. The test board used for the calibration measurements is also inducing a possible impedance mismatch and uncertainties, as well as the connectors, especially the *D-Sub 37* one. A thorough investigation of the s_{11} and s_{21} parameters for all components was done to find the cause of the impedance mismatch. A flatter gain curve in the desired frequency band was obtained by using the component individually. Unfortunately, that result was not translatable yet to the more complex board with seven layers and eight amplifiers, where both aspects can induce complex cross-talk.

The measurements explained above are done for each of the 60 radioTads v2 produced for the first phase of deployment of the SAE, to identify if the curve behaves as expected. If a problem is identified in the gain curve, the board is put aside. Afterward, the calibration of the radioTads for one station is done with the TAXI board, explained later.

7.3.2 Temperature calibration

The same setup as described above, but with the radioTad v2 and the test board located inside the climate chamber is used to measure the influence of the temperature on the gain of the radioTad. The measurements are conducted at temperatures from -60°C to 60°C in steps of 20°C, with the reference radioTad v2. Only the results of channel 21N are shown in the plots of this section to disentangle the influence of the channel-to-channel variation from the temperature effect. The total variation including temperature influence and channel deviation can be found in Appendix C.

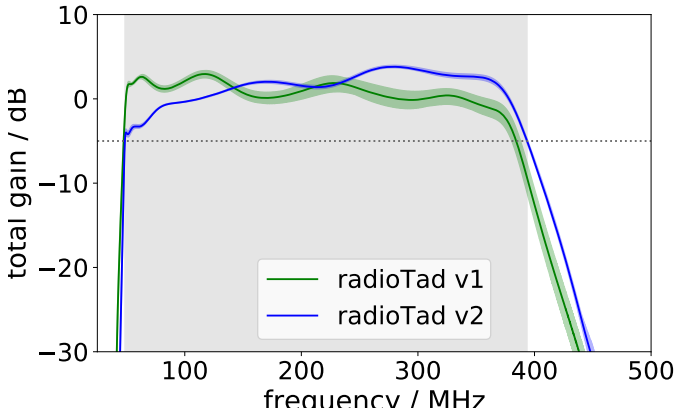


Figure 7.7: Total gain of the reference boards obtained by the numerical addition of the differential pairs for both radioTads. The band represent the maximum and minimum from the different channels of the boards and the line is the average of those channels. The gray band represents the section where the total gain is above a threshold of -5 dB.

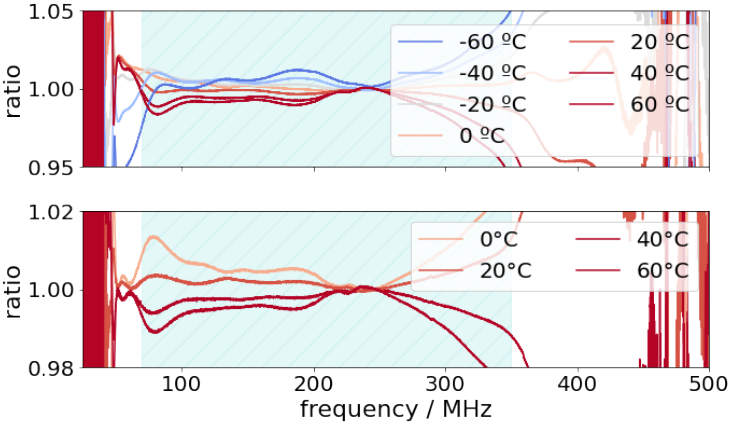


Figure 7.8: Amplitude ratio of the gain of channel 21N of the reference radioTad v2 at each temperature over the average of the measurements shown. The top plot shows the measurements at temperatures between -60ºC and 60ºC, and the bottom plot, only the measurements of the expected operating temperature of the radioTad. The pale blue band is a visual help highlighting the nominal frequency band.

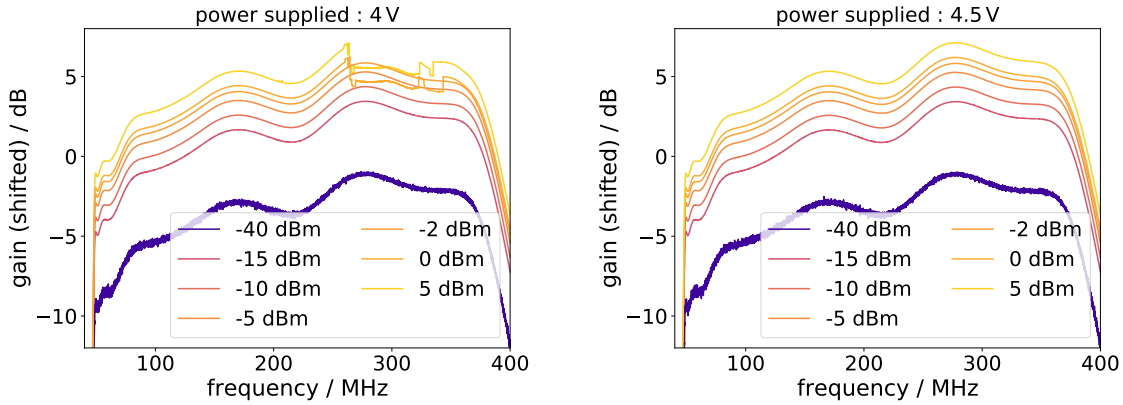


Figure 7.9: Gain curve of one radioTad v2 at different output power from the vector network analyser. The left plot is with a supplied voltage of 4.0 V and the right plot is for 4.5 V. The curves are shifted by 10 times the VNA output power for better visibility.

Figure 7.8 illustrates the ratio of the amplitude of measurements at different temperatures compared to the average of these measurements. In the top plot, the temperature is from -60°C to 60°C whereas on the bottom the temperature is from 0°C to 60°, the expected operating temperature range in the fieldhub.

The influence of the temperature on the gain becomes more important at frequencies above roughly 300 MHz. Below that, the influence on the gain is small. For all temperatures, with frequencies between 50 and 370 MHz, the total standard deviation is 1%, and for frequencies between 50 and 300 MHz it reduces to ±0.5%. When looking at the operational temperatures, the total standard deviation for the nominal frequency band is ±0.6%, and for the frequency band 50-300 MHz, at ±0.3%. Overall the influence of the temperature on the radioTad v2 is well below ±10%

7.3.3 Clipping

It was noticed during the gain calibration that clipping of the gain curve occurs at high test port output power when the radioTad is supplied with 4 V instead of 4.5 V. This parameter is a measure of the power transmitted by the VNA, and is calculated via

$$G/\mathrm{dBm} = 10 \cdot \log_{10} \frac{P}{1\ \mathrm{mW}} \tag{7.4}$$

where P is the power transmitted by the VNA. Using a setup similar to the previous one, the gain of one channel is recorded and the output power of the VNA is variated from -40 dBm to 5 dBm. The measurements are made with a voltage of 4 and 4.5 V supplied to the reference radioTad. The gain curves in Figure 7.9 are shifted accordingly to 10 times their dBm values for better clarity of the clipping. The results for a supplied voltage of 4 V is shown on the left, and the results for 4.5 V are on the right. In the 4.0 V plot, clippings appear at values of 0 dBm and above, for frequencies higher than ~250 MHz. When powered with 4.5 V no clipping is visible in the whole range for the radioTad v2.

A careful reader may have noticed that the instrumentation uncertainty and the correction for the experimental setup are not explicitly removed from the plots presented above, except for the plot comparing the LNAs. The reason for this is twofold. Firstly, in both the LNA and radioTad experimental setups, an additional board that cannot be calibrated is needed due to the unpractical connectors. Therefore, it is more accurate to state that "all plots above still have the experimental setup response included" rather than saying that "all plots above have half the experimental setup removed". The latter statement may lead someone to wrongly assume that the absolute gain plot of these measurements can be taken. Additionally, the effect of the cables in these setups is relatively small

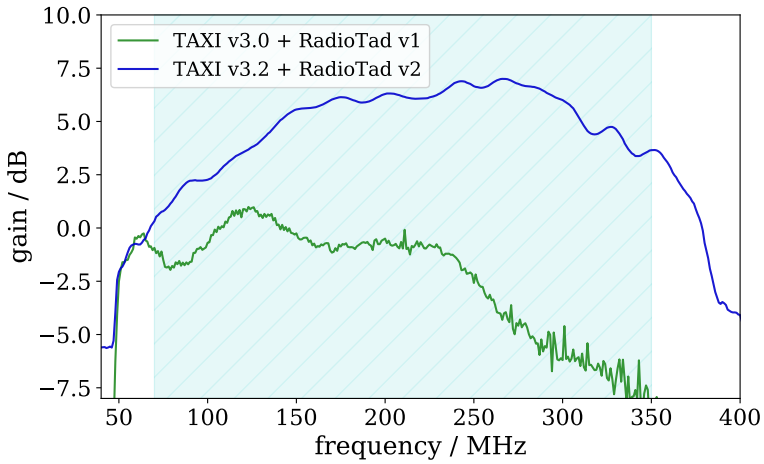


Figure 7.10: Total gain of the TAXI and the radioTad for TAXI v3.0 with radioTad v1, and TAXI v3.2 with radioTad v2.

compared to the response of the experimental boards, which are expected to contribute significantly more to the measurement uncertainty. Secondly, with the exception of the absolute gain of the LNA which is still being researched, all the measurements presented above are relative rather than absolute. This means that the temperature is compared to the temperature, channels are compared to other channels, and so on. For these relative measurements, the experimental setup response cancels out, indicating that the experimental setup is not influenced by the device under test. For example, during temperature tests, all devices are placed outside the climate chamber, with only the DUT inside. In the case of the electronic response of the radio chain, the radioTad and the TAXI are calibrated together, eliminating the need for any connector board and providing an accurate gain response of the radio system. More details on these measurements are explained in the next section.

7.4 Joint calibration of radioTad and TAXI

To minimize uncertainties and avoid the need for intermediate boards, the radioTad and TAXI are calibrated together. Otherwise, both the TAXI and radioTad would require a connecting board, which can introduce additional sources of uncertainty.

To calibrate the two devices together, three radioTad are connected to one TAXI board. A function generator SDG6052C by Siglent is connected to a 6-way splitter ZBSC-615+, from Mini-Circuits, which is then connected to six cables of equal length each attached to a DC-blocker BLK-89-S+ also from Mini-Circuits. The response of the cables, the splitter and the DC-blocker all together was measured for the 6 polarisation channels. It is then removed numerically from the gain plots individually for each of the antenna polarisation channels. The function generator sends a chirp pulse from 1 µHz to 450 MHz for 500 s of 300 mVpp. During the time that the chirp is sent, the TAXI records many waveforms in soft trigger mode (see Section 6.2.4). The frequency associated with the part of the chirp pulse that is recorded in one waveform is inferred from a sinus fit to the waveform.

The measurements in Figure 7.10 are conducted using this methodology at room temperature. Once with the TAXI v3.0 and the radioTad v1 and once with the TAXI v3.2 and the radioTad v2. From this figure, one can see that the total amplification of the TAXI v3.2 and the radioTad v2 is higher than the previous generation and lies between 0 and 7 dB in the band of interest, with a roughly constant 6 dB amplification between 150-300 MHz. The first generation attenuates strongly the high-frequency part of the band, whereas the second generation has a reduced amplification in the lower frequencies of the band

All the TAXIs and radioTads are calibrated in this fashion prior to deployment and this is the calibration used in the analysis described in Chapter 10. In addition, this measurement is conducted

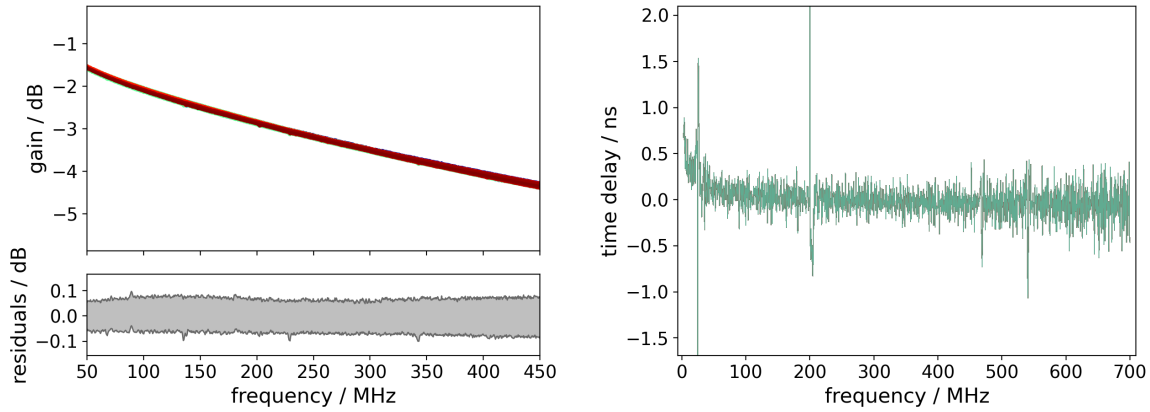


Figure 7.11: On the right, attenuation in function of the frequency of all the LMR-400 cables. On the left, group delay of all the cables relative to the reference cable.

for each TAXI in the climate chamber, where the temperature is varied from -70°C to 30°C in steps of 10°.

7.5 Cable calibration

Once more the VNA is used for the calibration of the LMR-400 cables, the longest one from the station, with a length of 50 m. Each cable is measured three times in a row, then a reference cable is measured. The attenuation of the cable is then averaged and normalized to its reference cable measurement. The addition of the reference cable to reduce the uncertainty from the VNA, it was noted afterward that the gain of the VNA change of about 0.15 dB when the working temperature of the VNA passes from the room temperature to 51°C, where the temperature as well as the gain stabilize. Two hundred cables were calibrated this way.

The left-hand side of Figure 7.11 shows the attenuation of the cable over the frequency on the top plot. The bottom plot represents the maximum and minimum values of all the cables. The attenuation at 350 MHz is around -3.8 dB, and exponentially smaller attenuation for lower frequencies. The standard deviation over all the cables and frequency is ±0.3%, or 0.03 dB. The 1.5 m LMR-240 is not yet calibrated for the complete deployment.

The group delay of the frequency relative to the reference cable is shown on the right-hand side of Figure 7.11. The standard deviation over all the cables and the frequencies is 0.30 ns. The sharp peaks in the group delay plots, e.g., at 20 MHz and 200 MHz are artifacts from the VNA.

7.6 Discussion

The calibration of the electronics needed for radio signal processing was detailed in this chapter. A clear improvement on the channel similarity is achieved with the radioTad v2 for the gain as well as for the phase. The channel-to-channel variations within one radioTad v2 reaches a maximum of 0.52,dB, with a standard deviation of ±1.7% at 300,MHz, which is where the difference between the channels is high. The LNA-to-LNA variation is in the order of 2.8%, which has a larger deviation around 300,MHZ.

The influence of the temperature for the LNAs is small, around ±2%, and almost negligible for the radioTads v2, where the standard deviation is roughly ±0.6%. Both for their nominal frequency band and their operational temperature.

From there, a rough estimate of the accuracy of the radio chain can be done. Assuming that the uncertainty adds linearly then

$$\sigma_{\text{partial}} = \sqrt{\sigma^2_{\text{LNA}} + \sigma^2_{\text{LNA-temp}} + \sigma^2_{\text{RT-channels}} + \sigma^2_{\text{RT-temp}} + \sigma_{\text{cable}}} \tag{7.5}$$

where σ_{cable} is the uncertainty on the cables, σ_{LNA} is the standard deviation between LNAs, $\sigma_{\text{LNA-temp}}$ the influence of the temperature on the LNA, $\sigma_{\text{RT-channels}}$ the standard deviation between the channels of the radioTad v2, and $\sigma_{\text{RT-temp}}$ the influence of the temperature on the radioTad v2. This results in a partial uncertainty of ±3.9%. Partial because the TAXI-radioTad calibration variation is not accounted for nor is the antenna's calibration. This leaves ∼9.2% for $\sqrt{\sigma^2_{\text{TAXI-RT}} + \sigma^2_{\text{Ant}}}$ in order to satisfy the total of ±10% accuracy of the system requirement. However, if this is not reached, a possible electronic response station-per-station in the processing of the data could be implemented.

Naturally, it is disappointing that the LNA measurements from different laboratory experimental setups do not agree. Nevertheless, the relative gain at different temperatures and the variations between different boards can be measured accurately without knowledge of the absolute gain, because the uncertainties from the instrumentation cancel. Further measurements that will be conducted in an anechoic chamber might be able to shed light on the absolute gain of the LNA as well as for the antenna. The calibration of the antennas was not discussed in this chapter, because calibrating an antenna is the work of a complete thesis. As discussed in Chapter 3, typically the antennas are calibrated on the field by either a calibrated pulser or by using the galactic radio emission, as in [132]–[134] for example. These two techniques also permit the calibration and cross-checking of the complete radio signal chain.

These three calibration methods, anechoic chamber, galactic background, and calibrated pulsers are being investigated at the moment. Until then, a simulated gain of the antenna response, shown in Chapter 5, is used for analyses. Another interesting measurement that will be conducted in the anechoic chamber is the LNA noise floor measurement. The radio silent chamber will be much more suitable to reach the noise floor than any other environment. Dry ice will be used to reduce the thermal noise.

On the radio board side, improvements could be achieved on the amplification circuit, for example, the addition of pull-down resistors for lower input voltage than the nominal 5 V [129] and by removing the oscillation in the gain curve due to the apparent impedance mismatch. Overall as long as the gain is in the wanted region, is constant between all the boards, and the channels have less than 10% difference between them, the radioTad meets the requirements. Nevertheless, the most important parameter to take into account is the similarity of the channels. Indeed, The total gain of the radioTad is in the desired range and the electronic response is taken into account in the post-processing of any physics analysis, and thus the shape is not crucial in the design. Ideally, the dependence on the temperature is low as well, although the temperature dependence can be corrected with the temperature sensor on the TAXI board.

Finally, the absolute gain of the TAXI v3.2 and the radioTad together are more uniform between the 70-350 MHz frequency band than for the TAXI v3.0 and the radioTad v1, where a strong attenuation is present for frequencies above 250 MHz. Nevertheless, as it is shown with the calibration of the TAXI, the board as it complies with the requirements for the SAE. It is considered to incorporate the processing of the radio signal directly in the DAQ board for the surface enhancement of IceCube-Gen2, which was briefly discussed in Chapter 6.

The next chapter will concern the commissioning of the radio array of the prototype station of January 2020.

CHAPTER 8

Commissioning Data from the Prototype Station

This section describes the commissioning of the prototype station of January 2020, beautifully pictured in Figure 8.1. After designing and testing the hardware in the lab, the station was deployed at the site. After deployment, the data from the system were examined to confirm that the system works as intended. In this chapter, all measurements are done using the TAXI v3.0 and the radioTad v1 from the initial prototype station. This chapter does not cover the replacement of the TAXI box at the South Pole in January 2022, whose new version contains the TAXI v3.2 and the radioTad v2.

The commissioning is composed of two parts, an observational study regarding the behavior of the radio signal chain in the system and a search for air showers. First, the acquisition of the raw data is briefly explained and inspected. Subsequently, the observation of the Galactic radio noise is shown using two different methods, one with in the time domain and the other in the frequency domain. Finally, an example of an air shower measured with the prototype stations is shown, and quadruple coincidences between the radio antennas, scintillation detector, the IceTop tanks and the in-ice detector are presented.

8.1 Raw data acquisition

The communication to the station at the South Pole is achieved via satellite. There are three satellites alternating for a total coverage per day of approximately 15 hours. The time of accessibility of the satellites changes everyday[1]. The binary files obtained as output by the DAQ system are transferred to a storage located in Madison, and then processed with the taxi-reader module to i3 files, the standard file

[1] The window of availablity of the satellite can be found here: https://live.icecube.wisc.edu/satellite/

Figure 8.1: Picture of the prototype station 2020 during the South Pole winter. Credit Yuya Makino.

Table 8.1: Summary of the used containers in the i3 files

key	subkey	object
I3EventHeader	runid	unique identifier for the 8 hour runs of IceCube
I3EventHeader	eventid	unique identifier of the event within the run
SurfaceFilters	condition_passed	flag for the type of trigger (soft or scint)
RadioTaxiTime		time when the TAXI received the trigger
RadioTraceLength		length of the trace (depends on cascading mode used)
RadioAntennaROI	antenna	roi value associated to one antenna
TAXIRadioWaveform	ant. & pol.	waveform for each channels (in ADC counts)
Laputop	direction & core	air shower reconstruction from IceTop
LaputopParams	logS125	energy estimator value from Laputop reconstruction
ScintRecoPulses	charge & time	values of the scintillation detectors

format of the IceCube software [98]. The processing pipeline of the data underwent multiple iterations since the station is installed, in this work the version V5 is used. In the i3 files, different containers are accessible via a key value are summarized in Table 8.1. These are all located Q-frames, which correspond, in the IceCube terminology, to DAQ-related information.

The definition of related terms such as trigger types, waveform stitching, and cascading mode can be reviewed in Section 6.2.4. The processing of the binary file into i3 files already includes a proper stitching of the traces. Therefore, the traces in the "TAXIRadioWaveform" container are assumed to be in the correct order. The amplitudes of the waveform in that container are given in ADC counts, where one ADC count is defined by the dynamic range of the ADC component on the TAXI board. This corresponds to $1\,[\mathrm{V}]/2^{14} \approx 61\,\mu\mathrm{V}$, assuming a linear behavior of the system until saturation. Information from the IceTop tanks, and data from the scintillation detectors, when available, are included in the processed i3 files as well.

Triggering flags are assigned to each event during the processing, but the triggering mode information is not transmitted yet by the DAQ system. A statistical method taking advantage of the constant Δt between soft triggers is applied. Eventually, the FPGA will be programmed to send out the trigger information alongside the waveform data directly in the binary files.

8.2 Basic checks

Once the data is processed to i3 files, several basic checks are performed on the radio measurements.

8.2.1 Waveforms from the TAXI

In a first step, raw waveforms are extracted from the "TAXIRadioWaveform" container and no further processing is applied. An example containing a few artifacts is chosen and exposed in Figure 8.2. For this event, the data was recorded on the 10^{th} January 2021 in non-cascaded and soft trigger mode. In the figure, from top to bottom, antenna 1, 2, and 3 are presented, each with both polarisation channels.

Due to the 14 bits ADC converters, the dynamic range of the system is 16,384 ADC counts. From the design of the firmware inside the TAXI, the position of the waveform should be located in the middle of this range (8,192 ADC counts) to allow for free oscillation of the amplitudes above and below the baseline. In this figure, one can see that the baseline is located around 7,500–7,800 ADC counts, slightly lower than the expectation. As a consequence, a small reduction in the dynamic range of the system is to be expected. The down oscillations will saturate faster, i.e. reach their maximum value. These clipped measurements are not trivial to use in analysis, because of the loss of information. As explained in Chapter 6, the DRS4 chips are composed of four physical channels per polarisation, and each of those channels have slightly different baselines. The different channels can be observed in

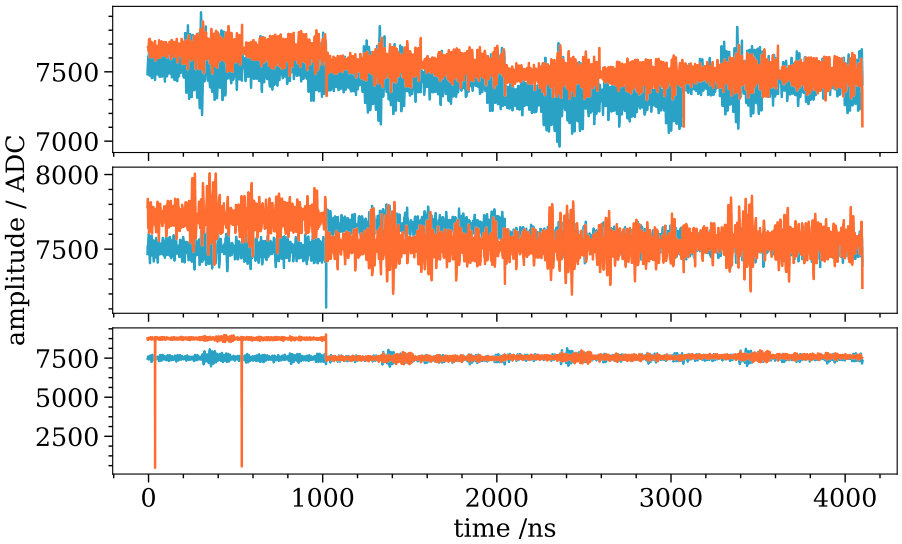


Figure 8.2: Non-cascaded waveforms of one soft triggered event. Top: antenna 1, middle: antenna 2, bottom: antenna 3, each with both polarisations. Note that although the data is recorded in non-cascaded mode, the waveforms from physical channels are put one after the other in the processing.

the waveforms, especially discernible in the first channel in antenna 2 and 3. Moreover, in the third antenna, there is the presence of two single bins with a value close to zero, which are called bin spikes. As the waveforms were recorded in non-cascading mode, the resulting waveforms from the 4 channels should be similar. The fact that the bin spikes occur in only one channel, and appear so prominently, speaks against RFI or any possible measured quantity as the root cause. A study, which can be found in Appendix G, was done on these bin spikes, and they are believed to arise from corruptions of byte registers. The most common effect is the flip of the most significant bit from a 1 to a 0, surprisingly the opposite, 0 to 1, is not observed in the data. In very corrupted files, a bit shift is also observed. The module "RemoveTAXIArtifacts", is designed from these observations and incorporated into radcube. The module corrects the values of the bin spikes to the true values, it works well if the corruptions in the file are not too numerous. The challenge within the module is the identification of bin spikes, and ensuring to not miscorrect a real pulse, and this reduces the efficiency when the number of corruptions is too severe. The number of bin spikes changes from TAXI to TAXI and was observed to be related to temperature, especially as the temperature of the board rises above 30°C. The bin spikes are the dominant undesirable feature in the TAXI waveforms, however, other particularities are seen in the data; please refer to [130] for a more detailed study of the TAXI v3.0 in a laboratory context.

8.2.2 Raw noise level

The noise level of the waveform was also examined. For this, the baseline of the "TAXIRadioWaveform" is moved to zero by filtering to the band between 1 and 500 MHz, and the artifact remover is run. The noise level is calculated with two different methods, one obtained by calculating the RMS (Equation (9.2)) between the bin 10 and 410 in each waveform, and the other by using the subtraces methods (explained in Section 9.2.3), with a subwindow of 64 bins. The data used was recorded in the period between the 1st December 2020 and the 31st December 2020 for a total of 35,292 waveforms. Note that the same set of measured noise is used in Chapter 9 and Chapter 10.

The resulting distributions for each channel are shown in Figure 8.3 where the filled histograms represent the noise level calculated using the subtraces method and the hollow histogram using the window method. The widths of the distributions are smaller with the subtraces method. Nevertheless,

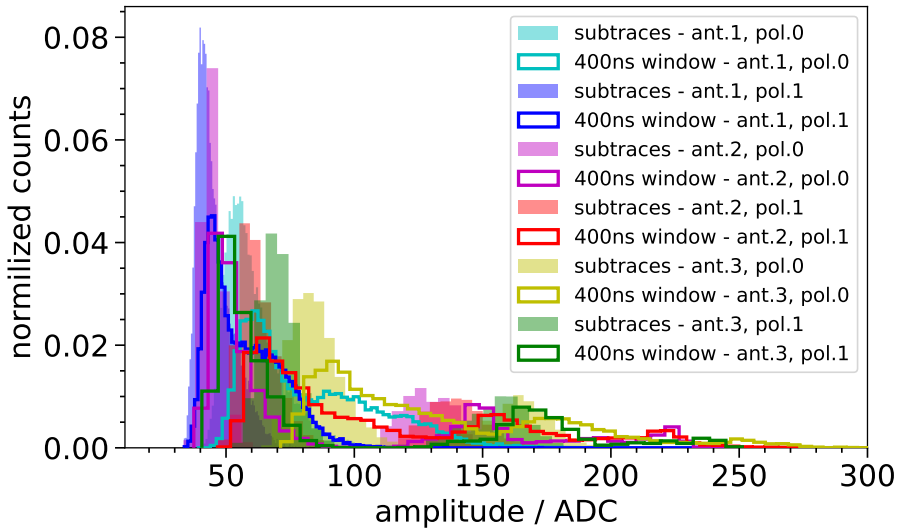


Figure 8.3: Distributions of the noise level in the traces for each antenna channel independently, calculated with the subtraces method in hollow histograms and with the standard RMS window in filled histograms.

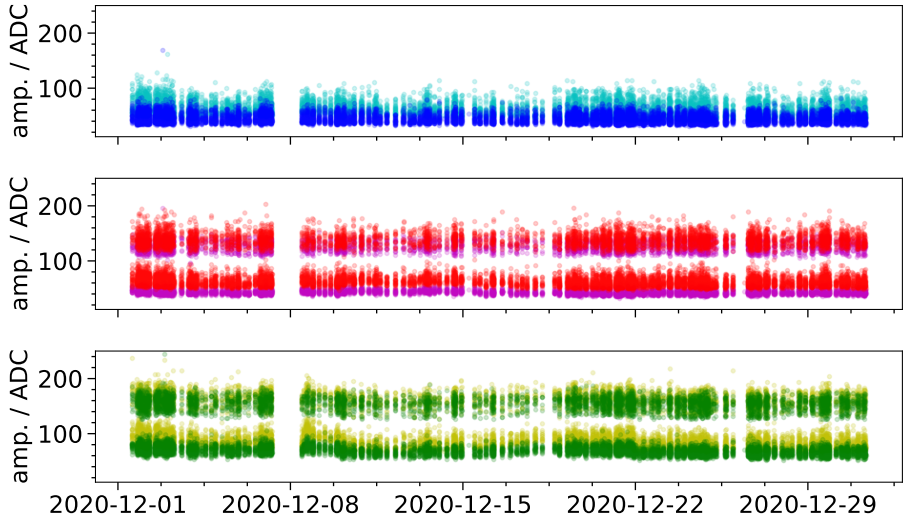


Figure 8.4: Noise level over one month of soft trigger data taking. From top to bottom: antenna 1, antenna 2, antenna 3, both polarisations in each. The x-axis is in the form year-month-day.

with both calculations a second distribution with a higher noise level is present. The average as well as the variation of the noise level is different from antenna channel to antenna channel.

Ideally, the noise level in the antenna would be constant, except for the effect of the galactic revolution that is shown later on. The fact that the subtraces method gives sharper distributions is a good indicator that possible RFIs are not accounted for in that noise estimation.

To investigate this second distribution, and examine the stability, the noise level is plotted against time. This is represented in Figure 8.4 where antenna 1 is plotted on top, antenna 2 in the middle, and antenna 3 on the bottom. The bimodal distribution of the noise is present in antenna 2 and antenna 3, constantly over time. In antenna 1, only the lower noise distribution appears. Outside the bimodal

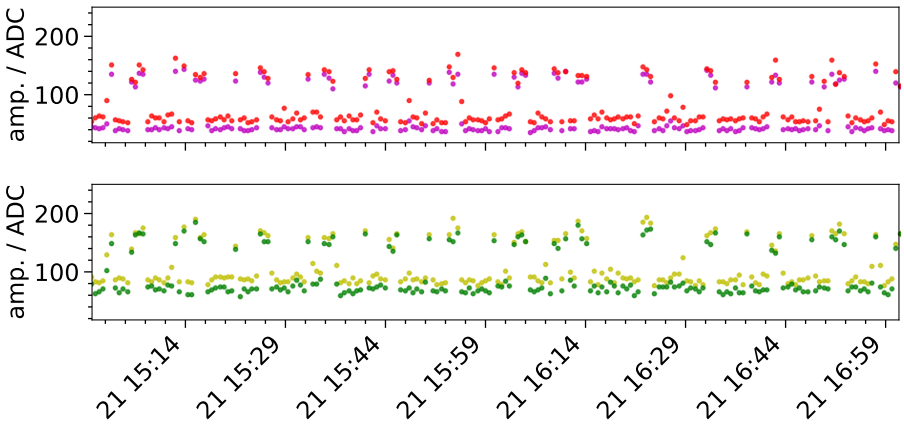


Figure 8.5: Close up on the noise level over one month of soft trigger data taking. On the top antenna 2 and on the bottom antenna 3, both polarisation each. The x-axis is in the format day hour:minute.

distribution, the noise level is relatively constant over time. From the positioning of the antennas in the prototype station, antenna 1 is the most southward antenna and the closest to the IceAct telescope and the fieldhub. This can be viewed in Figure 4.7.

From the previous plot, a close-up of 2 hours of operation time is presented in Figure 8.5 for antenna 2 on top and antenna 3 on bottom. The increase of the noise level happens periodically in both antennas simultaneously at roughly every 10 minutes and lasts about 2-3 minutes. The reason of the presence of the second distribution is, as of yet, unclear.

8.2.3 ROIs stability

An additional parameter to examine in the prototype station is the `ROI_value` (see Section 6.2.4), which is related to the position where the read-out of the waveform begins. Each of the DRS4 chips features its own `ROI_value`. The soft triggering here is a better parameter to observe the behavior of the `ROI_value`, due to its cyclic repetition. Therefore, a repetitive pattern in the `ROI_value` is expected. The data used was recorded on the 10^{th} January 2021 in non-cascaded mode, and the "RadioAntennaROI" key is used. Due to the specific cascading mode, the value of the `ROI_value` in this example should be comprised between 0 and 1023, which is to say, the position in the individual capacitor in one DRS4 chip where the read-out can start.

An example of the values of the `ROI_value`, for several soft triggered events is shown in Figure 8.6. The top plot represents the value of the `ROI_value` as a function of the event number, in the middle and bottom ones, the value of antenna 1 is subtracted from antennas 2 and 3, where a close-up between -5 and 5 `ROI_value` is done for the bottom plot. A repetitive pattern is present in the top plot for all of the DRS4 chips. On the bottom plot, small deviations, in the order of 2 to 5 ns, between the three DRS4 chips can be observed. This is believed to arise from the traveling time of the command in the system, i.e., antenna 1 is first to receive the command, then antenna 2, and finally antenna 3. The behavior, best seen in the middle plot, is, however, unexpected. The values of antenna 2 and antenna 3 diverge largely compared to antenna 1. It remains unclear for the moment whether this behavior emerges from bit corruption in the registers, similarly as for the waveforms, whether another aspect is the cause, or whether this is simply the normal behavior. Tests were conducted in the lab, but they, remain inconclusive so far. A similar trend with higher unexpected values rising with temperature was, however, observed.

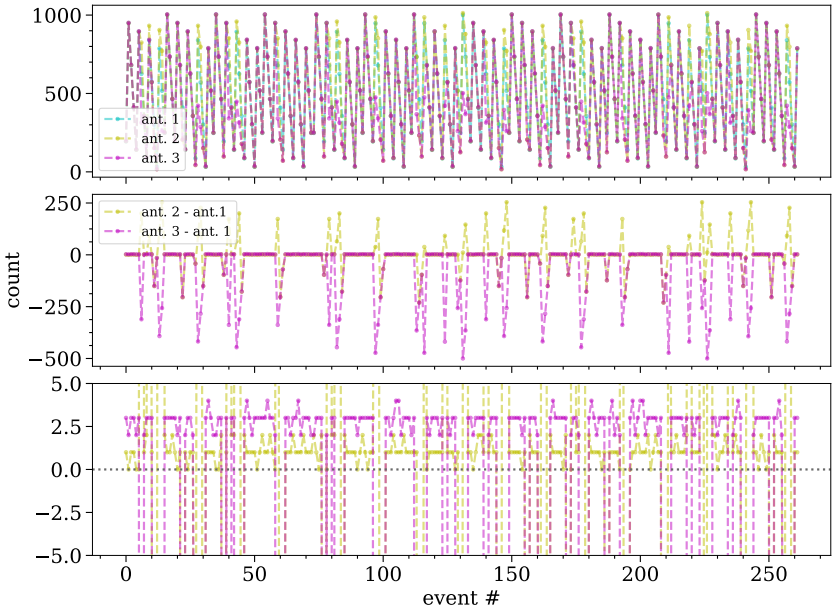


Figure 8.6: ROIs of the DRS4 chips as a function of the event number for the three antennas. Each DRS4 chip is associated with one antenna. On top the raw ROIs values are plotted, on the middle and bottom, the ROI values of antenna 1 are subtracted from those of antenna 2 and 3.

8.3 Radio background from the Milky Way

There are two different techniques to observe the radio noise emitted from the Galaxy. The first one uses the time domain data, in other words, the waveforms, and the second one, the frequency spectrum. In this section, data recorded with the soft trigger mode of the data acquisition system is used.

8.3.1 Galactic rotation around the antennas

At the South Pole, the galactic center is always located at a zenith angle of approximately 61° and performs one complete revolution around the geographic South Pole during a sidereal day. The recorded height of an incoming radio wave is maximal when its Poynting vector is perpendicular to the face of one polarisation arm of the antenna. Assuming that the center of the galaxy emits more radiation [135], then it is expected to see an increase in the noise level when one polarisation arm faces the galactic center. Thus, an increase (and decrease) of the noise level should be seen two times a day in one polarisation, where the noise level of one antenna would follow a sinusoidal function. As the antennas have two perpendicular polarisations, the sinusoidal function would be mirrored between the two polarisations. The revolution of the galaxy around an antenna at the South Pole is illustrated in Figure 8.7.

Since the effect from this rotation is small, the data needs to be processed to enhance it. In a first step, only soft triggered events are taken in the data to avoid contamination by air shower pulses. Afterward, the artifacts remover is run to remove possible bin spikes. Finally, the electronic response is removed and the data is filtered between 100 and 300 MHz. The removal of the electronic response is not necessary to observe this effect, it, however, makes the results to be comparable between different electronic systems. Finally, the noise level is estimated by using the subtraces method in combination with the RMS values.

The results for three days in April 2020 are shown in Figure 8.8, for each channel. The individual

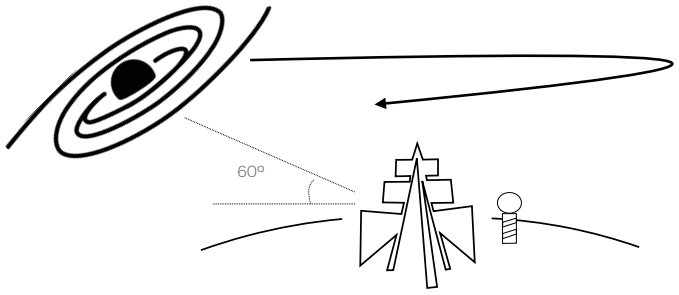


Figure 8.7: Illustration of the galactic center revolving around one antenna.

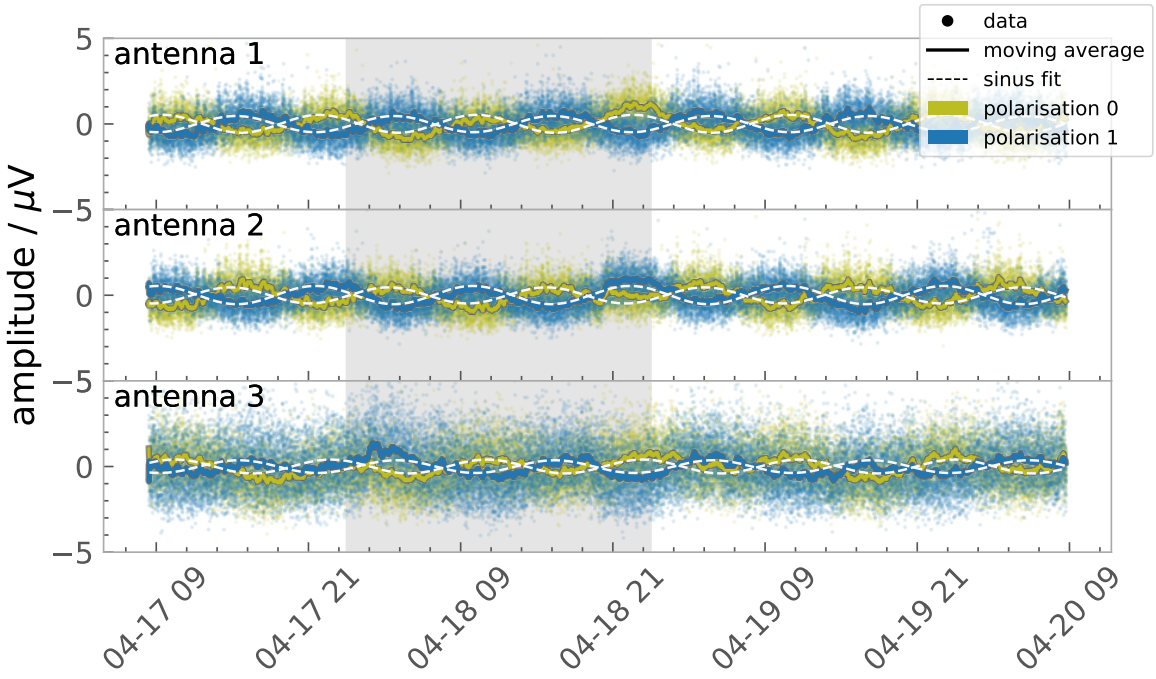


Figure 8.8: Galactic center oscillation observed in the noise level of the different antenna channels. Each channel has its average value subtracted. The time of the x-axis is in month-day hour format.

data points are subtracted by the mean of each channel individually and the results are represented by the scatter points. A moving average is calculated for a window of 150 points and the variation for one day is highlighted in gray. A sinusoidal fit is applied to the data and is shown by the dashed line. The values from the fits are summarized in Table 8.2.

One can see visually that, as expected, both polarisation oscillates oppositely to each other. This is also corroborated by the sinusoidal fit which has a difference in phase ($|\Delta(\omega)|$) close to 90°. In the field, the antennas polarisation are not necessarily aligned to each other, e.g. antenna 1 polarisation 0 is not perfectly parallel to antenna 2 polarisation 1, which explains why the maxima (or minima) between antennas are not aligned. The presence of two maxima per day per polarisation is seen in each antenna channel. Antenna 1 and antenna 2 have a comparable amplitude of the oscillation and similar overall behavior, whereas antenna 3 data is more scattered.

Table 8.2: Sinusoidal fit summary

antenna	polarisation	amplitude [μV]	T [hour]	ω [deg]	$\lvert\Delta(\omega)\rvert$ [deg]
1	0	0.52	12.2	132	85
	1	0.47	12.1	47	
2	0	0.47	12.2	56	76
	1	0.53	12.2	132	
3	0	0.41	12.2	17	84
	1	0.37	12.1	101	

8.3.2 Background frequency spectrum at the South Pole

Discrete Fourier transform in radcube

Fourier transform functions are used to transform a signal in the time domain to the frequency domain. The principle is to represent a function in the time domain by a set of sinusoidal functions of different intensities. The continuous Fourier transform is described by

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i2\pi t\omega} \mathrm{d}t. \tag{8.1}$$

where t represent time and ω frequency. Equivalently, it is possible to transform a function in the frequency domain to the time domain via an inverse Fourier transform in the form

$$f(t) = \int_{-\infty}^{\infty} F(\omega) e^{i2\pi t\omega} \mathrm{d}\omega \tag{8.2}$$

by convention, a capitalized letter represents the function in the frequency domain and a small letter in the time domain. In most of the physical systems, the measured functions are discrete in time (and thus in frequency), and the equations above become

$$X_k = \Delta t \sum_{n=0}^{N-1} x_n e^{-i2\pi \frac{nk}{N}}, \qquad x_n = \Delta f \sum_{k=0}^{N-1} X_k e^{i2\pi \frac{nk}{N}}. \tag{8.3}$$

where Δt represent the sampling time of the system, here 1 ns, and N the length of the waveform in units of bins. These equations are called discrete Fourier transform. The fast Fourier transform (FFT) is a particularly fast computational implementation of a discrete Fourier transform. The calculation of the Fourier transform is executed with the tools available in radcube [136], using this specific normalization.

8.3.3 Frequency spectrum

The FFT is calculated for each of the soft triggered waveforms of one day, 17$^{\text{th}}$ April 2020, subsequently the median value of each bin in the spectra is taken. The amplitude of the FFT is converted to power with the `GetDbmHzFromFourierAmplitude` from the radcube module. In addition, the spectral power expectation is calculated from the Cane model [137] with an addition of 40 K thermal noise. Both emissions are multiplied by the electronic response of the system.

The spectra recorded from the prototype station as well as the expectation are illustrated in Figure 8.9. The spectrum in each channel is comparable and follows the curve of the minimal noise background expectation at the site. A higher noise level between around 50 and 100 Mhz can be

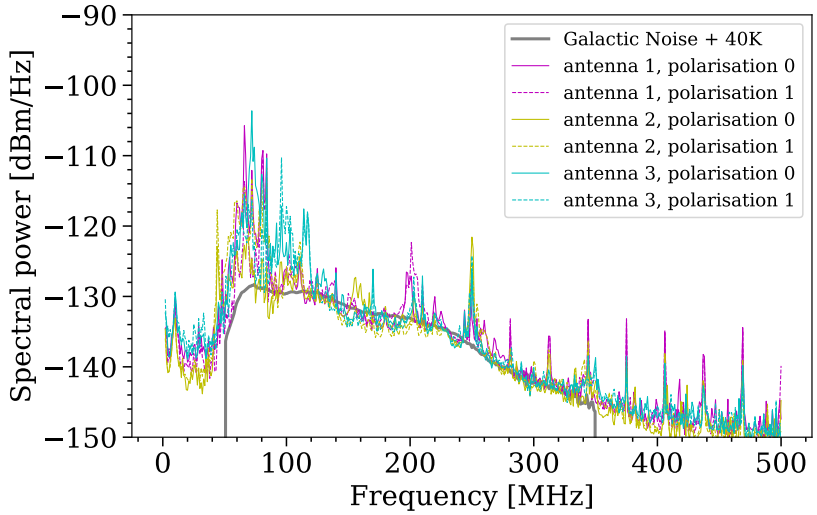


Figure 8.9: Spectrum of the median of one day (17.04.2020) in each channel (in color) compared with the galactic noise derived from the Cane model with an addition of 40 K thermal noise (in grey) [119].

observed. This feature motivated improvements on the TAXI board explained in Chapter 6. Sharp increases in the spectrum every ∼20 MHz are also visible. From an investigation conducted during the deployment, this is believed to be caused by RFI emitted from the server room in the ICL [103].

8.3.4 Monitoring

A monitoring web page[2] displays the median frequency spectrum for every day as well as keeping a record of the previous spectra, hence, tracking the stability of the spectrum baseline over time. Furthermore, the system automatically sends an alert if the baseline of the spectrum for frequencies above 400 MHz rises above ∼35 dBm/Hz, which would indicate either a malfunction or an exceptionally high background.

8.4 Trigger rates

As mentioned, the TAXI system operates with two different types of triggering, soft and scint. triggers. In this section, both triggers are studied for the years 2020 and 2021. In the i3 files, the key named "RadioTAXITime" is used, which represents the time at which the FPGA recorded the trigger and the key "SurfaceFilters" to choose the trigger type. The rate of a certain trigger type is calculated via

$$\frac{1}{\Delta t} = \frac{1}{t_i - t_{i-1}} \tag{8.4}$$

The distributions of event rates for 2020 and 2021 is shown in Figure 8.10. The scint. trigger is represented by the hollow histograms and the soft trigger by the filled histograms. The distribution from 2020, in blue, follows a similar trend as the distribution of 2021, in orange, albeit slightly offset. Although the usual settings of the prototype station are fixed to 6 scintillation panels above a threshold of 4095 ADC counts, different settings were tried during the years, and in particular a lower threshold setting, equivalent to approximately 0.5 MIP [114], from approximately the 4$^{\text{th}}$ April to the 1$^{\text{st}}$ September 2021. The line in green corresponds to the data for the year 2021 with the low threshold period removed, which brings the distribution closer to the 2020 one. Considering that several tests were performed amid the years, their disentanglement is not trivial.

[2] Spectrum monitoring for the Surface Array Enhancement

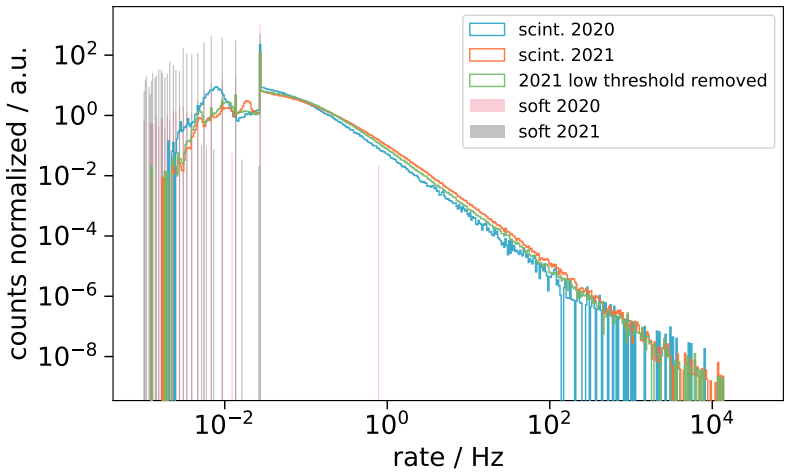


Figure 8.10: Distribution of the scint. trigger in hollow and of the soft trigger in filled histograms.

The distribution of the scint. triggers are smooth for rates above $\sim 7 \times 10^{-2}$, for rates below, some features can be seen. The highest peaks in the soft triggers as well as scint. triggers distribution are located in the 2.67×10^{-2} Hz bin, corresponding to a soft trigger read-out every ~37 s. The soft trigger distribution should, in principle, be composed of only one bin, if the settings are constant over time, which they mostly are. A change of settings can be seen for example in the pink line at a rate close to 1 Hz. This is not, however, what explains the presence of a rate below 2.67×10^{-2} Hz. In fact, these are an artifact of the triggers separation method, where contamination occurs in the soft triggers from the scint. triggers, and vice-versa.

The maximal rate for the years 2021 and 2022 are 7,921 Hz and 13,312 Hz respectively, which corresponds to 126.2 µs and 75.1 µs. An experiment conducted in the laboratory, using a function generator sending two pulses separated by a defined time, showed that the dead-time of the TAXI v3.0 board, for recording two radio events subsequently is (14.15±0.01) kHz, or ~70.7 µs [130], which is closed to what is seen from the prototype station.

Another aspect to examine is the stability of the trigger rate over time. Assuming a fixed threshold, one would expect, on average, a constant scint. triggering rate. In Figure 8.11, each Δt between two consecutive scint. triggered events are plotted as a function of the time of the event, illustrated in blue in that figure. The black line with white squares represents a one-day moving average. From this, one can see the contamination of the soft/scint. trigger separation method, which gives rise to the features below $\sim 7 \cdot 10^{-2}$. Secondly, one can see the effects of different settings for the scintillation panels, e.g. around the 1st June 2021 when a sharp step occurs in the moving average. One can also see the longer low-threshold period in this figure. The white spaces are due to the system not recording or triggering in that period for diverse reasons, mainly related to scintillation detectors and firmware testing. Finally, outside all of that, one can see fluctuations, large and small, over time, which are especially visible in the moving average. A tendency for a higher trigger rate around the Austral winter months, and lower in Austral summer months.

Studying that observed tendency, the trigger rates are plotted as a function of the outside temperature at the South Pole [138] in Figure 8.12. A 2 h average is calculated for the trigger rates as well as for the temperature, the colors represent months in the year. An inverse linear correlation between the outside temperature and the trigger rate can be seen. In other words, the lower the temperature, the higher the triggering of the radio read-out. The cloud of points located far from the correlation are due to the different settings used at that time, in particular, the largest cloud is related to the low threshold period.

The temperature dependence of the scintillation panels, which trigger the radio read-out, can

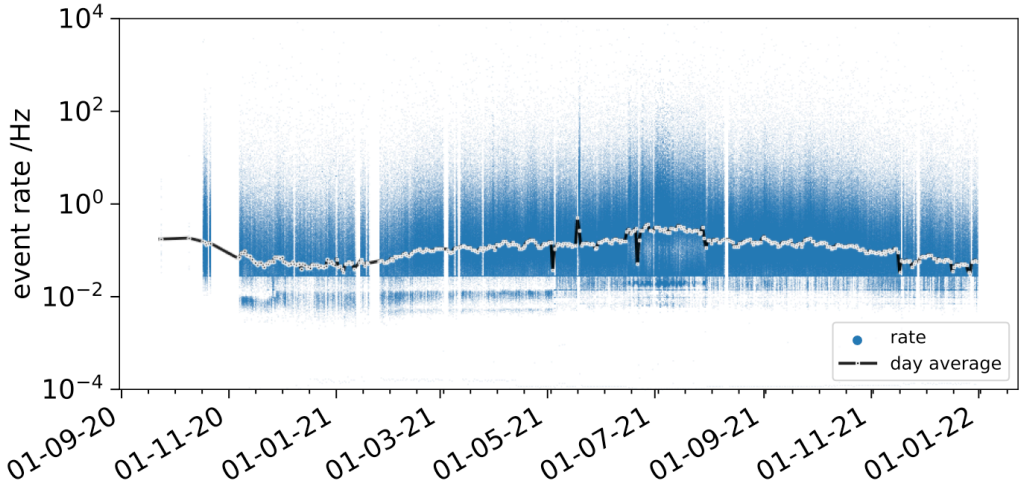


Figure 8.11: Scint. trigger rates with a 2 h moving average as a function of time during the lifetime of TAXI v3.0

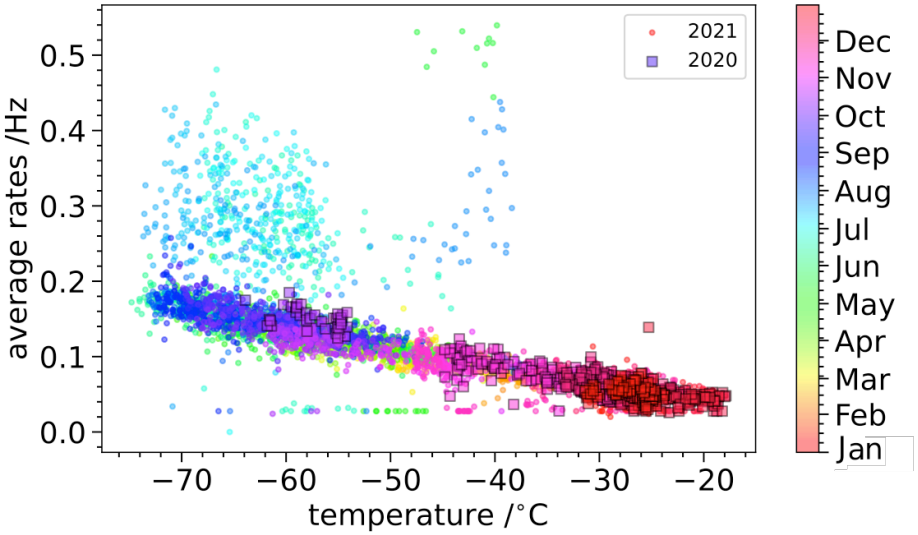


Figure 8.12: Correlation plot between the 2 h average triggering rates for the radio read-out and the 2 h average temperatures at the South Pole.

explain this effect. As illustrated in Figure 8.13, the gain of the scintillators depends on both the temperature and the bias-voltage (AUXDAC in that figure). For the standard operation of the prototype station, the applied voltage is constant, resulting in a change in the gain of the scintillation panels with temperature. Additionally, the threshold in ADC counts of the scintillators for radio read-out is constant during standard operation. Therefore, if the outside temperature is lower, the gain of the scintillation panels is higher, which makes the threshold for radio read-out easier to reach. This, in turn, effectively triggers lower-energy air showers.

One can also notice a slight flattening of the curve in December-January. This is caused by the unstable baseline of the scintillators in the summertime due to light leaking inside the detectors. This issue was fixed with light-tight tape during the change of the TAXI box in January 2022. In the future, the scintillators will have an in-situ calibration using a temperature sensor located close to the SiPM in order to reach a constant gain.

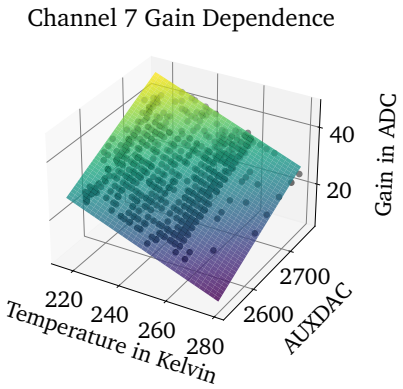


Figure 8.13: Calibration plot for one scintillation panel relating the temperature of the SiPM, to the applied voltage (AUXDAX), and to the gain of the panel. From Ref. [119].

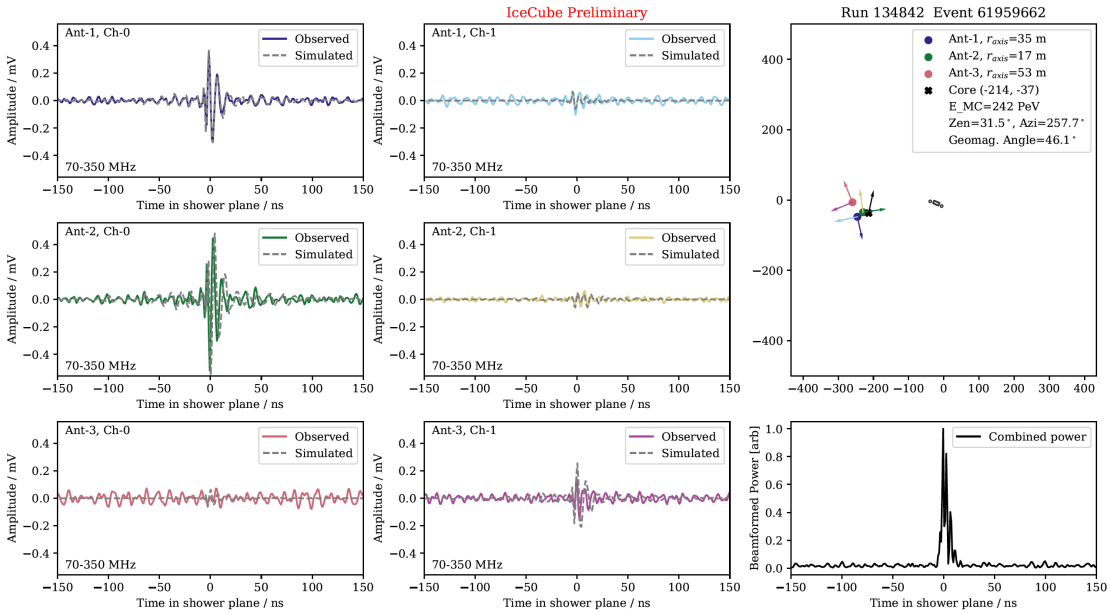


Figure 8.14: An example of one identified radio event for which a MC simulation of air showers was performed using the parameters reconstructed with IceTop. On the left and middle columns, the recorded waveforms of each polarisation channel (in solid lines) are plotted with their analogous simulated waveforms (in dashed lines). In the top right plot, the antenna locations and the Poynting vector of their polarisations are shown in colors as well as the reconstructed shower core position and direction in black. The bottom right plot shows the combined radiation power normalized of all the channels beamformed according to the IceTop direction reconstruction. From Ref. [123].

8.5 Air showers from the radio array

During the operational time of the prototype station, several air showers were found in the radio array. The details of the search algorithm in radio as well as coincidences with scintillation and IceTop (IT) detectors in the data can be found in [123]. A great way to confirm that the recorded pulses come from an air shower event and not RFI, and, at the same time, to verify that the processing chain implemented in radcube describes well the system, is to compare a simulated air shower to the recorded air shower.

In Chapter 10, MC simulations are done for a few of these air showers. Those simulations use the

reconstruction from IceTop as input, and with the radcube modules, the instrument response is applied to the simulated electric field to resemble the measured data [136]. In Figure 8.14, the simulated waveforms are compared to the recorded waveforms for each polarisation channel individually in the plots in the left and middle columns. The solid lines represent data and the dashed lines, simulation. Good agreement regarding the size and the shape of the pulses between the data and simulation for all channels can be noted. On the top right side, in color, the locations of the antennas and their associated Poynting vectors are shown. In black, the core location of the air shower and its arrival direction are illustrated. Finally, the waveform on the bottom right shows the beamformed signal (see Section 10.7 for more details about beamforming) using the time delay according to the IT directional reconstruction.

8.6 Quadruple coincidence – radio antennas, scintillators, IceTop tanks and in-ice detectors

By construction, if there is a radio shower, there are at least 6 scintillators with a signal over threshold, as they trigger the antenna read-out. Similarly, for IceTop, with a larger array size and lower (CR-) energy threshold than the radio antennas, an IceTop event should be coincident when radio air showers are found. Hence, in principle, radio signals from air showers are the bottleneck when searching coincident air showers with all the surface instrumentation. In this vein, to find quadruple coincidence events that also include an in-ice signal, the pool of identified radio air showers is taken, which amounts to 121 air showers, recorded between the 6^{th} December 2020 to the 7^{th} April 2022.

8.6.1 Search for in-ice coincidences

Once more, the IceTop direction reconstruction of the shower axis is taken for each identified radio event and prolonged inside the ice. A perimeter is drawn around the uppermost layer of the in-ice detector and the events which are located inside that perimeter are tagged. This principle is depicted in Figure 8.15, where, on the left plot, all the shower axes are pictured in gray and the axes which cross the perimeter area, in blue. The two upper plots on the right side illustrate the same information but in the x-z and y-z planes. The red line acts as a visual reference and shows a vertical line crossing the centers of the detectors. Lastly, the bottom right plot represents the core position of the radio events at the level of the uppermost layer of the in-ice detector.

In the set of 121 air showers with radio signals, five events have a possible in-ice counterpart, equivalent to approximately 4% of the radio events. Once those events have been identified, their characteristic runid and eventid are used to find an associated in-ice event.

From those five events, the event associated with rundid 134950, eventid 81008640 is picked as an example, the four others can be found in Appendix F. Figure 8.16 shows a three-dimensional representation of the event with all the detectors, with an isometric projection on top and on the bottom in a bird's-eye view. The colors represent the timing between the charges of the IceTop and in-ice detectors, in the bird's-eye view plot the in-ice timing is plotted in gray for better visualization. This view focuses on the surface detectors, with the scintillation panels represented by squares, the IceTop tanks by circles, and the antennas by crosses. The size of the marker represents the charges deposited in the individual detectors for the in-ice DOMs, the IceTop tanks, and the scintillation panels. For the antennas, the size of each bar depicts the amplitude of the maximum of the Hilbert envelope in the respective polarisation channel, for a waveform filtered between 80 and 300 MHz, after removal of the pedestal and conversion to mV. The pulses are found using the methodology explained in Chapter 9. The waveforms related to this event are shown in Figure 8.17. For IceTop charges in tanks, the "OfflineIceTopHLCTankPulses" key is taken from the i3 frame, for the scintillators the "ScintRecoPulses" key, and for the in-ice DOMs the "SRTInIcePulses".

In Figure 8.17, the three rows represent the three antennas with both polarisation channels. The left plots are a close-up on the peaks. The gray band represents the search window and the black band the pulse window. Clearly, for this event, all antenna channels recorded a distinct radio pulse.

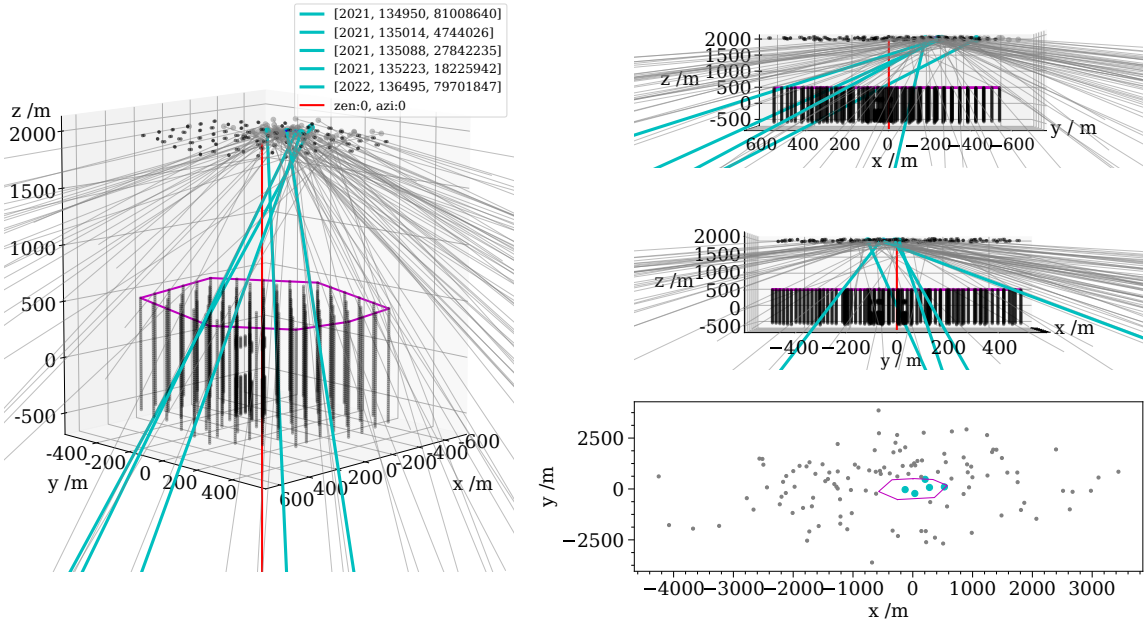


Figure 8.15: Identification process of a possible quadruple coincidence. The left and the two upper right plots show, in gray, the axes of the IceTop direction reconstruction of all 121 air showers detected with the radio antennas, and, in blue, the axes which pass through the area drawn by the purple borders, i.e., the top layer of the in-ice detector. The bottom right plot depicts the core positions of all showers at the depth of the uppermost layer of the in-ice detector.

One can wonder how can there be only three scintillators with charges when six are required for triggering the radio read-out. The reason is that the communication between the scintillators and the TAXI was not optimal and there were some issues in the data transfer, e.g. a full buffer. In this event, 4 more panels had a signal over threshold but unfortunately, the data is not available. The IceTop reconstruction is shown as well as two reconstructions of the in-ice signal and they are in agreement. In the future, analysis methods should be developed that make use of the rich information content of these multi-hybrid events.

8.7 Discussion

A prototype station was deployed in January 2020, consisting of 3 antennas, 8 scintillation panels, and a new DAQ system as described in Chapter 6. This prototype station closely resembles the specifications that will be used for the complete surface enhancement array. Based on the fact that the galactic background, which is the hard limit on the noise level, can be witnessed in both the time domain and frequency domain, this chapter emphasizes that the operation of the radio array is successful with only minor fine-tuning needed. Furthermore, during approximately two years of operation, 121 air showers were identified in the data, which are in agreement with the IceTop reconstruction and simulations. The detailed method of finding air shower radio signals and the comparison with IceTop can be found in [123], while a more in-depth analysis of the scintillation detectors' performance for this prototype station is described in [114]. A forthcoming publication will describe the performance of this station in regard to the radio antennas and scintillators.

The few issues highlighted in this chapter, such as the bin spikes in the waveforms, the noise level of the system in general, and the possible corruption of the binary files, are all under active research within the group. A working solution to reduce the bin spike issue is already implemented. The version of the prototype station that is analyzed in this chapter is not existing anymore. An improved version

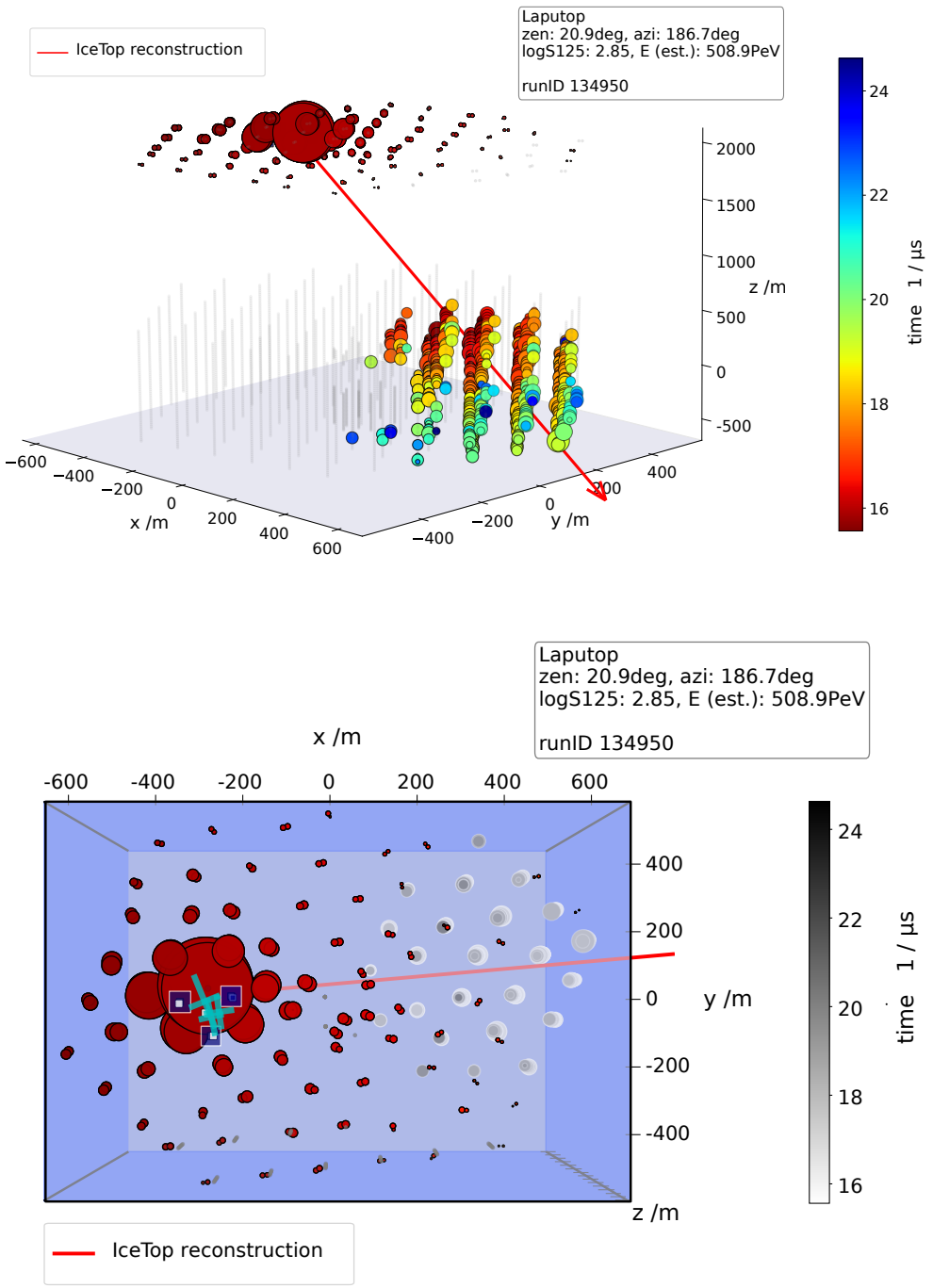


Figure 8.16: Visualization of all detector charges (size) and timing (color) in two perspectives: (top) isometric projection and (bottom) bird's-eye view. IceTop tanks are represented by red circles, in-ice DOMs by gray circles, scintillation panels by blue squares, and antennas by cyan crosses. The colors of the antennas and scintillation panels are not related to timing.

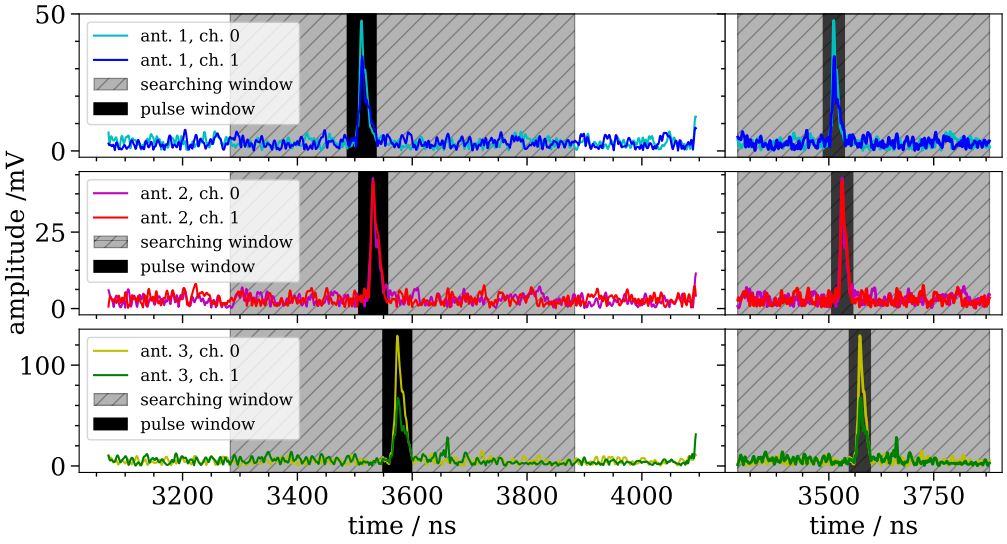


Figure 8.17: Measured radio waveforms of the event shown in Figure 8.16. Data is taken in non-cascaded mode and only the last physical DRS4 channel is shown.

of the DAQ using TAXI v3.2 with the radioTad v2 is in operation at the South Pole since January 2022. The tests outlined in this chapter were also conducted with the new system, but the noise level of the system is higher than expected, the cause is due to a capacitor baseline correction issue. The tools are now in place to correct for it, and the mass re-processing of the data will be executed in the next few months. There is however no purpose in showing the results of the newer version in the chapter, since this issue is identified, and it furthermore renders any analysis that would be done with the processed currently available data obsolete.

Therefore, also the analyses presented in the next chapters use the data recorded with the TAXI. A study of how the noise affects the true data is explained in Chapter 9, and this will lay the building blocks for the energy and X_{max} reconstruction using a LLH of Chapter 10.

CHAPTER 9

Influence of the Background Noise on the Radio Signal

In the era of high-precision physics, any high-level reconstruction or air shower parameters necessitate a careful study of the impact of the noise on the signal. For a radio array, the noise depends mainly on the location and on the electronics, which renders it detector-specific. In this chapter, the effect of the modeled noise as well as the measured noise from the prototype station on the signal of two simulated air shower events will be investigated. In a first step, the parameters for the Monte Carlo simulations are introduced. Subsequently, the methodology for injecting noise, and searching for pulses in the waveforms is explained. Finally, the results of the injection of noise on two measurable quantities in the waveform are presented.

9.1 Monte Carlo simulation set

CORSIKA is a framework designed to generate air shower simulations based on a Monte Carlo approach to capture the variations between showers [139]. It uses knowledge about particle interactions at low and high energy. A shower library was created for comic ray analysis at the South Pole location. This library is created using CORSIKA v7.7401 [139], CoREAS V1.4 [140] for the radio emission simulation, Sibyll 2.3d [141] for the high energy interactions and Fluka2011.2x [142] for the low energy interaction. EGS4 [143] is used for electromagnetic interactions.

The air shower emission is recorded in certain defined locations. The antenna-positions[1] are distributed in an 8-arms star-shaped pattern. This layout is symmetric in the $\vec{v}\times\vec{B}$ and $\vec{v}\times(\vec{v}\times\vec{B})$ reference frame in the shower plane. The positions of the antennas from the star-shaped pattern are shown in Figure 9.1, where the shower core is located in the center of that reference frame. From the star-shaped pattern, the complete time series and the frequency spectrum at any position inside the radius of the star can be interpolated [136], [144]. However, in this study no interpolation is applied, only the value of the pulses at the given location on the star-shaped pattern is taken.

For this study, two examples MC simulated events are used. The first one is configured with the energy of $1.24\times10^{18.0}$ eV, with a zenith of 46.7°, an azimuth of 19.5°, and a proton as the primary particle. The second one has a higher energy of $3.9\times10^{18.0}$ eV, a zenith angle of 67.3°, an azimuth angle of 165.8°, and an iron nucleus as primary. The assumption is that, for this study, the only important aspect is acquiring numerous different pulses with a variety of strengths. To ensure that there is no unexpected behavior related to the CR-energy or zenith angle, the second event is used as a cross-check. The first exemplary library event uses a CORSIKA event steering file set up as shown in Table 9.1.

In this table, the SEED field is used for random seed numbers for starting the showers, which determine the interactions and decays, ERANGE is the energy of the primary particle in GeV, PRMPAR 14 means that this is a proton primary, THETAP and PHIP dictate the incoming direction of the cosmic ray in the CORSIKA reference frame, and THIN is the energy level at which below that particles are

[1] This refers to the position where the antenna would be positioned, however, this does not take any antenna model into account in the MC simulation for obtaining the electric field at this specific location.

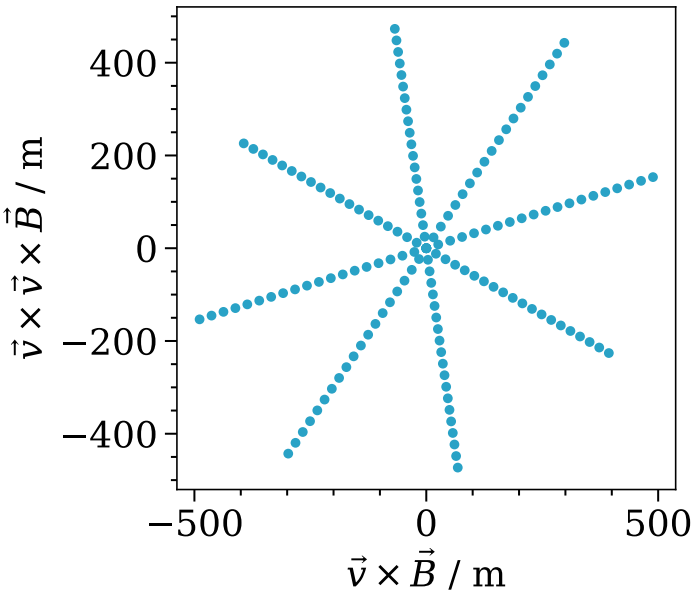


Figure 9.1: Antenna locations of the 8-arms star-shaped pattern in the $\vec{v}\times\vec{B}$ / $\vec{v}\times(\vec{v}\times\vec{B})$ reference system of the shower plane. The layout used for both of the example MC simulated events.

grouped together. MAGNET, ATMOD, and OBSLEV are parameters specific to the detector location and correspond to the geomagnetic field, the atmosphere, and the height of the detector, respectively. For a detailed description of the parameters please refer to the CORSIKA user guide [145].

The simulations are then treated with radcube [136], a module within the IceTray [98] software which is specially developed for the treatment of the radio data, simulations, and measurements. The radio emission is convoluted with the electronic response of the prototype station 2020 and filtered with a box filter to the band of 80 and 300 MHz, the same band that the reconstruction of the next chapter is using. The waveforms are treated in this way to resemble the data recorded by the TAXI. The processing of the waveforms obtained via the Monte Carlo simulations, using radcube, is summarized in Table 9.2

This processing is for noiseless simulations. For the waveform with noise, the two example events are read n time and at each of these readings noise is injected. For the modeled noise, the injection happens between steps 5 and 6. For the "measured noise", obtained from measurements with the prototype station, the noise injection is done between steps 7 and 8, because it already has the electronic response incorporated. During the injection of the noise the two time series, noise and signal, are added together. The definitions of two noise types are given later in this chapter. The number of readings n defines the number of times a random noise is applied in each waveform from one simulation, e.g. if the simulation is read 1000 times, there will be for each single pulse, a thousand equivalent pulses with noise.

9.2 Definitions and methodology

The methodology used for investigating the noise, which encompasses: noise calculation, pulse identification, types of noise, and the choice of measure quantity, is used similarly in the next chapter. Ergo, the results of this chapter are utilized for the implementation of the log-likelihood procedure for energy and $X_{\max}$ reconstruction presented in the next chapter.

9.2.1 Types of noise

Two types of noise are injected in the traces: a modeled noise and measured noise.

Table 9.1: Example library steering file for CORSIKA

RUNNR	0
EVTNR	1
SEED	18798715 0 0
SEED	19798715 0 0
SEED	20798716 0 0
NSHOW	1
ERANGE	1237.96[...]e+6 1237.96[...]e+6
ESLOPE	-1.0
PRMPAR	14
THETAP	46.6699 46.6699
PHIP	78.8086 78.8086
THIN	1e-06 1237.96[...] 0.0
THINH	2.00E+02 10.000000
ECUTS	0.02 0.01 4.0E-04 4.00E-04
CASCADE	F F F
ELMFLG	T T
OBSLEV	284000.0
ECTMAP	1.e11
SIBYLL	T 0
SIBSIG	T
FIXHEI	0. 0
HADFLG	0 1 0 1 0 2
STEPFC	1.0
MUMULT	T
MUADDI	T
MAXPRT	1
MAGNET	16.75 -51.96
LONGI	T 10. T T
RADNKG	2.E5
ATMOD	33
EXIT	

Table 9.2: Radio emission treatment in radcube

steps	**processing without noise**
1	import the electronic response and the antenna response
2	read CORSIKA simulation
3	pad E-field with zeros to obtain a fix of length for the waveform
4	convolve E-field with antenna response
5	re-sample for bins of 1 ns
6	add electronic response
7	digitize waveform
8	remove pedestal and convert to voltage
9	filter between 80 and 300 MHz
10	write i3 file
11	extract pulses from i3 file
12	any subsequent analysis

Modelled noise. The modeled noise calculated from the non-thermal galactic radio emission in the polar regions, using the Cane model [137], and in addition, a 40 K thermal noise, which corresponds to the level expected from the LNA [124]. The amplitude of the frequency spectra from these two contributions are combined, and then phases at each frequency are chosen randomly to obtain a white noise. The modeled noise is shown in the previous chapter in contrast to the recorded spectrum and the shape of both spectra agrees well with each other, except for the RFI peaks that can be seen in the measured spectrum.

Measured noise. This noise refers to real background waveforms recorded in the soft trigger mode by the antennas of the prototype station. The frequency spectrum and the distribution of the measured noise can also be seen in the previous chapter. The "measured noise" mentioned here is taken with the TAXI v3.0 and the radioTad v1, using the soft trigger mode, between the 1st December 2020 and 31st December 2020, and recorded in non-cascaded mode. The noise is not associated with a specific antenna channel. Therefore, when the measured noise is injected a thousand times in a specific waveform from the simulation, that noise can be coming from any of the channels of the prototype station. The data is treated with the "binArtifactRemover" (see Appendix H), which removes some artifacts shown in Appendix G, present in the waveforms.

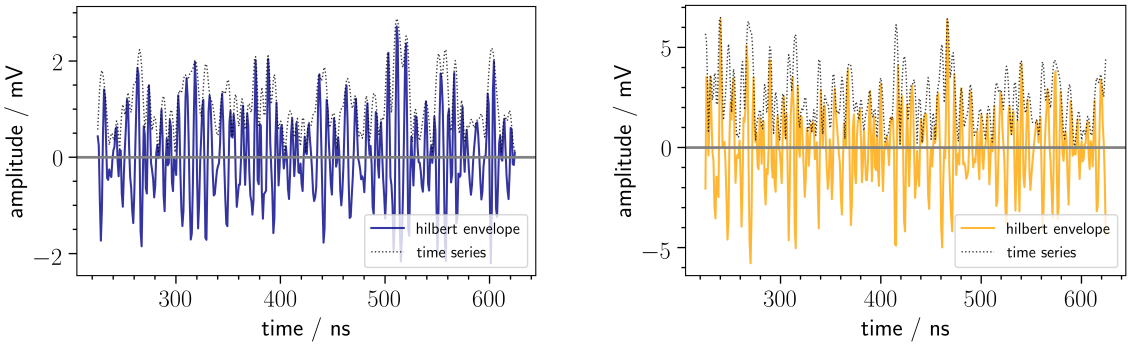


Figure 9.2: Example of a noise waveform filtered to the band between 80–300 MHz. The modeled noise is shown on the left and the measured noise on the right. The y-axis scales are different, the measured noise amplitude is higher than the modeled noise by roughly a factor of 2.

Figure 9.3 shows an example of a waveform with a modeled noise on the left and one with the prototype station on the right hand side. One can see, although the spectral shape is similar, the average noise level is smaller with the modeled noise than for the measured noise by almost a factor of two.

9.2.2 Measurable quantities

Two classes of measurable quantities can be extracted from the pulse in a waveform: amplitude measurable and integrative measurable.

Amplitude measurable refers to a quantity that is measured from the amplitude of the signal, the magnitude of the signal, or the Hilbert envelope of the signal. For this study, the amplitude quantity used is the maximum of the Hilbert envelope, A_{Hilb}.

Integrative measurable denotes a quantity that is measured from the integration of a certain section of the signal. Here, the power within a 50 ns window, P_{50}, is used.

For both cases, the measurable quantities are taken from each polarisation channel of each antenna individually. The power is obtained with the `GetPower()` function of radcube module which calculates the power via

$$P = \frac{\Delta t}{Z\, t_{\text{tot}}} \sum_{i}^{b_{\text{win}}} A_i^2 \tag{9.1}$$

where Z is the impedance and is set to $50\,\Omega$, Δt is the sampling time and is equal to 1 ns in this case. A_i is the amplitude of the waveform at bin i and b_{win} is the number of bins over which the summation is made.

9.2.3 Noise level calculation

The standard way of measuring the noise level in a waveform is to take a defined-length window outside the expected signal region and calculate the RMS inside that region. The RMS can be calculated by

$$x_{\text{RMS}} = \sqrt{\frac{1}{n}\left(x_1^2 + x_2^2 + \cdots + x_n^2\right)}. \tag{9.2}$$

where x_i is the signal amplitude in each bin. The RMS gives a measure of the average power contained in the noise of a waveform but in units of amplitudes, i.e., volts. It is obtained via the `GetRMS()` function in radcube. The main downside of this technique is that the presence of RFI within that noise window would alter the result. If transient noise occurs inside the noise window, the noise will be over-estimated. To counteract that, a new method is implemented. The traces are split in multiple parts of equal length, these parts are named subtraces. The RMS of each of these subtraces is calculated and then the median of the 10 smallest RMS values is taken. The motivation is simply to remove possible biases in the noise estimation from impulsive RFI or coincident showers. This technique further allows for the measurement of the noise level for any waveforms without prior knowledge about possible air shower pulse locations within these waveforms. A question is remaining: how small can the subtraces be in order to make an accurate estimate of the RMS.

To answer that question, this technique is applied to a sinusoidal signal. The signal is sent using a function generator, SDG6052C by Siglent, and the waveform is recorded with the TAXI v3.0. The only processing done on the trace is the removal of the pedestal, i.e., moving the baseline to 0 ADC counts. The frequency of the sinus is 100 MHz, a frequency in the lowest range of what the system records after the analog processing of the signal. Similarly, this is equivalent to the longest wavelength, hence the limiting frequency for finding the smallest subtrace sizes.

To estimate the noise level of one waveform, the waveform is cut in multiple subtraces, this is illustrated in the left plot of Figure 9.3, where the above-mentioned method is applied. For visual help, the histogram of the absolute value of the measured amplitude of the sinus signal in each bin of the waveform is shown in the right plot. The black vertical line represents the RMS calculated over the complete waveform and the magenta one, the noise estimation from the subtraces method.

This value is recalculated with different lengths of the subtraces. The results of the noise estimated with the developed method compared to the RMS of the waveform is summarized in Figure 9.4. The noise estimation with subtraces rises until the subtrace size is equal to 64 ns. It then stabilizes at around 1270 ADC counts, whereas the RMS of the complete waveform is about 1280 ADC counts. Therefore, subtraces of 64 ns are used throughout this work, unless specified otherwise.

Noise measured in this way, i.e. using the RMS and the subtraces method will be referred to as σ_{A} and used with quantity measured in amplitudes. The effects of the noise on the power comprised within a pulse will also be analyzed in this chapter. When dealing with power, the same subtraces method is used, however, instead of calculating the RMS in each subtrace, the power is calculated (with Equation (9.1)). This noise level will be referred to as σ_{P}.

9.2.4 Specific used definition of signal-to-noise ratio (SNR)

In order to be able to compare the results of both quantities, the strength of the pulse from the simulations is defined by the following definition of signal-to-noise ratio (SNR):

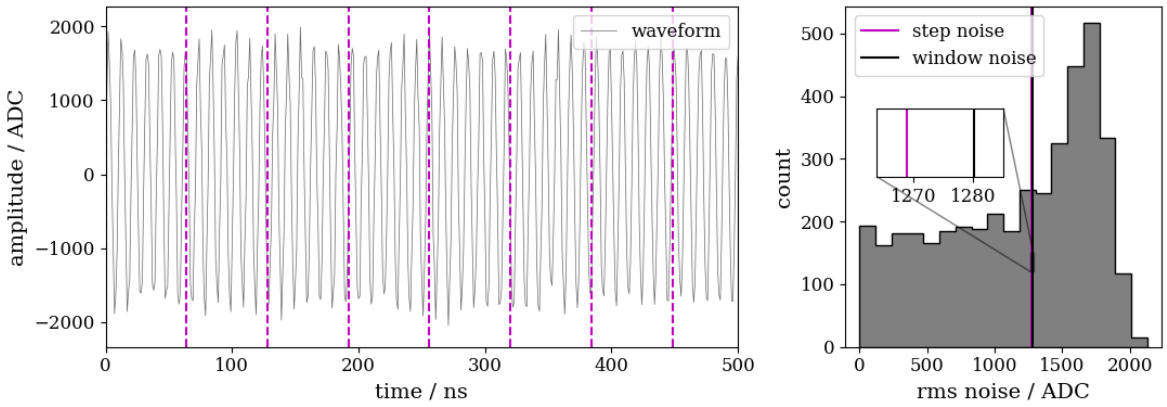


Figure 9.3: Sinusoidal signal with subtraces of 64 ns indicated on the left plot. The right plot is the histogram of the absolute amplitude values with the RMS value calculated with the subtrace technique and the RMS of the whole trace.

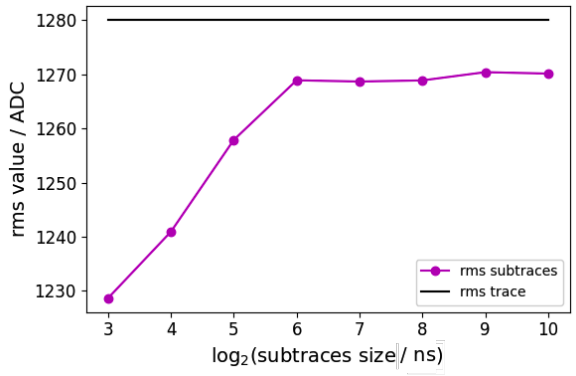


Figure 9.4: Final value of the RMS using the subtraces method against the size of the subtraces.

$$\mathrm{SNR} = \left(\frac{S_\mathrm{A}}{N_\mathrm{A}}\right)^2. \tag{9.3}$$

There is a few different ways of expressing the SNR. In this equation, S_A corresponds to the maximum of the Hilbert envelope of the signal without noise injected, and N_A represents the noise level in the waveform. Usually, the noise level will be calculated in each waveform for which the SNR is associated with. This way is, for one, difficult to apply on the results of this study because multiple waveforms with injected noise are associated with one signal (S_A). For another, with the type of numerical filter used, a "leaking" of the pulse in the rest of the trace is observed. The consequence is an overall artificial rise of the noise level, increasing with the strength of the signal pulse. To solve both of these issues, N_A is a constant. This constant is calculated by using the noise level σ_A, which, as previously mentioned, is calculated with the (RMS) subtraces method for each individual waveform. Thereafter an average is made over the σ_A for one example event, and one noise type injection.

$$N_\mathrm{A} = \overline{\sigma_\mathrm{A}} \tag{9.4}$$

9.2.5 Identifying a pulse in a waveform

The intention is to find the pulses in a way that is applicable to real measurements as well. The assumption is that the pulse time is approximately known from the direction reconstruction, which

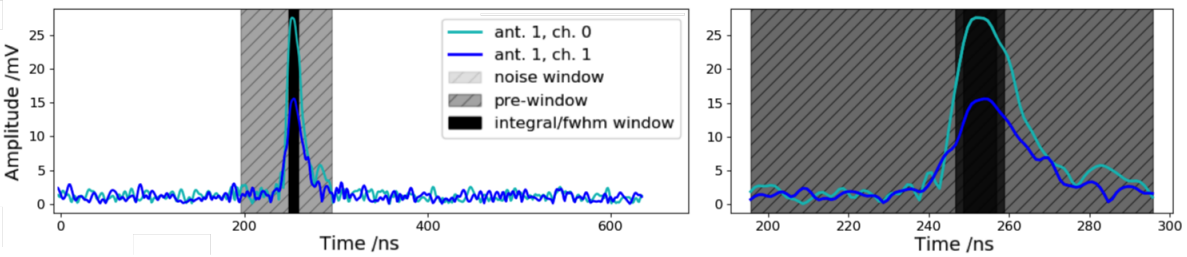


Figure 9.5: Example of a pulse search two simulated waveforms belonging to one antenna. The modelled noise is injected. The gray band represents the 200 ns, pulse window, and the black band the number of bins taken for the integrative measurable. For this example, the number of bins is 10.

can be obtained from IceTop or from the scintillation detectors (see. Section 10.7). In simulations, the position of the pulse in the waveform is known by construction to be always at the same index, in this chapter only simulations are used and this index is set as the initial location of the pulse.

It is possible, depending on the shower geometry, that one polarisation channel has a large pulse and the other one has a small pulse, which could be hidden in the noise. To reduce wrongly identifying the position of the pulse within the search window, when the signal-to-noise ratio is low, we take advantage of the two polarisation channels. Thus, the Hilbert envelopes of both polarisations are summed together for the pulse search. This aspect of correctly identifying an air shower pulse is especially important in the next chapter where the individual polarisations are compared to the Monte Carlo simulations. Once the polarisations are summed, a symmetrical search window of 200 ns around the expected initial pulse position is created. Subsequently, the position of the highest value of the Hilbert envelope in that window is found.

For the maximum of the Hilbert envelope, A_{Hilb}, the value and the time at this position are simply saved for each polarisation, with the assumption that the shower pulse would be located at the same position for both polarisation channels of one antenna. For the integrative measurable, P_{50}, a symmetrical window of (b_{win}) around the maximum is created, and that section of the waveform is saved for each polarisation as well. Although mitigated by the addition of the polarisation channels, if both channels have a small signal-to-noise ratio, which is likely in these two example events, the noise maxima will be found instead of the true pulse.

The different window uses for the identification of pulses are shown in Figure 9.5. The blue and cyan colors represent the waveforms in both polarisation channels of one antenna. The gray band depicts the search window and the black window is the pulse window. The waveform is taken from a simulated event with modeled noise injected, the pulse window represented here is of 10 ns, or 10 bins.

9.2.6 Injecting the noise

From each of the two example events described before, the pulse in 160 antennas, with two polarisation channels each, is recorded, resulting in a total of 320 different pulses per example event. On each of these pulses, each type of noise is injected 1000 times, thus obtaining 1000 similar waveforms out of one pulse. Two sets are created this way, one with modeled noise injected and the other with measured noise injected.

Figure 9.6 shows two examples of the modeled noise injection in a pulse, where on the left the true pulse is smaller than the noise and, on the right, larger than the noise. These are the 50 ns window waveforms, determined in the pulse search. The top left plot confirms that the method finds a noise maximum in case of low signal, whereas in the top right plot, the true pulse is found each time. This can be affirmed by the similar shape between truth (black) and noisy (purple). This is also confirmed with the start time of the distribution of the pulse visible in the bottom row where: when the pulse is large, the pulse window starts always at the same time, inversely when the pulse is small, there is large fluctuations about where the pulse window starts within the search window. Another interesting

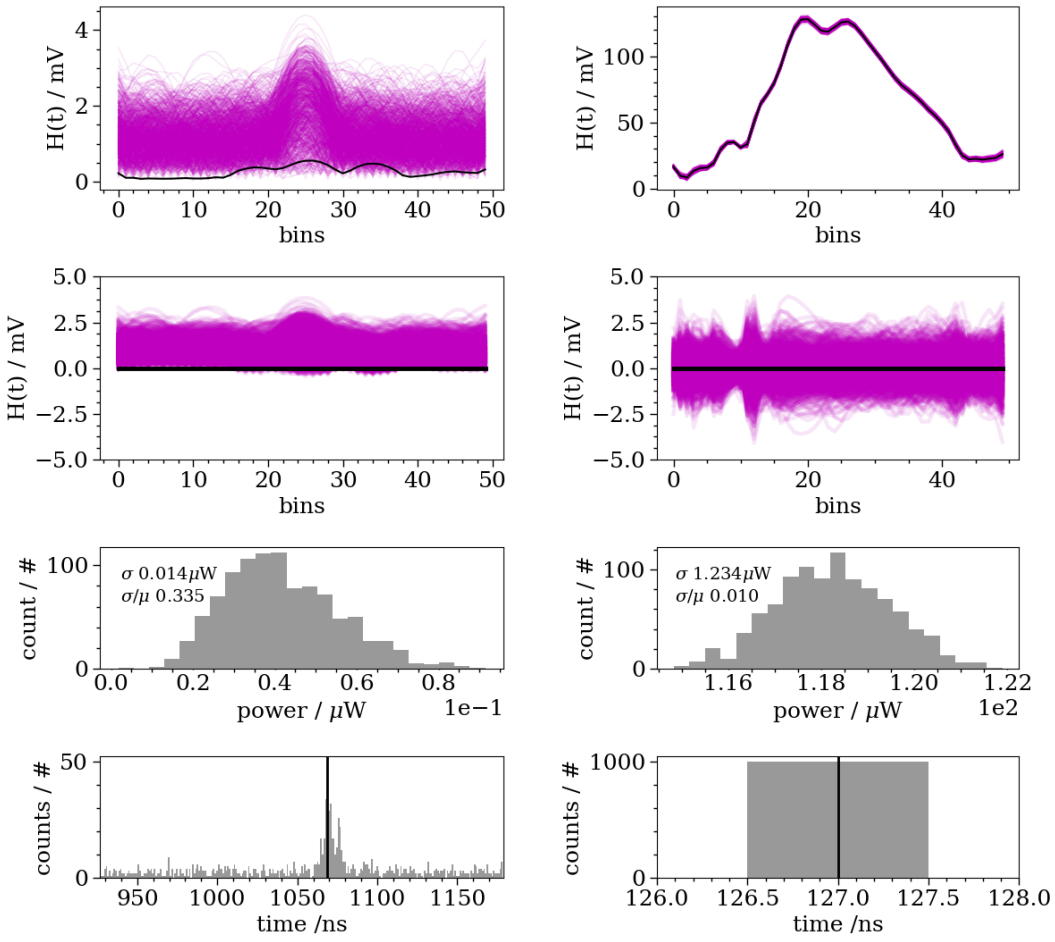


Figure 9.6: Examples of the injection of the modeled noise for a pulse smaller than the noise level on the left and for a pulse larger than the noise level on the right. In the two upper plots, the purple lines represent the Hilbert envelope 1000 pulses with noise injected in one polarisation channel, and in black the corresponding noiseless true pulse. The top row shows all pulses unaltered, and the second to top one, the noisy pulses minus the true pulse. The distribution in the third row represents the values of the integrative measurable for all the purple pulses shown in the top plot. μ and σ are the mean and standard deviation respectively. The histogram in the last row represents the time of the beginning of the pulse window. In gray for the pulses with noise injected and in black for the true pulse.

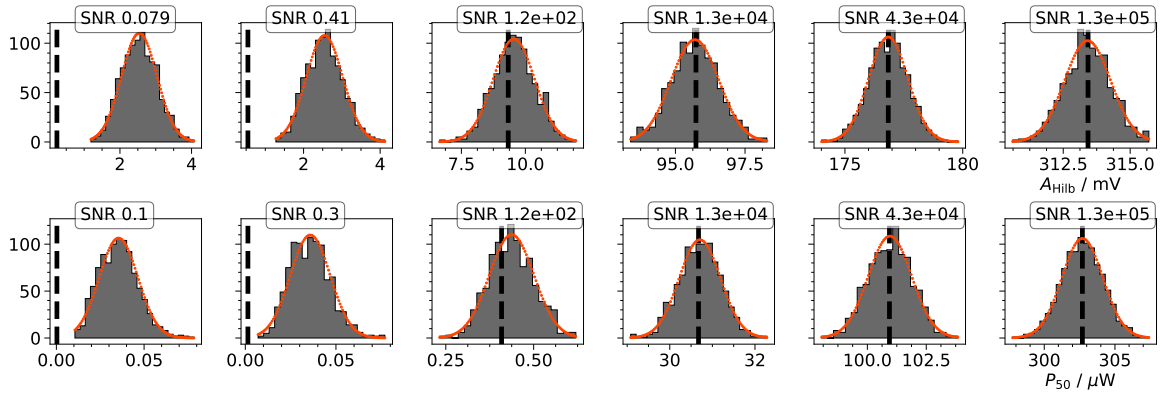


Figure 9.7: Seven distributions showing the value of A_{Hilb} at the top and P_{50} at the bottom for the waveforms with modelled noise injected. The vertical line represent the value of the measurable quantity for the same pulse without noise injected. The distributions are ordered by ascending SNR values from left to right. In these plots the y-axis represents the count number.

observation arises from the plots in the second row. With the small pulse, all the pulses with the noise injected lie above the true pulse and have fluctuation from approximately 0 to 2.5 mV. With the large pulse, the waveform with the noise injected fluctuates above and below the true pulse, spanning approximately from -2.5 mV to 2.5 mV. Finally, the second to bottom plots represent the distribution of the average power (P_{50}) calculated for each of the waveform with noise injected within the shown window of 50 bins (or 50 ns). The standard deviation σ calculated for those histograms is 0.014 µW for the small pulse and 1.234 µW for the large pulse, i.e., the standard deviation of the large pulse is bigger than that for the small pulse. However, the relative standard deviation (σ/μ) is ~0.34 for the small pulse and ~0.01 for the large pulse, thus the relative deviation of the large pulse is smaller than for the small pulse. These interesting aspects are analyzed and explained in more detail in the next section.

9.3 Results

All the methodologies explained above are applied to the two example events. The results are shown first with the modeled noise injected in the traces and subsequently with the measured noise from the prototype station injected.

9.3.1 Modelled noise

The distributions of A_{Hilb} and P_{50} extracted from waveforms with modeled noise injected, similar to the plot on the third row of Figure 9.6 shown earlier, are taken for each of the 320 pulses. Seven exemplary distributions are shown in Figure 9.7 and they are ordered by increasing SNR values from left to right. The top plot shows the amplitude measurable quantity A_{Hilb} and the bottom plot the integrative measurable quantity P_{50}. A Gaussian distribution is fitted to each distribution, and the vertical line represents the value of the measurable quantity from the true pulse, i.e. the same pulse but without noise injected. The Gaussian distribution formula used is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \tag{9.5}$$

where μ and σ are seeded in the Gaussian fits by the mean and standard deviation calculated using the Numpy library [122]. The fit is done with the iminuit library [146].

One can see on this figure that for small SNR values, the mean of the measurable distribution is higher than the true measurable value of the noiseless pulse. For higher SNR, the true value and the mean coincide. Another observation is that the distribution can be fit with a Gaussian distribution

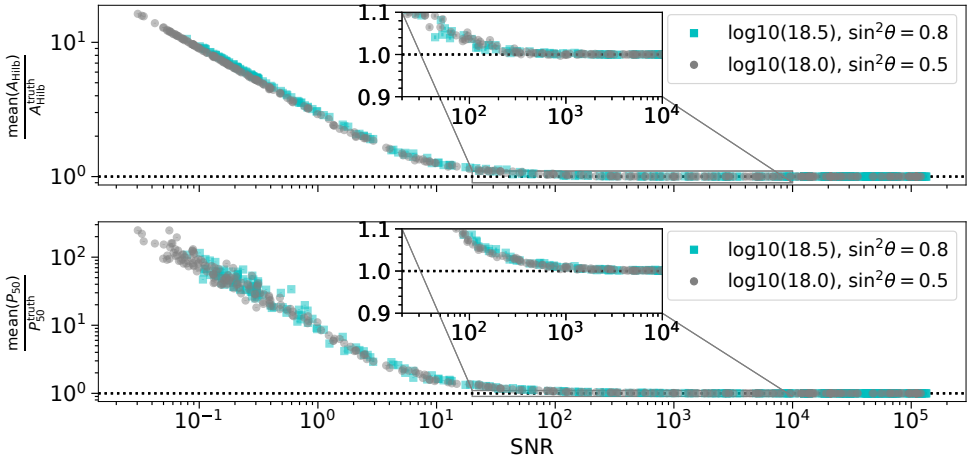


Figure 9.8: Mean of the Gaussian fit on the waveforms divided by its associated true value with modelled noise injected as function of the SNR. On top for the maximum of the Hilbert envelope and on the bottom for the power in a 50 ns window. The different colors represent the two example events.

although at low SNR a skewness of the distribution towards the left side can be observed, especially for the integrative quantity P_{50}. Nevertheless, for the rest of the study, a Gaussian distribution is assumed for every SNR.

Once the histograms above are obtained, the mean and standard deviation of the Gaussian fit from each distribution are plotted as function of the SNR. Figure 9.8 illustrates the behavior of the mean of the Gaussian fit for A_{Hilb} on top and P_{50} on the bottom. The mean is divided by the true value of the pulse for that measurable. The error bars, which are hardly visible because of their small size, represent the uncertainty from the Gaussian fit.

From these plots, one can see that both example events from the MC simulations have a similar behavior, the gray and light blue points overlap. Both A_{Hilb} and P_{50} can be described by three different regimes: a linear decrease for pulses with a associated SNR of approximately less than 1, a constant for pulses roughly above a SNR of 1000, and a transition regime between the two. The behavior of the mean in function of the true value can fitted with a broken power law (see Appendix B), which is used in the next chapter.

Similarly to the previous plots, Figure 9.9 depicts the behavior of the standard deviation of the Gaussian fit for A_{Hilb} at the top and the P_{50} at the bottom. The standard deviation is divided by a constant noise contribution to obtain a more intuitive relative value. The noise contribution of the power quantity N_{P} is calculated in a similar way as N_{A}, describe in Section 9.2.4, but with using the power in the subtraces σ_{P} instead of the RMS.

$$N_{\text{P}} = \overline{\sigma_{\text{P}}} \tag{9.6}$$

The results for A_{Hilb} can as well be categorized into three regimes: a constant regime at around 0.6 times the noise estimation for pulses with a SNR of less than 1, another constant regime at approximately 1 times the noise estimation for pulses with a SNR value of around 300, and a transition regime in-between. The standard deviation of the measurable is described by a sigmoid function in the next chapter (see Section 10.4.2). Concerning the relative standard deviation of the P_{50} measurable, the deviation is constant at around 0.6 times the noise estimation for SNR of less than ∼1 and then rises linearly as a function of the SNR.

It is noteworthy to acknowledge that for all these observations the SNR can be directly translated to the strength of the true pulse due to the definition of the SNR given earlier, Equation (9.3), where

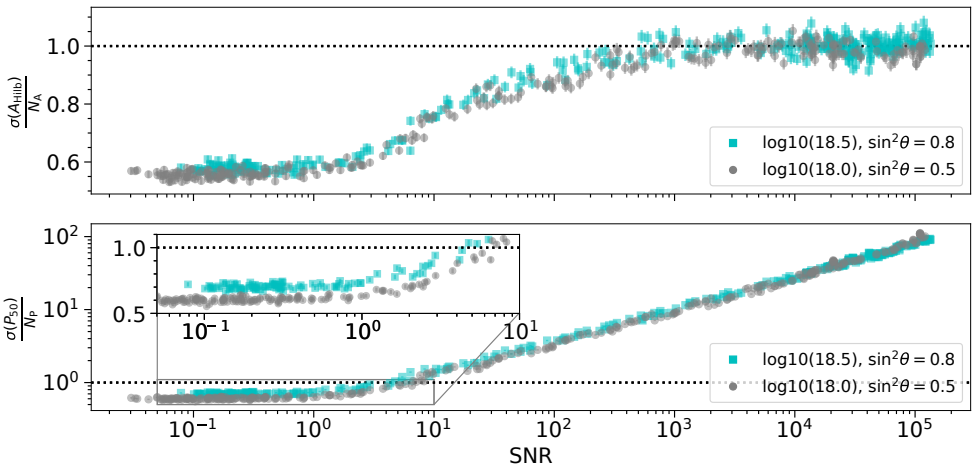


Figure 9.9: Standard deviation of the Gaussian fit on the waveforms with modeled noise injected as a function of the SNR. On top for the maximum of the Hilbert envelope and at the bottom for the power. The different colors represent the two example events.

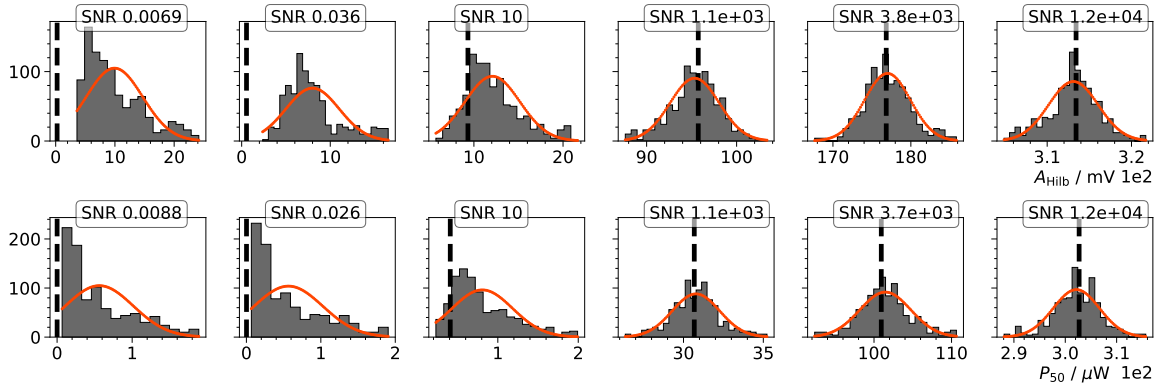


Figure 9.10: Seven distributions showing the value of the surface measurable at the top and the integrative measurable at the bottom for the waveforms with measured noise injected. The black dashed line represents the value of the measurable for the same pulse without noise injected. The distributions are ordered by ascending SNR values from left to right. In these plots the normalization is removed for better visualization, y-axis represents the count number.

the noise is a constant. In fact, for the implementation of the LLH in the next chapter, the function is plotted against the true value for one measurable quantity.

9.3.2 Real noise

The same methodology as before is applied but with noise from the prototype station injected in the traces instead of the modeled noise, again for a 50 ns window.

Figure 9.10 shows distributions of both measurable quantities with the measured noise injected one thousand times. The different distributions represent different initial values of the true pulse, the SNR value is given as reference. The top figure shows the A_{Hilb} quantity and the bottom figure the P_{50} quantity. For the Gaussian fit as well as for the shown histograms, outlier values of more than five standard deviations are removed. The distribution on the right, for high SNR values, are well represented by a Gaussian distribution. Similarly as for the modeled noise, but more dominantly here,

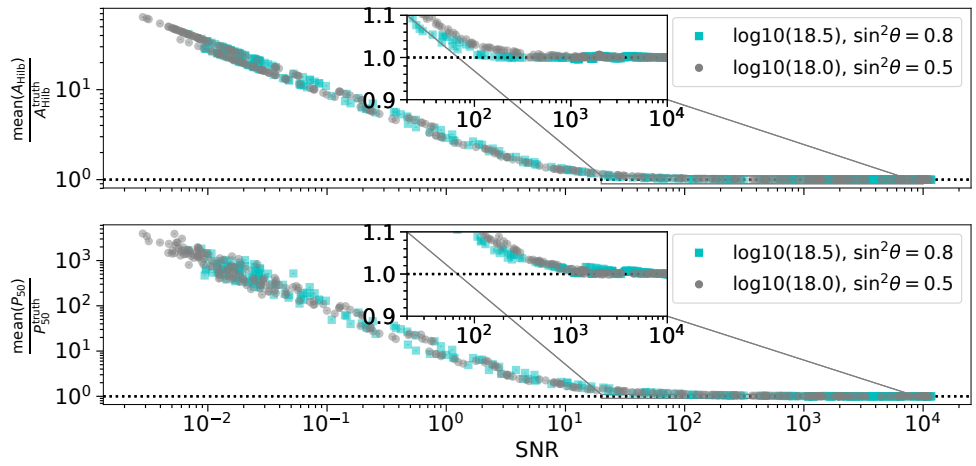


Figure 9.11: Mean of the Gaussian fit on the waveforms with measured noise injected as function of the SNR. On top for A_{Hilb} and at the bottom for P_{50}. The different colors represent the two example events.

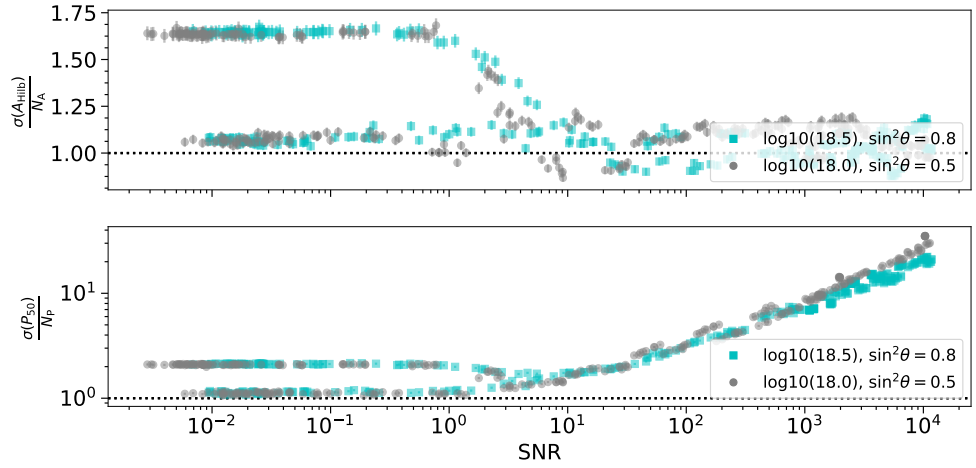


Figure 9.12: Standard deviation of the Gaussian fit on the waveforms with measured noise injected as a function of the SNR. On top for the A_{Hilb} and at the bottom for P_{50}. The different colors represent the two example events.

the three first histograms on the left have a clear skewness towards smaller values.

As for the modeled noise, Figure 9.8 shows the values of the ratio of the mean of the distributions for one measurable quantity over the true value of that measurable quantity as a function of the SNR. The top plot shows A_{Hilb}, and the bottom plot, P_{50}. As for the modeled noise, the same three regimes are present. Since it is plotted over the SNR, taking implicitly the noise level into account, the bounds of the regimes are comparable to the modeled noise results. The scaling of the x-axis, however, differs, and by extension the y-axis as well. From the fit of these curves on can see that the function for the modeled noise is similar to the one for the measured noise.

Once more, Figure 9.12 is the equivalent as Figure 9.9 but for the measured noise instead of the modeled noise. Here the results diverge from the previous results. Below a SNR of about 10, a clear separation between two distributions is seen, for the maximum of the Hilbert envelope as well as for the power. For the maximum of the Hilbert envelope, a reversal in the sigmoid function is seen, where the pulses lower than 1 in SNR have a higher standard deviation than pulses with larger SNR, especially

for the upper distribution. Also for the power, unlike for the modeled noise, the standard deviation is larger than the noise estimation. For the integrative measurable, on the bottom plot, the broken power law is seen again. The regimes are bound by similar values of SNR as for the modeled noise.

9.4 Discussion

First, let's define some terms associated with the different regimes seen in the results section:

- **Small pulse regime**: pulses with SNR values of less than 1
- **Transitional regime**: pulse with SNR values between 1 and 100
- **Large pulse regime**: pulse with SNR values above 100

The bounds in this definition are loose but follow from the anterior observations. The type of distributions seen in the results, the behaviors of the mean, and the standard deviations are discussed in this section for both measurables.

9.4.1 Distribution

The high standard deviation, higher than the noise level, observed when the measured noise is injected, likely arises from the improper Gaussian fit of the measured quantities for the small-to-transitional regimes. This can be seen on Figure 9.10, where the Gaussian fits have a larger width, than a 68% containment, or a more accurate fit, would lead to a smaller estimation of the standard deviation.

One hypothesis for the skewness of the distribution comes from the way the pulses are extracted. As mentioned earlier, the pulse search effectively finds the maxima of the noise in the small pulse regime. The distribution of the extrema in a system can normally be attributed to a Gumbel distribution. This distribution is described by

$$f(z) = \frac{1}{\beta} \exp -(z + e^{-z}), \quad \text{where } z = \frac{x - \mu}{\beta}. \tag{9.7}$$

In the same spirit, the ANITA balloon collaboration reported that the observed amplitude of self-triggered radio pulses at a given frequency, which is the sum of the signal and a thermal noise background, follows a Rician distribution [147]:

$$f(x, b) = x \exp -\frac{x^2 + b^2}{2\sigma^2} \tag{9.8}$$

where b is a shape parameter. Interestingly, a Rician distribution is equivalent to a Gaussian distribution in the limiting case when $b \gg 1$, which, as shown, describes the distribution well for large pulses.

Figure 9.13 shows a Gumbel distribution, a Gaussian distribution, and a Rice distribution with different values of the shape parameter. A skewness to the left, similar to the experimental data, can be observed in the Gumbel distribution as well as in the Rice distribution. One advantage of the Rice distribution over the Gumbel is the possibility to tend towards a Gaussian distribution. Therefore, all regimes can be described by one distribution. Whereas, if using a mix of Gaussian and Gumbel, an arbitrary cut would have to be applied. Although, it might not be so hard to find the change in regime. In the end, it boils down to which distribution fits the data better, and that should be researched with the new TAXI v3.2. For simplicity, the first implementation of a reconstruction technique described in the next chapter assumes a Gaussian distribution everywhere.

9.4.2 Behavior of the noise

A naive expectation, assuming $P_{\text{meas}} = P_{\text{true}} - P_{\text{noise}}$ thus in average $\sigma_{P_{\text{meas}}} = \sigma_{P_{\text{noise}}}$, is that the standard deviation of the measured power quantity, when noise is injected multiple times into the same pulse,

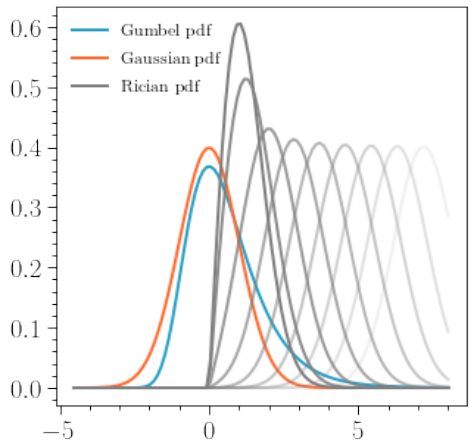


Figure 9.13: Gaussian, Gumbel and Rice distributions with a parameter related to the location at zero and a parameter related to the width at 1. The Rice distribution has different values of the scale parameter between 0 to 8.

would be constant and at the noise level of the trace. This behavior would be seen in the bottom plots of both Figure 9.9 and Figure 9.12 as a straight line: $\sigma(P_{50})/N_{\mathrm{P}} = 1$. However, this is not what is observed. The constant line is instead seen in the high pulse regime of A_{Hilb}, where $\sigma(A_{\mathrm{Hilb}})/N_{\mathrm{A}} = 1$.

Amplitude measurable quantity (maximum of the Hilbert envelope). The three regimes of the A_{Hilb} plot can be explained by the noise contribution either interfering constructively, therefore increasing the pulse, or adding destructively, effectively reducing the pulse at one given time, or one bin. If the pulse is very small, let's say zero, the amplitude will fluctuate around the baseline. If the baseline is at zero, and the absolute values of the Hilbert envelope are taken, then the fluctuations of the noise are only positive, and thus reduced by about half. This describes the small pulse regime. If the pulse is very large, much larger than the noise, then the pulse acts as a baseline for the noise. This baseline is now located far from zero, allowing the noise to freely fluctuate around the pulse, which is then not affected by the Hilbert envelope calculation. This describes the large pulse regime. Then, naturally, there is a transition between the two regimes, which corresponds to the third regime, the transitional regime.

This effect can be seen qualitatively on Figure 9.6 introduced before, where in the plot on the second row, the noisy pulses are almost entirely above the true value for a small pulse, and on both sides of the true value for a large pulse. This is also shown quantitatively on Figure 9.6, which shows the standard deviation of pulses with modeled noise injected, as a function of the strength of the pulse. In this plot, the small pulse regime is equal to slightly more than half the noise level, and the large pulse regime to the pulse level, ergo corroborating the hypothesis. The fact that the standard deviation is not exactly half the noise level for the small pulse regime might be a feature of the way the pulses are found, i.e. per construction, the algorithm often captures an over-fluctuation in the noise.

In summary, if the pulse is smaller than the noise, then its variation can only be half the total contribution of the noise because the possibility of destructive interference is limited for small pulses. For large pulses, the noise can interfere completely destructively without reaching values below the baseline, and the variation in the values of the Hilbert envelope maximum equals the RMS of the noise. In between, there is the transition region where the noise has only a partial destructive behavior.

Integrative measurable quantity – power in the waveform The understanding of the linear rising of the standard deviation as a function of the SNR for the integrative measurable is slightly more complex. To achieve that, one has to resort to the Gaussian propagation of uncertainties.

First, from the power equation Equation (9.1), one can assume the power in the pulse for one bin to be equal to the amplitude squared times a constant c,

$$P^{\mathrm{F}} = c\,(A^{\mathrm{F}})^2. \tag{9.9}$$

where, the superscript F represents the final state of the pulse, i.e. the measured pulse. Therefore, the final state of the pulse comprises the true amplitude A_{T} and a noise amplitude A_{N} contribution,

$$P^{\mathrm{F}} = c\,(A^{\mathrm{T}} + A^{\mathrm{N}})^2. \tag{9.10}$$

Using the Gaussian propagation of uncertainties on Equation (9.9), one obtains

$$\sigma(P^{\mathrm{F}}) = c\sqrt{\left(\frac{\partial P^{\mathrm{F}}}{\partial A^{\mathrm{N}}}\,\sigma(A^{\mathrm{N}})\right)^2 + \left(\frac{\partial P^{\mathrm{F}}}{\partial A^{\mathrm{T}}}\,\sigma(A^{\mathrm{T}})\right)^2} \tag{9.11}$$

The second term vanishes because $\sigma(A^{\mathrm{T}}) = 0$, then solving the differential equation leads to

$$\sigma(P^{\mathrm{F}}) = c\,2\left((A^{\mathrm{N}} + A^{\mathrm{T}})\sigma(A^{\mathrm{N}})\right) \tag{9.12}$$

The value of A_N is unknown. If replaced by its average contribution, which is $\sigma(A_{\mathrm{N}})$, the equation becomes

$$\sigma(P^{\mathrm{F}}) = 2\,c\,\sigma(A^{\mathrm{N}})^2 + A^{\mathrm{T}}\,\sigma(A^{\mathrm{N}}). \tag{9.13}$$

The first term, $2\,\sigma(A_{\mathrm{N}})^2$, describes the flat region at low SNR, where the noise contribution dominates. The second term, $A_{\mathrm{T}}\,\sigma(A_{\mathrm{N}})$, explains the linear rise of the uncertainty on the power as function of the strength of the pulse A_{T}.

Replacing $\sigma(A_{\mathrm{N}})$ by the definition of the noise that is used in this chapter, σ_{A}, and substituting c for $\Delta t/Z\,t_{\mathrm{tot}}$, where for one bin $\Delta t = t_{\mathrm{tot}}$, then

$$\sigma(P_1^{\mathrm{F}}) = \frac{2}{Z}\,(\sigma_{\mathrm{A}}^2 + A_1^{\mathrm{T}}\,\sigma_{\mathrm{A}}). \tag{9.14}$$

Equation (9.14) is valid for one bin only, its extension for multiple bins is

$$\sigma(P_n^{\mathrm{F}}) = \frac{\Delta t}{t_{\mathrm{tot}}}\sum_{i}^{n}\sigma(P_i^{\mathrm{F}}). \tag{9.15}$$

One can convince himself of that by using the Gaussian propagation of uncertainties on the series of the power calculation formula. Under the assumption that, if the window is small enough, the amplitude of each bin are equal $A_{\mathrm{T}} = A_1 = A_2 = ... = A_n$, then

$$\sigma(P_n^{\mathrm{F}}) = n\,\frac{\Delta t}{t_{\mathrm{tot}}}\sigma(P_1^{\mathrm{F}}). \tag{9.16}$$

This is tested with the data from the results section. Figure 9.14 depicts, in gray, the standard deviation for the power of the waveforms with noise injected in a 10 ns window. In blue, Equation (9.14) is plotted, and in orange Equation (9.16). These equations do not take changes of regimes into account, nor non-Gaussian distributions, or any bin-to-bin variations, but still in a first order, they describe well the flat and linear behavior of the standard deviation that is seen in the simulations.

Furthermore, if we look back in Section 9.1 at how the noise is injected in the signal trace, by the addition of signals, these results are not so surprising anymore.

Remaining is the fact that in the low pulse regime $\sigma(P_{50})/N_{\mathrm{P}} < 1$. This might be partially caused by how the average noises N_{A} and N_{P} are calculated. The leaking of the pulses might overestimate the average noise level. This could be improved by changing the simple box filter to a causal filter which would not create oscillation before the pulse. This, however, cannot explain it all, because when the size of the window for the power is reduced to 10 ns, in the low pulse regime, $\sigma(P_{50})/N_{\mathrm{P}} \approx 1.1$, whereas it was ~0.6 with a 50 ns. Changing the power window size involves no change in the calculation of the noise level. The figure illustrating the deviation of P_{10} can be seen in Appendix B. This lead to an

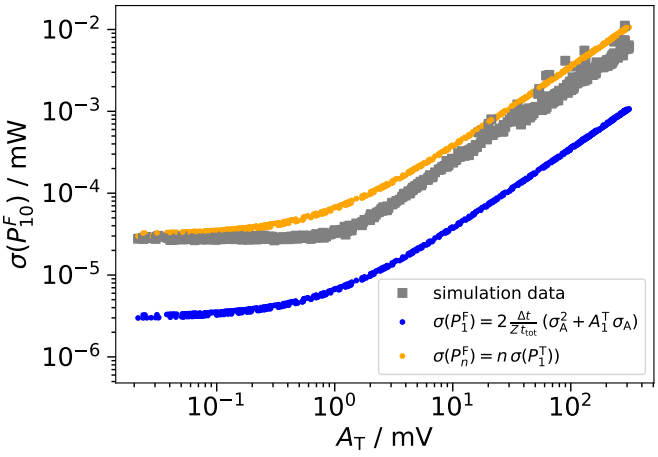


Figure 9.14: Standard deviation of the power of the waveform with modeled noise injected in relation to the Gaussian propagation of error calculation for the power in a 10 ns window. Thus, for the orange line, n=10.

interesting hypothesis, that the size of the pulse window, where the integration is done, affects the noise contribution of the signal. When analyzing the ideal size for the subtraces method for noise calculation (Section 9.2.3) the minimal window size, for stable estimation, is found to be at 64 ns, arguably this could correspond too to the minimal size of the pulses window for the standard deviation to be on average equal to the noise contribution.

Note once more that the N_A and N_P derived here are utilized for better intuition of the noise behavior, they are not used during the parametrization for the LLH method in the next chapter. Related works on the influence of noise on radio measurements can be found in [66], [148].

In conclusion, in this chapter, the noise behavior is studied for two different measurable quantities, A_{Hilb} and P_{50}, and two types of noise, modeled and measured for the prototype station. The distribution of the measured quantities with noise injected is described here via a Gaussian function, and the mean and standard deviation of these fits are plotted against the SNR, where the SNR can be directly related to the maximum amplitude from the true pulse. In all these plots, three different regimes can be observed which are categorized as the small pulse regime, the transitional regime, and the large pulse regime. All these functions are fitted, and the corresponding fits can be found in Appendix B. This chapter lays down the building blocks for the reconstruction of the energy and the X_{max} using a LLH minimization that is described in detail in the next chapter.

CHAPTER 10

Energy and X_{max} Reconstruction Technique

Unlike particle detectors, where the energy can be reconstructed accurately with a one-dimensional LDF, as e.g. for the IceTop tanks introduced in Chapter 4, the reconstruction of the energy with antenna arrays proves more complex due to the asymmetric radio emission around the shower axis (see Chapter 3). This asymmetry arises mainly from the interference of the geomagnetic emission and the Askaryan emission, whose polarisations differ. Reconstructing the energy from the radio emission analytically, using a 1D or 2D LDF, is thus complex, necessitates several fitting parameters, and is prone to large uncertainties.

A method developed by LOFAR [85] and further refined and tuned by other experiments [88], [149]–[151], which uses specific MC simulations, is currently the state-of-the-art reconstruction in this field. In addition to an accurate energy reconstruction, this technique also leads in X_{max} reconstruction with radio, a variable priced for mass composition studies, delivering accuracy comparable to the fluorescence telescopes (see Section 3.3). An additional feature of this method is that it implicitly takes all the peculiarities of the footprint into account without necessitating extra corrections. For the experiments which used it (AERA, LOFAR, and Tunka-Rex), the resolution of the energy reconstruction is of the order of 10% to 14%, and for X_{max} between 16 to 35 g/cm^2 , depending on the cuts and experiment-specific implementation of the technique. All these experiments use the 30-80 MHz frequency band, have strict quality cuts, and require at least five antennas with a signal above threshold.

In this chapter, the technique is boldly applied to the three antennas of the prototype station at the South Pole. As it is shown and discussed later on, there are some limitations to this application, but more importantly, this technique is applied for the first time in the 80-300 MHz bandwidth. This work is intended to be a proof-of-concept that this type of reconstruction can work with the higher frequency band of the IceCube's SAE. Furthermore, the software algorithms developed within the IceCube framework can be easily re-used as a first implementation for a more intricate reconstruction in the future. In this first implementation, efforts are put into the transition from a χ^2 to a LLH minimization, to describe the noise accurately. Nonetheless, the χ^2 methods are also implemented as a safety check to ensure the proper working of the LLH, the complete parametrization of the noise is, however, only done in regards to the LLH at a defined frequency band, whereas some assumption are made for χ^2 methods.

This chapter begins with an overview of the method, then describes the event selection and the "initial seed" reconstruction. Subsequently, the method is implemented using the different measurable quantities in the radio waveform and evaluated with a Monte Carlo to Monte Carlo (MC-to-MC) comparison. Finally, the method is applied to the real data recorded on the field.

10.1 Overview of the method

There is two philosophies for when doing an analysis using a physical detector. These are called forward folding and unfolding. The first one entails the addition of the system response and effects in the

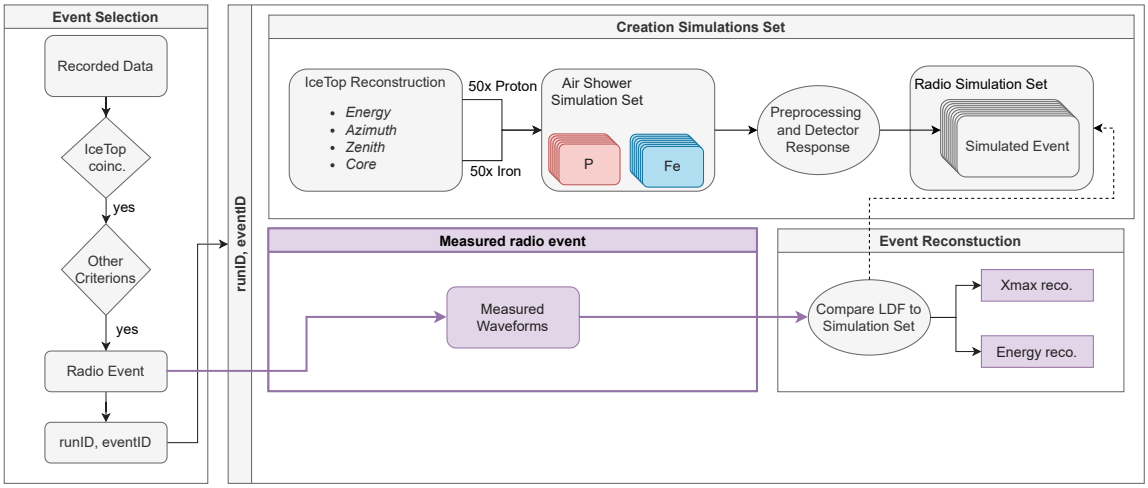


Figure 10.1: Overview of the reconstruction method using data. At the final step, a reconstructed value of $X_{\max}$ and of the energy for one recorded air shower are obtained.

simulated data, whereas the second one removes the system response from the measured data. Both methods have their advantages and complexities. For this analysis the forward folding is applied as it is a natural transition, after working on the detector, calibration, and design of the system, to translate the simulations into measured data.

The method entails the creation of a unique set of Monte Carlo simulations for each air shower event recorded by the radio antennas. This implies that a first reconstruction of the energy, direction, and core position must be estimated in order to create the simulations. Thus the IceTop standard reconstruction presented in Chapter 4, with the parameters shown in Appendix A, is used as initial reconstruction. For each simulation set, protons and iron nuclei are used as primary particles to cover the extended $X_{\max}$ phase-space. The simulated electric field created by the air shower at each antenna position is then convolved with the detector response to be comparable to the measurements by the antennas on the field.

Afterward, a certain measurable quantity within the waveform is chosen and extracted from each simulated radio waveform as well as from the recorded radio data. In the end, this chosen measurable quantity from the recorded event is compared to each simulation of its unique simulation set using a minimization technique.

Before reaching this last step, the technical performance is evaluated using a MC-to-MC comparison. The idea is to extract one simulation from the set, inject noise, and compare it to the rest of the set. This extracted simulation thus mimics real data. This step is applied for each simulation in the set and provides an estimate of the accuracy of the reconstruction for that specific event.

Figure 10.1 gives a visual representation of the steps of the reconstruction using real data where the boxes in purple show the steps which are specific for the reconstruction with data. For the estimation of the accuracy, these steps are slightly modified as will be shown later.

10.2 Events and initial reconstruction

The initial events sample is created out of the first 16 showers that are found in coincidence with the IceTop tanks and the scintillation detectors. Their identification as air showers was previously done on an event-per-event basis. Further improvements related to the scope of the quality cuts, for example the SNR level required, the position of the core, or the noise level of the trace are currently ongoing, but outside the scope of this work. Nevertheless, some insights about useful quality cuts increasing the

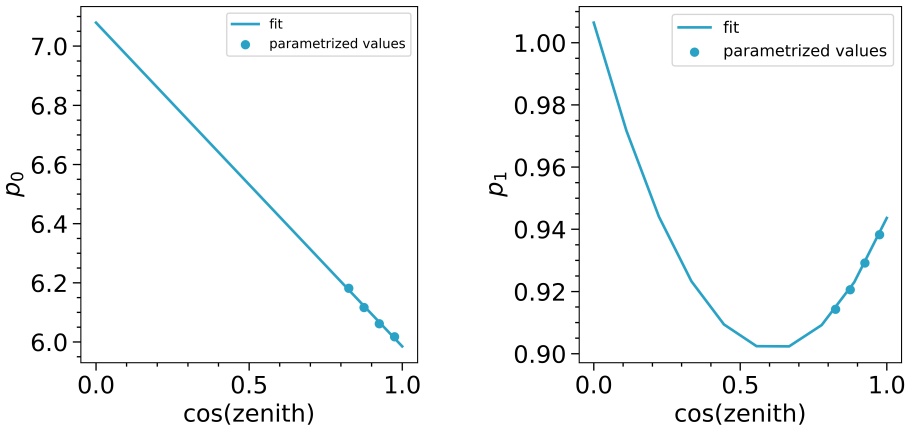


Figure 10.2: Extrapolation of the fit parameters p_0 and p_1 relating the energy to S_{125}. Those two fits are the ones used for energy extrapolation.

accuracy of the X_{max} and energy reconstruction using this method are given in the discussion section at the end of this chapter.

10.2.1 IceTop reconstruction

The Monte Carlo simulations for a specific air shower were initiated using the Laputop reconstruction algorithm applied to the IceTop data associated with the event, with the parameters of the air shower including its energy, direction, and core position serving as the "initial seed". Further information regarding Laputop can be found in Appendix A, with additional details provided in Chapter 4.

However, it should be noted that many of the showers recorded by the radio antennas do not fall within the phase-space for which IceTop is calibrated. In particular, the relationship between S_{125} and energy is only calibrated for zenith angles θ of $\cos\theta$ between 0.8 and 1.0. To estimate the energy for more horizontally-oriented EAS, a simple extrapolation of the known parametrization is utilized.

The relation between the energy and the estimator S_{125} is linear and can be described by

$$\log_{10}(E) = p_1 * \log_{10}(S_{125}) + p_0 \tag{10.1}$$

where p_0 and p_1 are the two fit parameters. The values of p_0 and p_1 for four different bins in zenith are taken from the published energy spectrum in [152] and are represented by the points on Figure 10.2. Thereafter, a polynomial fit of either first or second order is applied to the points in order to obtain the value of p_0 and p_1 for the whole range of the zenith angle.

Figure 10.2 shows on the left the parameterized values and the fit for the parameter p_0 and similarly on the right for the parameter p_1. The shown fits are linear in p_0 and quadratic in p_1 and are used as extrapolation. Naturally with only four points relatively grouped together, different fits result in similar agreements. To quantify the effect of the fit, the two parameters are derived using every combination of linear and quadratic functions.

These combinations are shown in Figure 10.3. The y-axis represents the ratio of the energy obtained with a certain extrapolation to the energy value defined by the "standard extrapolation", shown in Figure 10.2. The red points represent the status quo option which simply uses the lowest parameterized value for the whole zenith range. This option gives the lowest estimate on the energy, which is about half the energy from the standard extrapolation estimate at a cosine of the zenith of about 0.45. The highest estimate of energy is given by fitting both parameters with a quadratic function. This gives an about two and a half times higher energy estimation for the lowest zenith angles in the shower sample. This extrapolation is not optimal but is assumed to be enough for an initial reconstruction. The final

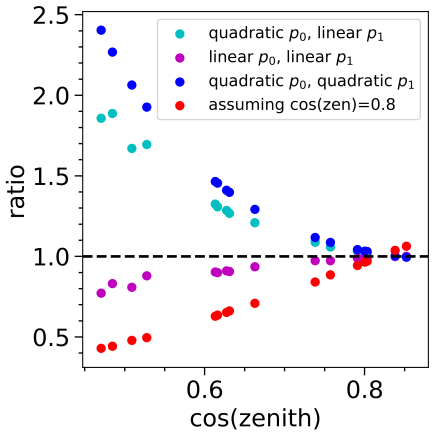


Figure 10.3: Ratio of the energy values reconstructed using different order polynomials fits over the energy reconstructed by the one used. The orange one is without extrapolation, simply using the value associated with the smallest parameterized value.

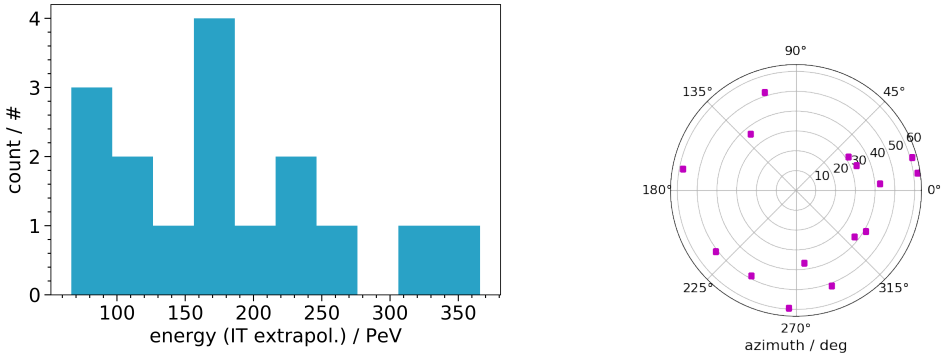


Figure 10.4: IceTop reconstruction using Laputop and the extrapolation for the energy on the left and for the zenith and azimuth angles of the incoming direction on the right of the 16 showers sample selection.

reconstruction of the energy will be obtained with the radio array from the method described in this chapter.

As for the core and direction reconstruction, IceTop reconstruction is taken as well. The core reconstructed by IceTop is more accurate than the one from the scintillation detector array purely due to the larger size of the array. For the direction, the reconstruction from the three detectors is similar by construction, i.e., it is required for the event in the radio data to be flagged as an air shower, to have a timing of the pulses agreeing with the arrival direction of the IceTop tanks and of the scintillation panels. The reconstructed energy and arrival direction from Laputop for the 16 showers is shown on Figure 10.4 and the reconstructed core positions in Figure 10.5.

The uncertainties on these reconstructions are estimated using IceTop simulations of 2012. The reconstruction of the simulated air shower is taken only when its reconstructed core is located inside the array and its reconstructed energy is between 90 and 100 PeV. A Gaussian fit is done on the histograms of the reconstructed value subtracted by the individual true value for all the parameters (zenith, azimuth, core position) of all the shower geometries from the sample. The results of this fit are shown in Figure 10.6, where the points represent the mean of the Gaussian fit and the error bar is the standard deviation. More details on the procedure can be found in Appendix B.

The reconstruction of the arrival direction shows a good resolution for the shower sample with

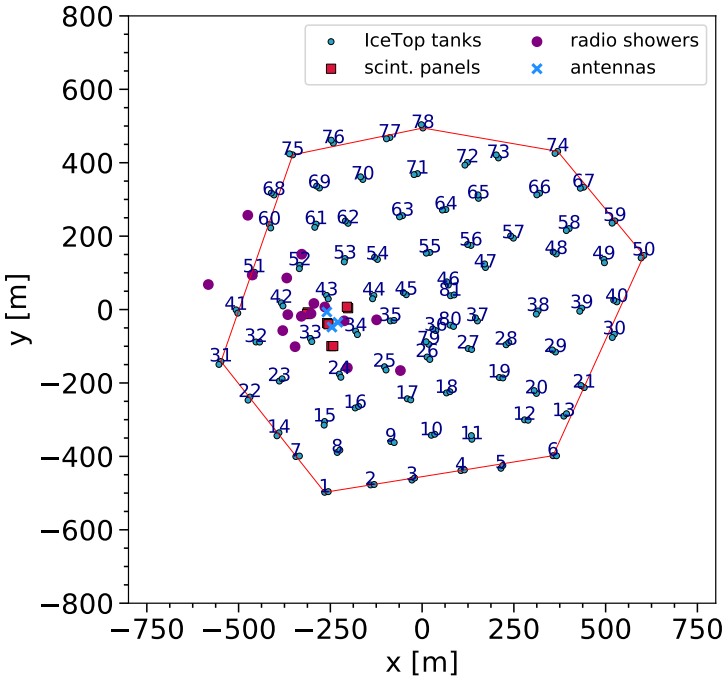


Figure 10.5: IceTop reconstruction using Laputop for the core positions of the 16 showers sample. The numbers are the associated IceTop tanks numbers.

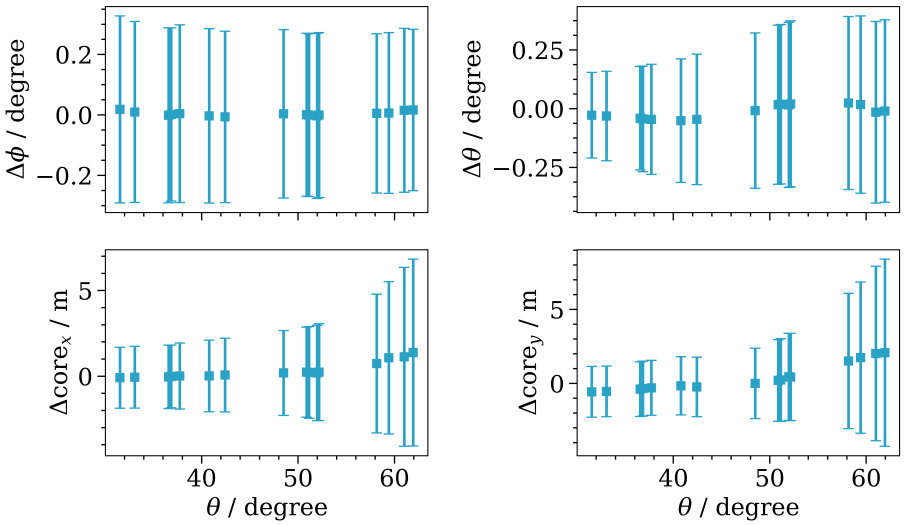


Figure 10.6: Accuracy of the Laputop reconstruction for the air showers sample. The points are the mean of the Gaussian fit and the error bars the standard deviation.

Table 10.1: Example library steering file for CORSIKA

RUNNR	0
EVTNR	1
SEED	30168043 0 0
SEED	31168043 0 0
SEED	32168044 0 0
ERANGE	135.01(...)e+6 135.01(...)e+6
PRMPAR	14
THETAP	33.0596 33.0596
PHIP	105.5091 105.5091
THIN	1e-06 135.01(...) 0.0
(...)	
ATMFILE	.../atmosphere/atmos_runId134826_eventId53455709.txt
DIRECT	.../coreas/runID_134826_eventID_53455709/proton/000000/

only a fraction of a degree of uncertainty, being constant at about 0.25° for the azimuth reconstruction and ranging from roughly 0.25° to 0.35° for the zenith reconstruction, increasing with higher zenith angle. Similarly, for the core reconstruction, the accuracy worsens for more horizontal showers, with a resolution ranging from roughly 2 m to 7 m for both x and y direction. The bias of the shower incoming direction is close to zero, whereas for the core, a bias, increasing for the more inclined showers is observed, but smaller than 1 m except for zenith angels over 50°.

10.3 Monte Carlo simulations using CORSIKA

10.3.1 Simulations setup

A special set of Monte Carlo simulations is created for the energy and X_{max} reconstruction technique, using the same parameters of the library mentioned in the previous chapter except for the energy and direction. The Monte Carlo simulations are also created with CORSIKA v7.7401 [139] using CoREAS v1.4 [140], Sibyll 2.3d [141] and fluka2011.2x [142]. An example of a steering file for one proton simulation associated with one event is shown in Table 10.1, where the (...) represents the omitted parameters which are exactly the same as in the steering file in Chapter 9. The complete steering file can be found in Appendix A.

In addition to the direction and energy, the atmospheric parameters are changed compared to Chapter 9. In Chapter 9, a standard atmosphere at the South Pole is used, here the specific atmosphere at the day of which the event was recorded is used instead. The atmosphere is created based on the data from the global data assimilation system (GDAS). GDAS is a database containing measurements from several diverse weather instruments located all around the world [153]. A script called "gdastool" included in the CORSIKA suite fetches the data at a certain time and place and creates a look-up table with a 5-layer atmosphere model that can be utilized by CORSIKA during the simulation. It was shown for LOFAR that a change from the standard atmosphere to the specific GDAS results in an average difference of 2g/cm^2 for the X_{max} reconstruction and up to 15g/cm^2 during extreme weather [154].

As in the previous chapter, the air shower emission is simulated at the locations of an 8-arm star-shaped pattern shown in Figure 9.1. However, this time, the signal is interpolated with radcube at the location of the antennas on the field relative to the shower core [136], [144].

10.3.2 Simulations set

For each of the 16 air showers in the sample, one simulation set is created with the same steering file except for the primary particle. For each event, 50 simulations are produced with a proton as primary

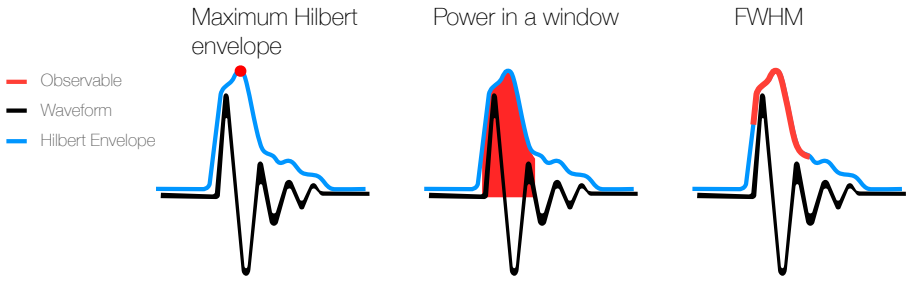


Figure 10.7: Illustration of the different measurable quantities studied. The maximum of the Hilbert envelope and the FWHM are considered as amplitude quantities and power as an integrative quantity.

inducing the air showers and 50 with an iron nucleus. The resulting 100 showers is expected to be enough to cover the $X_{\max}$ phase-space for one event.

10.3.3 Processing of the electric field with radcube

The electric field is then processed with radcube in the same way as in the previous chapter (Chapter 10). In brief, it is convolved with the antenna and electronics response and then filtered between 80 and 300 MHz. No noise is added to the signal at this point. The waveforms from the Monte Carlo simulations are now comparable to the measured data, except for the noise contribution. At the end of this step for each Monte Carlo simulation, a total of six waveforms are obtained, through two polarisation channels per antenna and three antennas in total. This totals 600 waveforms per set.

10.4 Methodology

The simulations set is created for each air shower of the event sample, and processed to resemble the measured data. From there, the identification of pulses in the simulations is done using the procedure described in Section 9.2.5.

10.4.1 Measurables evaluated and minimization used

Similarly to the previous chapter, amplitude measurable quantities and integrative measurable quantities are studied. In addition to the maximum of the Hilbert envelope, the FWHM measurable is added as amplitude measurable, which is a natural extension from the maximum. It is defined by the value of the signal for a certain interval of points around the signal maximum. The idea behind this new quantity is to capture information about the phase and the spectrum of the pulse, on the grounds that the pulse shape carries information about these by the way they add up together to create a pulse. The name is motivated by the initial implementation which used a 10 ns window, corresponding to the average full width at half maximum (FWHM) of the pulses in the shower sample. This is also the reason motivating the study of a 10 ns window. In addition to this short window, a substantially larger window of 50 ns is also implemented and studied. Figure 10.7 pictures the different measurable quantities that are studied with this reconstruction technique.

For each of the measurables a χ^2 minimization in the general form of

$$\chi^2 = \sum_i^n \left(\frac{x_i - \mu}{\sigma} \right)^2 \tag{10.2}$$

is implemented. In addition to the χ^2 minimization method, which is used in all the experiments referenced in the introduction, a more general LLH minimization is developed for each measurable. The χ^2 function is actually a specific case of the LLH function under the assumption that the probability

distribution is Gaussian and that its standard deviation σ is constant. This can be shown by taking the definition of a likelihood function

$$\mathcal{L}(\Theta) = \prod_i^n P(x_i;\Theta) \tag{10.3}$$

which is a product of the probability density functions $P(x_i;\Theta)$ for all the data points x_i, where Θ is the variable to be minimised. The likelihood function can be transformed to a LLH by applying $-2\ln$ (or just ln depending on the normalization) on both side and becomes

$$-2\ln\mathcal{L} = -2\ln\left(\prod_i^n P(x_i;\Theta)\right) = -2\sum_i^n \ln P(x_i;\Theta). \tag{10.4}$$

If the probability density function $P(x_i;\Theta)$ is Gaussian, and the general variable Θ replaced by the mean μ, then

$$-2\ln(\mathcal{L}(\mu)) = -2\sum_i^n \ln\left(\frac{1}{\sigma\sqrt{2\pi}}\right) + \ln\left(\exp\left(-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2\right)\right). \tag{10.5}$$

For a constant standard deviation σ, the first term can be ignored. the equation becomes

$$-2\ln(\mathcal{L}(\mu)) = \sum_i^n \left(\frac{x_i-\mu}{\sigma}\right)^2. \tag{10.6}$$

which is the definition of the χ^2 shown in Equation (10.2). Finally, to obtain the best value of μ, the χ^2 or the LLH are differentiated with respect to μ and equated to zero to find a minimum (or maximum). All the equations described in the next section are derived from these basic equations, where μ is substituted by a factor multiplied by the measurable quantity extracted from the simulations and x corresponds to the measurement from the data.

10.4.2 Log-likelihood and χ^2 formulas

In this section, all the formulas used for either a χ^2 or a LLH minimization are introduced. In the cases of the χ^2 minimization, naive assumptions are made. One is that, when the pulses are large enough, because of the constructive and destructive interference nature of radio waves, on average the noise contribution for the maximum of the Hilbert envelope equals zero. On one side, from the previous chapter, Figure 9.8, one can see that is true above a SNR of ~300. On the other side, regarding the prototype station, with only three antennas measuring, it is reasonable to assume that most of the signals are located in the large pulse regime. Thus, the maximum of the Hilbert envelope of the simulated waveforms (without noise) will be compared directly with the maximum of the Hilbert envelope measured (or simulated with noise added for the MC-to-MC comparison). Per extension, this reasoning applies to the FWHM method as well.

For the χ^2 with the integrative quantity (power), the reasoning is slightly different. The goal is to have something comparable to the experiments that use it, namely LOFAR and AERA. However, they opted for the unfolding approach, which makes the 1:1 comparison complicated. In brief, they removed the noise contribution directly in the unfolding method and then use the χ^2 method. Here, a noise contribution σ_P, calculated with the power in the noise of the same trace containing the signal, using the subtraces method (see Section 9.2.3), is removed from the data before comparison with the noiseless simulation.

For the LLH method, the whole parametrization of the noise is achieved based on the plot from the last chapter. The mean and the standard deviation are directly parameterized over the true values for the quantity examined. The parametrization values for the maximum of the Hilbert envelope as well as for the power with a 50 ns and a 10 ns window can be found in Appendix B. The parametrization is modeled and measured individually for both types of noise.

Maximum of the Hilbert envelope, is derived with a χ^2 function in the form

$$\chi^2_{\max} = \sum_{i=1}^{6} \left(\frac{\varepsilon_i - f \cdot \varepsilon_{\mathrm{MC}_i}}{\sigma_i} \right)^2 \tag{10.7}$$

where ε_i is the maximum of the Hilbert envelope in the data for the waveform i, $\varepsilon_{\mathrm{MC}_i}$ is the same quantity extracted from the simulated waveform i, which is taken from only one Monte Carlo simulated shower, f is a scaling parameter related to the energy, and σ_i is the noise estimation in the waveform i, calculated as in Section 9.2.3. The index i represents the channel in which the data is recorded, e.g. for polarisation 1 in antenna 2 this corresponds to a $i = 4$. In total, there are six channels.

The formula derived for the LLH of the Hilbert envelope's maximum is

$$-2\ln\mathcal{L}_{\max} = \sum_{i}^{n=6} \left(\ln\left(2\pi\sigma'^2(f, \varepsilon_{\mathrm{MC}})\right) + \left(\frac{\varepsilon_i - \varepsilon'_{\mathrm{MC}_i}(f, \varepsilon_{\mathrm{MC}})}{\sigma'_i(f, \varepsilon_{\mathrm{MC}})} \right)^2 \right) \tag{10.8}$$

where $\varepsilon'_{\mathrm{MC}}$ and $\sigma'(\varepsilon_{\mathrm{MC}})$ are the modification of the noiseless MC waveforms to account for the noise behavior and derived from the results of the previous chapter. The modified variable ε' is described by a sigmoid function of the form

$$\varepsilon'_{\mathrm{MC}} = \frac{a}{(1 + b\exp(-\alpha f \cdot \varepsilon_{\mathrm{MC}}))} + s \tag{10.9}$$

where a, b, α, and s are the parametrisation constants. Whereas, the variation $\sigma'(\varepsilon)$ is defined by a broken power law as

$$\sigma'(\varepsilon_{\mathrm{MC}}) = a\left(1 + \left(\frac{f \cdot \varepsilon_{\mathrm{MC}}}{x_0}\right)^k\right)^{(b/k)} \tag{10.10}$$

where a, b, k, and x_0 are different parametrisation constants. The values of these constants and the fits are Appendix B.

FWHM, is obtained in a similar fashion as for the maximum of the Hilbert envelope but in multiple bins around the maximum, the χ^2 formula is given by

$$\chi^2_{\mathrm{fwhm}} = \sum_{j}^{\mathrm{bins}} \chi^2_{\max}. \tag{10.11}$$

The LLH equivalent is in the form

$$-2\ln\mathcal{L}_{\mathrm{fwhm}} = \sum_{j}^{\mathrm{bins}} -2\ln\mathcal{L}_{\max}. \tag{10.12}$$

For this method, the same parametrization of the noise is used as for $-2\ln\mathcal{L}_{\max}$.

Power, is defined as the average power of the pulse (see Equation (9.1)) in a certain window with an average power of the noise σ_{P} subtracted. The χ^2 becomes

$$\chi^2_{\mathrm{power}} = \sum_{i}^{n=6} \left(\frac{(P_i - \sigma_{p_i}) - f^2 \cdot P_{\mathrm{MC}_i}}{\sigma_{p_i}} \right)^2 \tag{10.13}$$

where P_i is the power in the waveform i from the data and P_{MC_i} is the power in the waveform i of one simulation. The noise contribution σ_{p_i} is equivalent to the average power in the subtraces (explained in Section 9.2.3).

The transformation into the LLH takes the form

$$-2\ln\mathcal{L}_{\mathrm{power}} = \sum_{i}^{n=6} \left(\ln\left(2\pi\sigma'^2(f^2, P_{\mathrm{MC}})\right) + \left(\frac{P_i - P'_{\mathrm{MC}_i}(f^2, P_{\mathrm{MC}})}{\sigma'(f^2, P_{\mathrm{MC}})} \right)^2 \right) \tag{10.14}$$

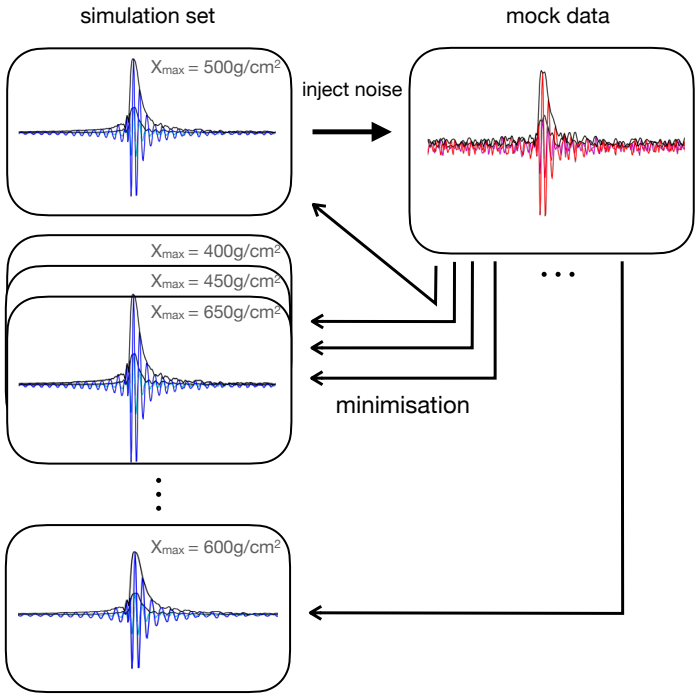


Figure 10.8: Schematics demonstrating the creation of mock data for the MC-to-MC comparison.

where the parametrization of both $\sigma'(f^2, P_{MC})$ and $P'_{MC}(f^2, P_{MC})$ are derived from the previous chapter as well, and both are described by a broken power law in the form of Equation (10.10) with different constant values.

The parameter f that is minimized in all those equations, is a scaling parameter that relates the size of the pulse to the energy, which can be approximated by

$$E_{reco} \approx f\, E_{MC} \tag{10.15}$$

where E_{MC} is the primary energy of the best-fitting simulated air shower, and E_{reco} is the estimated total energy contained in the recorded air shower using this reconstruction method.

10.5 Monte Carlo to Monte Carlo comparison

Before using the data with the previous methodology, the methods are evaluated with a MC-to-MC comparison to evaluate their respective efficacy.

One simulation is taken out of the set and noise is injected into order to mimic a recorded waveform. This simulation is then relabelled with mock data and fed into the reconstruction algorithm. The previous step is repeated for every simulation in the set, thereby reconstructing a specific event $n-1$ times with air showers of $n-1$ different X_{max} values. In reality, the reconstruction is applied n times, under the assumption that the injection of the noise alters enough of the waveform to not bias the reconstruction.

Figure 10.8 explains the creation of mock data by the injection of noise in one waveform in the set. The blue waveforms represent multiple air shower simulations of one set for a single channel. The red waveform represents the mock data waveform, i.e. the waveform from which energy and X_{max} will be reconstructed. The MC-to-MC comparison is executed for all the different measurables and with the modeled and the measured noise.

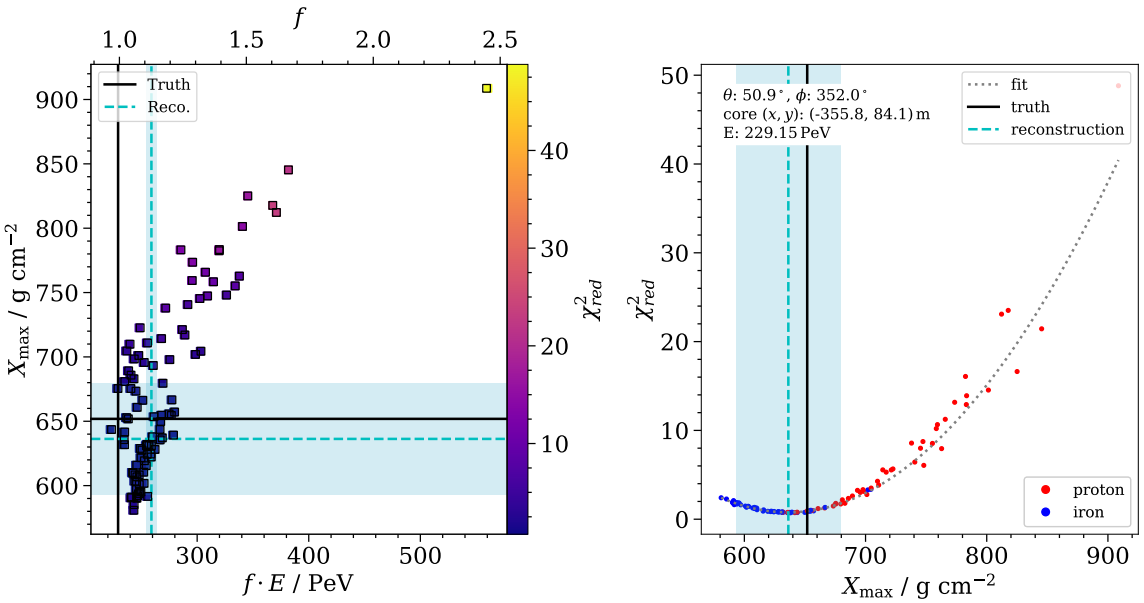


Figure 10.9: Reconstruction of one mock data air shower, with injected modeled noise, using the χ^2_{max} method. In both figures, the black line represents the true value, the blue dashed line is the reconstructed value, and the blue band is the uncertainty. The uncertainty of X_{max} is defined by one χ^2 value from the minimum, and for the energy, it is the fit uncertainty of f. The mock data is created from one air shower simulation initiated by a proton. (Associated with the event: runid 134630, eventid 53455709.)

10.6 Results for Monte Carlo to Monte Carlo comparison

First, the results from the MC-to-MC comparison are laid out and in the next section, the method is applied to real data. The exemplary plots shown in this section are for modeled noise. The end result is shown for the measured noise as well.

10.6.1 Reconstructing one event

For one mock data air shower, each time the mock data is compared to one simulated air shower in the set, a value of X_{max} , reduced χ^2 , and f is obtained. Once compared to all the simulated air showers in the set, an array of 100 X_{max} , reduced χ^2 , and f are created. The reconstructed value for energy is extracted by taking the f value associated with the smallest χ^2 value in the array, then this value is multiplied by the primary energy E_{CR} of that simulated air shower. To reconstruct X_{max} , the reduced χ^2 values are plotted against the X_{max} values which result in a parabolic shape. A parabola is fitted to the 50 points with the lowest χ^2 values. The minimum of the parabola is found and the X_{max} value is taken at this location.

Figure 10.9 illustrates an example of one mock data air shower compared to all the MC simulations in one set. The reconstructed X_{max} is shown with a blue dashed line and is located at the minimum of the parabola, shown by the gray line. The blue bands are the uncertainties on the reconstruction. The uncertainty of X_{max} is defined by the region within one reduced χ^2 above the minimum and the uncertainty of f is the statistical error associated with f during the minimization. The value of the lowest reduced χ^2 is equal to 0.72, close to the expected value of a reduced χ^2 of 1 at the minimum. The true value of X_{max} is located within the uncertainty band of the reconstruction, whereas the true energy is located outside the energy reconstruction uncertainty band.

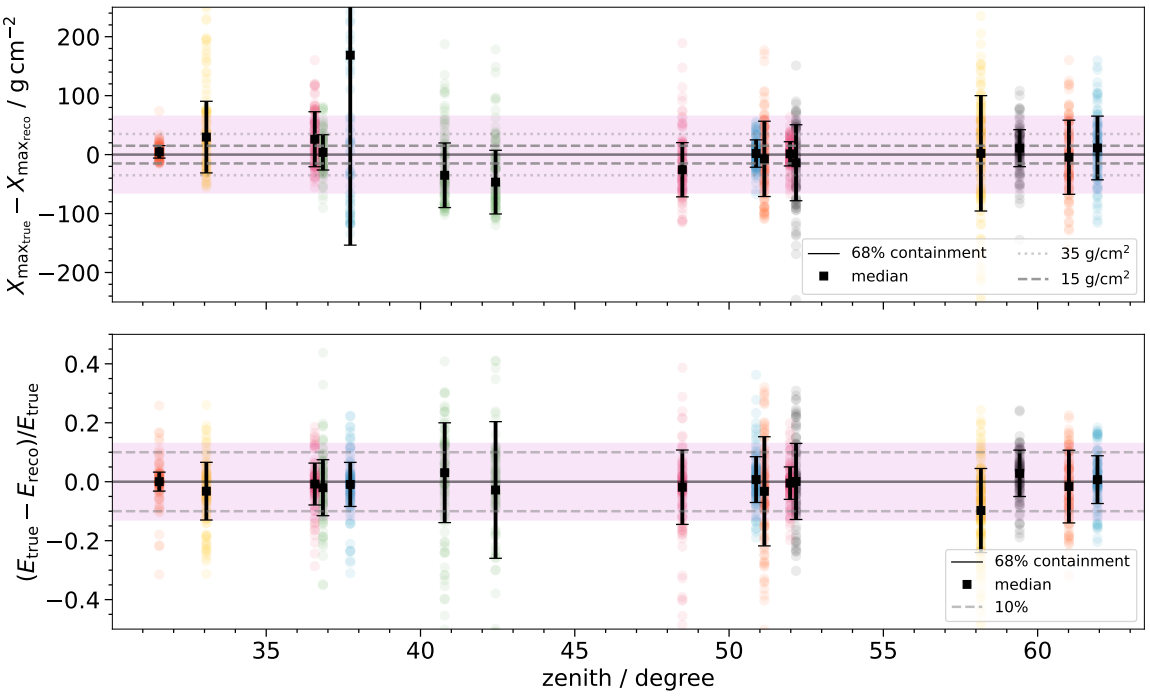


Figure 10.10: Bias and precision of the reconstruction on an event-per-event basis for all events in the shower sample using the χ^2_{max} method with modeled noise injected for the mock data. The dashed line represents the resolution obtained by other experiments and the purple band the precision of this method for all the events in addition to the scatter points of the individual reconstructions. The error bars represent the median of the values and a 68% containment.

10.6.2 Estimating the resolution

The next step is to obtain the reconstructed values, for each of the simulated air showers in the set. In other words, each of the simulations in one set is in turn used as mock data, and the step above is applied to each of them.

Figure 10.10 shows the results of this procedure applied to each event in the shower sample. The points are the individual mock data reconstruction, the black error bars represent the median for the event, and the 68% containment. The events with a zenith angle below 40° have a better resolution compared to the more inclined events. The energy reconstruction resolution varies between less than 10% to up to 30%. Regarding X_{max} reconstruction resolution, the variation is between less than 15g/cm^2 to 100g/cm^2 . In general, the events with a zenith angle below 40° result in a more accurate and precise reconstruction.

The data points of all the events are combined in a cumulative histogram and a Laplacian function is fit to the distribution using the iminuit library. This is shown in Figure 10.11, where $\sqrt{2}b$ is equivalent to the standard deviation σ.

The overall reconstruction resolution for the energy with the χ^2_{max} method is equal to about 13%, which is close to the 10% obtained by the other experiments, whereas for X_{max} the overall resolution is in the order of 58 g/cm^2 , this is worst than for the other experiments [149]–[151], which, however, use more than three antennas for X_{max} reconstruction and strict quality cuts. No obvious bias is seen, some are above zero and others below in the event base plot. In the cumulative histogram, the bias of the energy is small at 0.8% and the bias of X_{max} is negligible at 1g/cm^2 . The histograms for all the different methods are in Appendix B.

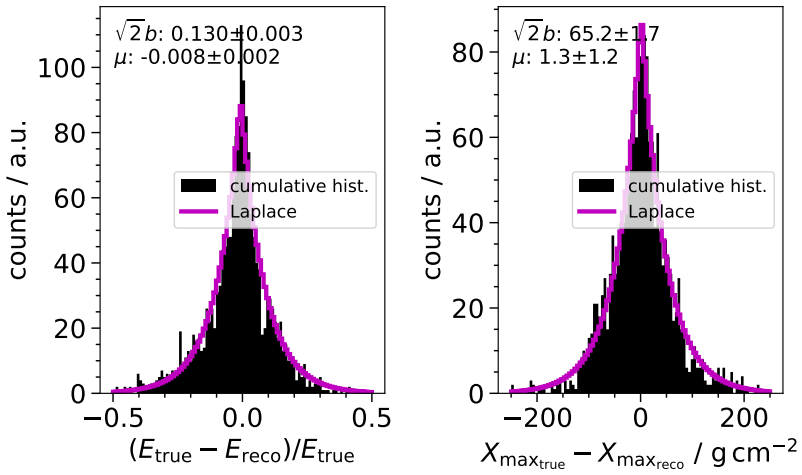


Figure 10.11: Distribution of all the events using the χ^2_{max}. A Laplace function is fit to the distribution where μ and $\sqrt{2}\,b$ represent the equivalent of the mean and standard deviation in a Gaussian distribution respectively. Simulated noise injected.

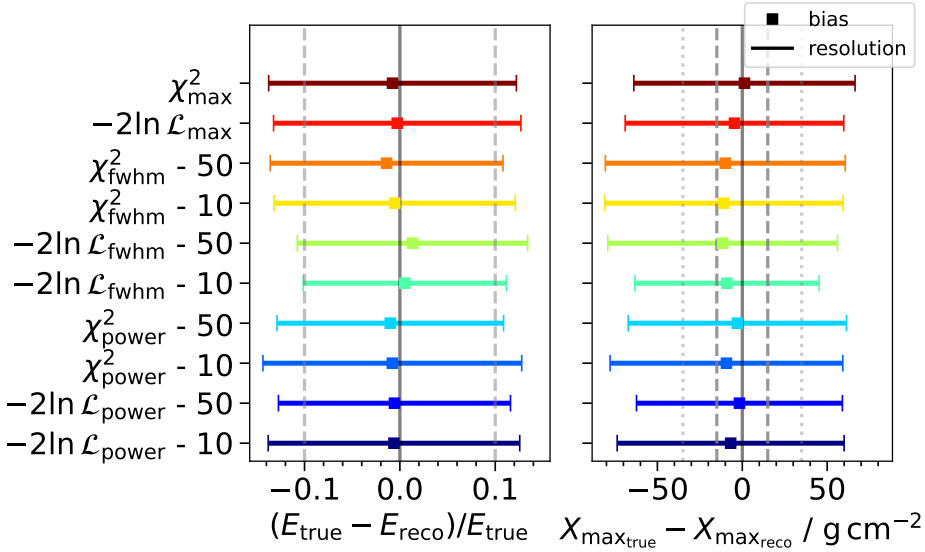


Figure 10.12: Bias and precision for energy and X_{max} of all the methods using all the events. The simulations have modeled noise injected for the creation of the mock data. The error bars represent the mean and standard deviation of the Laplace fit.

10.6.3 Comparing methods

The previous steps are applied to all the different methods, with a 10 ns, and 50 ns window, with modeled noise injected. The standard deviation and the mean of the cumulative histograms (e.g. Figure 10.11) are summarized in Figure 10.12.

The first observation from this figure is that all the reconstruction methods are more or less equivalent in X_{max} as well as in energy reconstruction. All the methods achieve roughly 13% resolution and a negligible bias in energy. For X_{max} a small bias is present for all methods, pulling the reconstructed value to be overestimated. The larger window of 50 ns gives, in general, a better resolution of the method when compared to their 10 ns counterpart. The resolutions are bounded

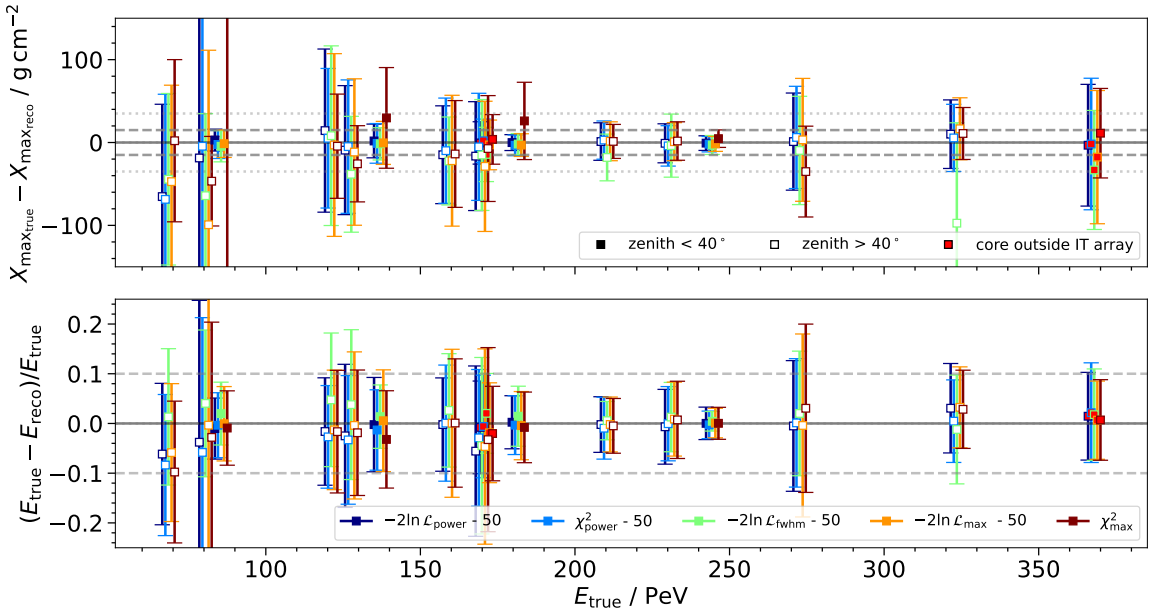


Figure 10.13: Bias and precision for the energy and X_{max} reconstruction with various methods on an event-per-event basis with median and 68% containment interval shown as error bars. Plotted against the true energy of the simulations. Simulated noise injected.

between 54.3 (± 2.0) g/cm^2 and 70.6 (± 2.0) g/cm^2 for the FWHM using the LLH minimization with a 10 ns window and the same variable using a χ^2 for a 50 ns window respectively.

A closer look on an event-per-event basis, with a plot similar to Figure 10.10, picking a few different methods, can inform on the consistency between the methods as well as on possible quality cuts that could be applied to the data.

Figure 10.13 illustrates the median and the 68% containment, using the MC-to-MC comparison, for all the events in the shower sample, with different methods. The showers with a zenith angle smaller than 40º are indicated by a full square and the showers whose core position is located outside the IceTop footprint with a bright red square. One can observe that the events are consistently reconstructed similarly by the different methods and that the low-energy events have larger uncertainties.

It is impossible to include all the shower variables on one plot, therefore Figure 10.14 shows the same information but as a function of the distance to the average X_{max} instead of the energy. The X_{max} is averaged because each of the simulations from the set has its specific X_{max} . The mean X_{max} of the 100 simulations made for each event, however, still provides a measure for the distance between the array and the shower maximum. The shower with the smallest distances to X_{max} seems to have smaller uncertainties. The distance to the X_{max} is, however, correlated with the zenith angle. Whether the air shower core is located outside of the array does not seem to influence the reconstruction significantly. A complementary plot can be found in Appendix B, where the same events are plotted in a shower-per-shower fashion.

10.6.4 All the methods with measured noise injected

Up to now, the results are with the modeled noise of the system, the theoretical minimum noise that can be achieved when taking into account the galactic noise and the thermal noise from the antenna and the LNA. In this section, all the previous steps for the MC-to-MC comparison are reapplied to all the methods but, at this point, with measured noise from the prototype station injected instead of the modeled noise.

Figure 10.15 is similar to Figure 10.12, which shows the efficacy of all the methods with all the

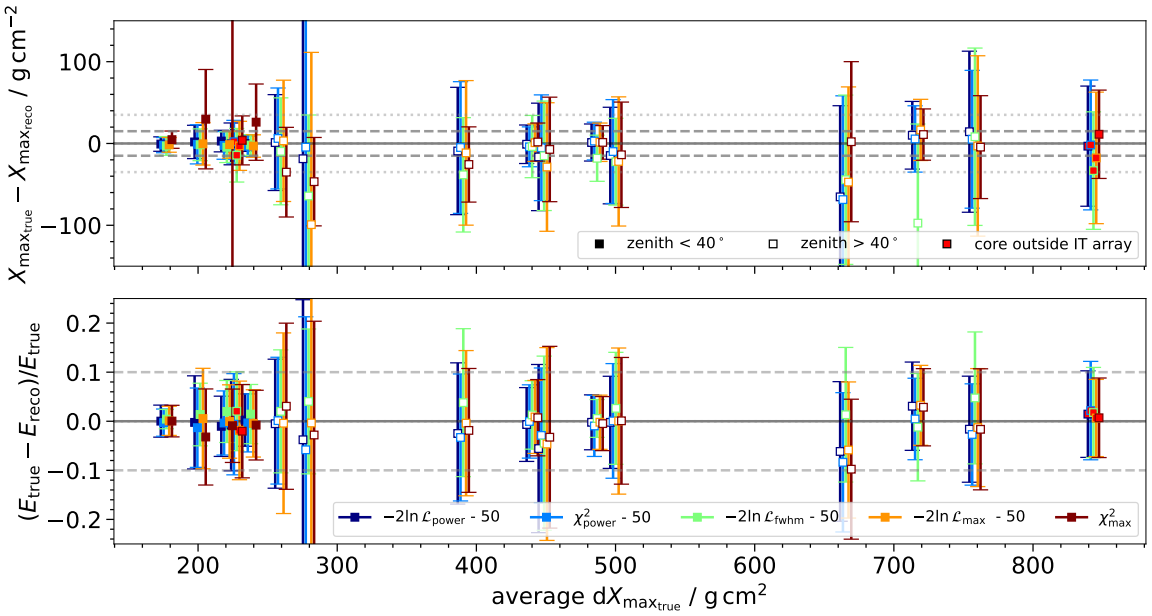


Figure 10.14: Bias and precision for the energy and X_{max} reconstruction with various methods on an event-per-event basis with median and 68% containment interval shown as error bars. Plotted against the distance to the maximum of the shower. Simulated noise injected.

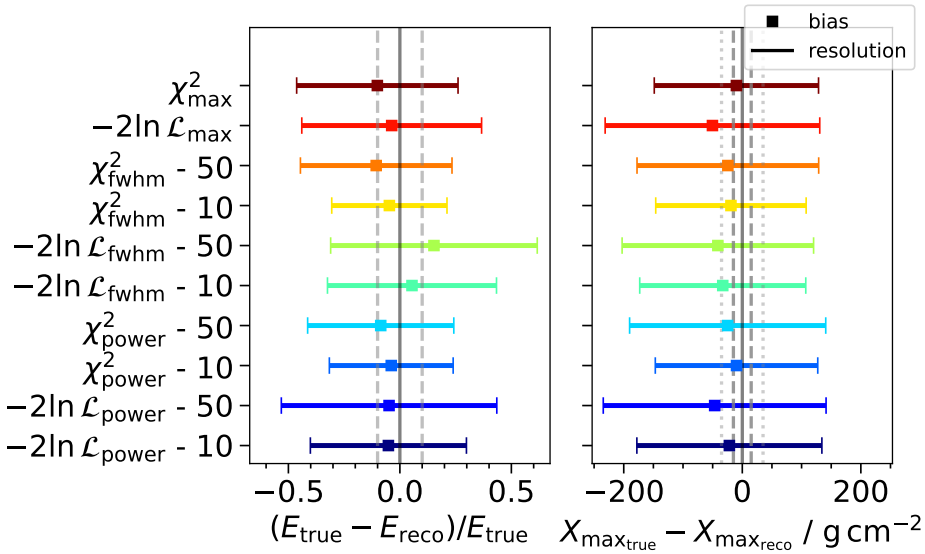


Figure 10.15: Bias and precision for energy and X_{max} reconstructed by all the methods using all the events. The simulations have measured noise injected for the creation of the mock data. The error bars represent the mean and standard deviation of the Laplace fit. (Note the different x-axis scales compared to Figure 10.12.)

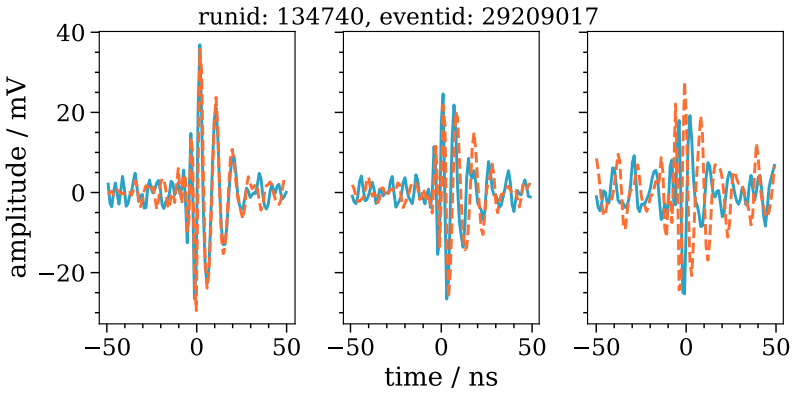


Figure 10.16: Waveform shifted to zero according to the time from the directional reconstruction of IceTop. Waveforms from a measured event

events combined, but with measured noise. The first observation is that the energy as well as the X_{max} resolutions worsen with the measured noise. Regarding the energy reconstruction, a bias in the order of 10% is present in most of the methods, fluctuating around zero for the different methods. Depending on the method, the resolution lies between 26.0 (± 0.7)% and 51 (± 2)% for the maximum of the Hilbert envelope using χ^2 and the power using a LLH and a 50 ns window respectively. As for the X_{max} reconstruction, similarly, as with modeled noise, a general bias, overestimating the X_{max} reconstruction, is present in all the reconstruction methods. Unlike for the modeled noise, for energy as well as for X_{max}, the methods using a 10 ns window perform better than their equivalent with a 50 ns window. The X_{max} resolutions are bounded by the FWHM in χ^2 with a 10 ns window for a value of 128 (± 3) g/cm^2 for the lower bound and by the power in LLH with a 50 ns window with a value of 196 (± 7) g/cm^2 for the higher bound. The large uncertainties are not surprising, as previous experiments have shown that at least 5 antennas are needed for a decent X_{max} reconstruction.

10.7 Results on data

In this section, the mock data is substituted by the real data recorded by the prototype station. The only difference in the methodology for the real data is that beamforming of each antenna is applied to find the pulse location in the trace. Beamforming means that the traces are shifted each by a time corresponding to the arrival direction. Beamforming techniques can be applied in two ways: first, the traces can be slowly shifted relative to each other searching for a coherence maximum, which can then informs about the presence of a shower and its direction; second, the method can also be applied reversely by knowing the direction of arrival of the shower and shifting the traces accordingly [155], which is the case here. The reconstructed direction is taken from the Laputop reconstruction of IceTop, the traces are shifted, the Hilbert envelopes of all the six channels are summed up, and the position of the maximum in the trace is found. This value of the pulse time is used as input for the pulse finder described in the previous chapter (see. Section 9.2.5), on each channel individually.

The Figure 10.16 shows the measured waveforms from the antenna 1, 2, and 3 from left to right, with the two polarisation channels. The shape of the waveform between the polarisation channels is in good agreement with antenna 1 and 2. All the pulses for the air shower are correctly shifted to zero from the IceTop air shower, so the direction reconstruction would agree well. In addition, Figure 10.17 illustrates as an example the summed Hilbert envelopes of the 6 channels after beamforming on four events measured by the prototype station. In the figure, the resulting waveforms are normalized to the maximum amplitude and shifted by a factor of 1.5 in the y-axis direction for clarity. The resulting waveforms for all events in the sample can be found in Appendix B.

Once the pulses are extracted from the waveforms of each channel, the data can be compared

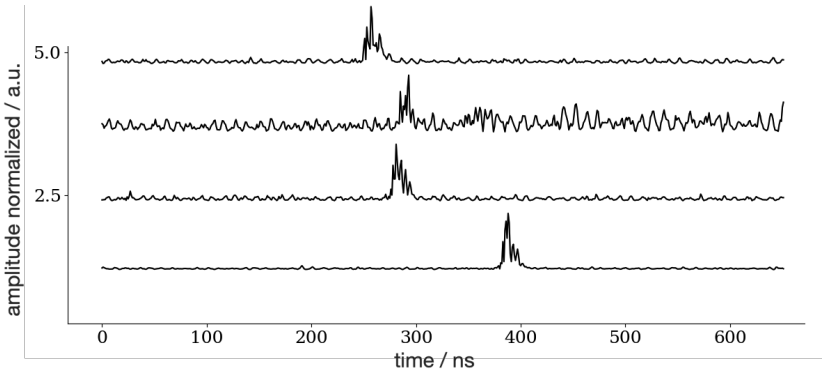


Figure 10.17: Example of the beamforming of the Hilbert envelopes for four measured events in the sample. The resulting waveforms are the sum of the six Hilbert envelopes time-shifted; they are normalized to the maximum height and shifted by a factor of 1.5 in the y-axis direction for clarity.

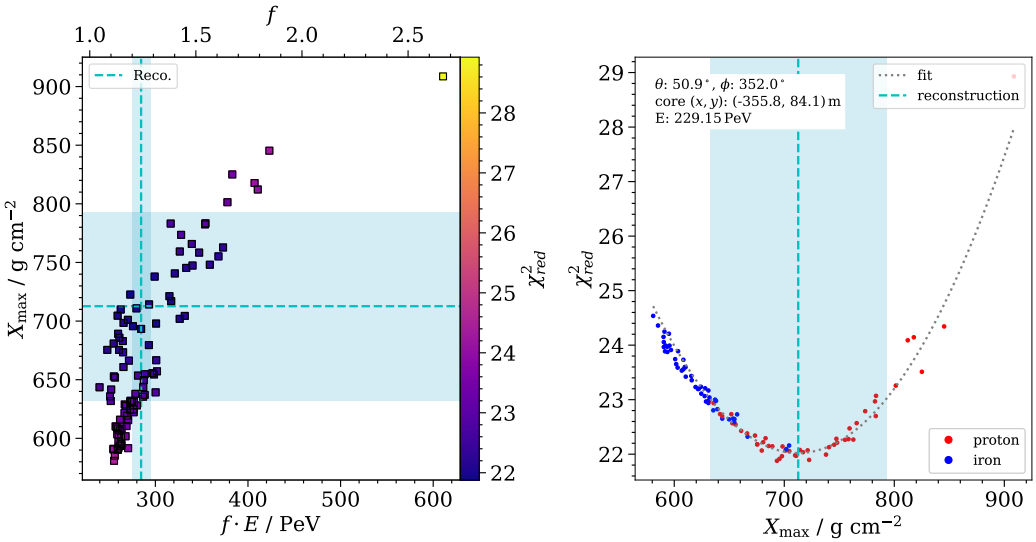


Figure 10.18: Reconstruction applied on measured data using the χ^2_{power} method (runid 134630, eventid 53455709),

to the noiseless MC simulations. The reconstruction method, applied to one event (runid 134630 – eventid 56738115, the same as the example shown with the mock data earlier), is shown in Figure 10.18. Although the uncertainty associated with the reconstruction is around 88 g/cm^2 a nice parabola in the right plot is obtained with the data. For this example, the χ^2_{power} method is chosen due to achieving qualitatively better results on data. For comparison, the same event reconstructed using the different methods can be found in Appendix B.

When applied to data, the χ^2_{fwhm} method gives unexpected results, splitting the proton and the iron in the reconstruction. The reconstruction often fails with all the methods when a 10 ns window is used, splitting the parabola into two parabolas. For the following methods: χ^2_{max}, $\mathcal{L}_{\text{max}}$, χ^2_{power}, and $\mathcal{L}_{\text{power}}$, the results are similar.

Analyzing further the χ^2_{power} method on data, about a third of the events show a nice parabola. The second third has only one branch of a parabola and the reconstruction falls just outside the simulated X_{max} range, two examples are shown in Figure 10.19. Finally, for the last third of events, the reconstruction, for different reasons, fails, mainly because the event is so far from the simulation that there is no parabola and only a linear function. Table 10.2, summarize the reconstruction for all the

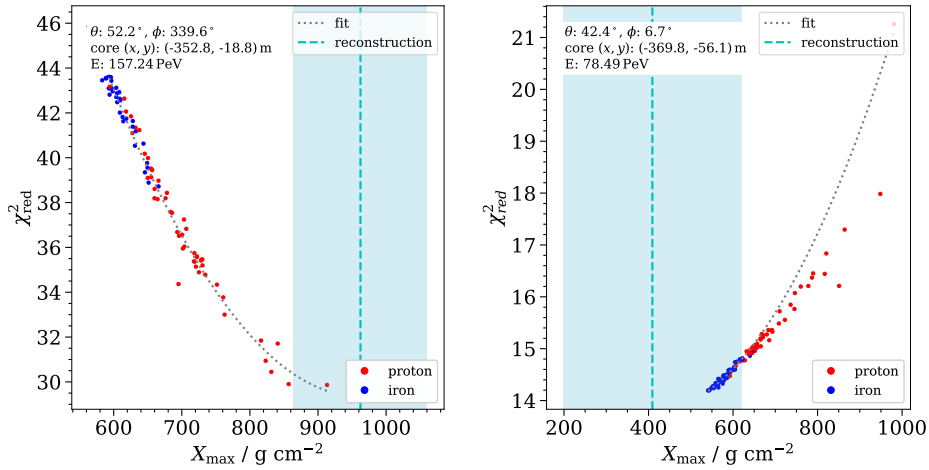


Figure 10.19: Example of reconstruction located outside the X_{max} space using χ^2_{power}.

Table 10.2: Reconstruction of the events in the sample using χ^2_{power}

runid	eventid	f	E^{reco} [PeV]	X^{reco}_{max} [g/cm^2]	Qualitative performance
134842	61959662	1.06 ± 0.002	256.53 ± 0.45	925 ± 5.5	outside ○○
134760	36546049	0.685 ± 0.008	123.04 ± 0.98	705 ± 22	good ○○○
134751	1987768	0.629 ± 0.018	52.55 ± 0.93	664 ± 67	good ○○○
134751	29771191	2.190 ± 0.028	260.91 ± 7.20	-663 ± nan	bad ○
134739	8585668	1.790 ± 0.046	224.92 ± 10.28	1029 ± 187	outside ○○
134777	12754797	0.876 ± 0.010	137.77 ± 1.38	962 ± 89	outside ○○
134740	29209017	1.079 ± 0.005	394.83 ± 2.10	1779 ± 102	bad ○
134762	40916495	1.161 ± 0.009	373.25 ± 3.25	785 ± 52	good ○○○
134746	5775675	0.849 ± 0.052	66.65 ± 3.46	409 ± 211	outside ○○
134630	56738115	1.242 ± 0.034	284.62 ± 9.79	713 ± 80	good ○○○
134826	53455709	0.871 ± 0.047	117.63 ± 5.54	740 ± 150	good ○○○
134625	31078935	1.932 ± 0.133	327.16 ± 43.35	871 ± 205	outside ○○
134860	71042091	0.570 ± 0.006	154.20 ± 0.98	911 ± nan	bad ○
134869	76177953	1.2 ± 0.004	242.06 ± 0.86	-167 ± 39	outside ○○
134871	24946279	1.560 ± 0.045	103.72 ± 4.69	830 ± 1032	bad ○
134930	4865364	1.248 ± 0.015	209.53 ± 3.05	1462 ± nan	bad ○

events in the sample with a qualitative performance associated with it, which is loosely defined as:

Good ○○○ The reconstructed value for X_{max} is inside the parabola.

Outside ○○ The reconstructed value lies outside the X_{max} phase-space, but the points aligns.

Bad ○ The reconstruction fails, no parabola can be seen.

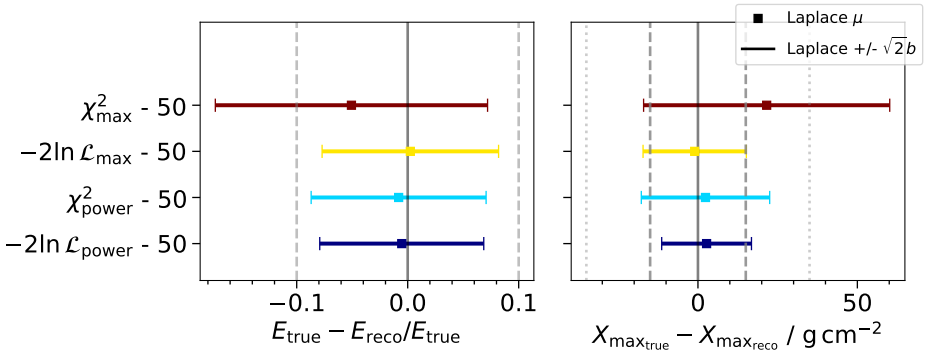


Figure 10.20: Bias and precision for energy and $X_{\max}$ reconstruction for the planned surface enhancement array comparing the χ^2 for the power and maximum methods versus the LLH, for the full array, with modeled noise.

10.8 Extrapolation to the complete array

Since the MC simulations are done with the star-shaped array and interpolated at the location of the prototype station, it is possible by using the same simulations to interpolate the signal at the location of all 96 antennas of the planned array. As the array is not yet built, the MC-to-MC comparison is used with the modeled noise injected. As the FWHM methods are computationally more expensive, and overall performed worse for the prototype station, the technique is put aside for this study. Only The maximum of the Hilbert envelope and the power are tested.

Figure 10.20 shows the results of the selected methods applied to the complete array. The χ^2 with the maximum and no noise correction here performs noticeably worse than the other techniques. For both measurable quantities, the usage of the LLH reduces the bias and the resolution compared to their corresponding implementation in χ^2 for both variables, energy and $X_{\max}$. The LLH power method produces a slightly better result than the LLH maximum with a resolution and precision in energy with values of 7.4 (± 0.2)% and -0.5 (± 0.1)% respectively, as well as in $X_{\max}$ with 14.1 (± 0.4) g/cm^2 and 2.7 (± 0.3) g/cm^2 . The method's equivalent in χ^2 results in a resolution of 7.9 (± 0.2) % in energy and 20.1 (± 0.5) g/cm^2 in $X_{\max}$, with an bias of -0.8 (± 0.1)% in energy and 2.4 (± 0.4) g/cm^2 in $X_{\max}$. This indicates that the target resolution of better than 10% in energy and 20 g/cm^2 in $X_{\max}$ will be achievable when more antennas will have been deployed.

10.9 Discussion

10.9.1 Initial reconstruction

Since the reconstruction method includes a scaling parameter f which is proportional to the energy and fitted to the data, it was initially assumed that the initial energy estimate can be coarse. In general, experiments using this technique require the scaling factor f to be between 0.5 and 2 [85], which directly translates to an initial seed in energy between 0.5 to 2 times the true value, and is in the order of what is achieved with the extrapolation. Nevertheless, the initial reconstruction seems to have an impact on the reconstruction which shows up by the fact that the vertical showers (<40°) have a better resolution in the MC-to-MC comparison, which is where the Laputop reconstruction from IceTop is optimized.

The initial reconstruction in energy is far from ideal, but moreover, as shown in Figure 10.21, the reconstructed energy begins diverging from the true energy at around a zenith angle of 40°, just after the parameterized values from IceTop. In the range from 40° to 65° of zenith, the difference in energy between the extrapolation used and other extrapolations is significantly smaller than what is observed, especially compared to the only linear extrapolation, with a maximal of ∼0.8 at 65° whereas

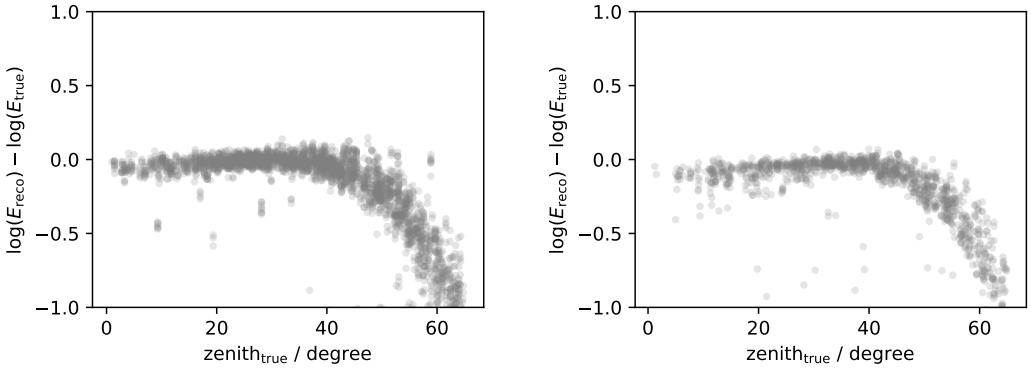


Figure 10.21: Initial energy reconstruction with the extrapolation compared to the true value of the simulation. On the right for energies of ∼50 PeV and, on the right for energies of ∼500 PeV Obtained from IceTop simulations using SIBYLL 2.1 [156].

in the comparison with the true value in reaches up to 10-fold. This is a strong indicator that a naive extrapolation cannot be done. The parametrization of the energy as a function of the S_{125} coefficient must, thus, be done for every zenith bin. In addition to the higher zenith angles, the snow attenuation should be included in the simulation. In this way, uncertainties of the energy reconstruction from the high zenith angles and from the snow accumulation will be significantly reduced. The simulations taking in account more recent snow accumulation are currently in production.

This reconstruction could be further improved by combining a first-order radio reconstruction using a simple LDF, and/or by using the energy reconstruction from the scintillation panels. Naturally, as the array expands, the reconstruction from those detectors will improve.

10.9.2 Quality cuts

Deriving quality cuts from 16 events is not an easy task, and the following recommendations should be reviewed in the future with more events. In radio analyses of air showers, the number of events is always limited due to the relatively high energy threshold. With a prototype station composed of three antennas and having a footprint of less than 100 m^2, the number of events is even more limited. For this reason, and due to the novelty of the station, these first events have no systematic quality cuts applied to them.

Having said that, from the results of this chapter, one can see that the quality of the reconstruction degrades for higher zenith angles. Inclined showers, although having fainter ground radiation due to the geometric attenuation (compared to a more vertical event of the same energy), are of interest because they feature a larger footprint, have their radiation energy completely released, and show a Čerenkov-like ring. Therefore, further research on the handling of inclined showers should be performed. As stated, this degradation of accuracy for inclined showers may be coming from the initial reconstruction being too far from reality, but it could also be related to other issues related to features that show up more dominantly with inclined showers, for example the displacement of the radio shower core relative to the particle core, due to refraction caused by the refractive index of the atmosphere [75], [76]. Although this effect should not affect the MC-to-MC comparison, due to its implicit consideration of the refractive index in the simulations, this would affect the reconstruction of measured events. In the short term, due to the limited number of events and antennas, a solution could be to focus only on air showers with a zenith smaller than 40°.

Another interesting observation is that the best-reconstructed showers are the ones with their X_{max} closer to the ground. This is a surprising result because the shower is not yet completely developed, thus its radiation energy is not totally liberated, and the Čerenkov-like ring is most likely not present. It is, however, not possible for the moment to disentangle the contribution from the proximity of

the shower maximum development from the effect of smaller zenith angles. Furthermore, the IceTop reconstruction is, as of yet, not calibrated for showers with a core outside of the IceTop array. This does not seem to significantly affect the MC-to-MC reconstruction. Nonetheless, the core position is bound to more uncertainties when located outside the array, and the position of the core directly influences the reconstruction of real data. Therefore, the events with the core located outside of IceTop should be removed from a high-quality sample of showers.

10.9.3 Simulation sets

The MC simulation sets are created with the same amount of protons and iron nuclei. In order to save computational time, the number of iron nuclei as primaries for the simulations could be reduced, due to the fact that the variation of $X_{\max}$ for an iron primary is smaller than for a proton primary, as shown in Chapter 3. Regarding the MC-to-MC comparison, cleaning out the limitrophe simulations from one set, i.e. not using as mock data the events whose $X_{\max}$ is located at the extremities of the $X_{\max}$ space, could improve the resolution estimation. The idea behind this is that these events cannot be reconstructed properly because they are not located within the $X_{\max}$ space and this results in situations similar to the examples shown in Figure 10.19. This was shortly examined here but did not lead to improvements in the 68% containment region for each event. A way to prevent that from happening in the reconstruction of real measurement is to create fast simulations using CONEX, picking the seed from specific targeted $X_{\max}$ values in the range, and then using these seeds in a complete CORSIKA simulation [157].

10.9.4 Effects of the noise

Looking at the reconstruction efficacy of the different methods in different circumstances gives interesting insights. The assumption at that, with only three antennas, all events belong to the large pulse regime seems to be true, all the methods behave very similarly when applied on three antennas with the modeled noise injected. This also confirms that the LLH implementations are sounds. Interestingly, when the methods are applied to the planned complete array, the parametrization of the noise influences the reconstruction positively. This can be attributed to a better handling of the noise especially for antennas in the small pulse or the transition regimes. Those regimes are accounted for by moving from the χ^2 technique to a more general LLH, where the behavior of μ and σ is parameterized. This is, however, only scratching the surface towards implementing a completely descriptive LLH. As example, the $P(x;\mu)$ probability density function used here could be transformed from a Gaussian to a Rician distribution as discussed in the previous chapter, and other suggestions are given in the next paragraphs.

Everything works well with modeled noise, however, when real prototype station noise is injected, things get more complex. Not surprisingly, the resolution noticeably worsens because the noise level is higher. Furthermore, the prototype station noise has a bi-modal distribution, which can be seen in Chapter 8 and 10, which is, at the moment, characterized only by one monomodal function. Another challenge concerning the noise is that it depends on the time of day. Overall, the frequency spectrum is relatively stable over the year for the prototype station, but some factors like the operation of the IceAct telescope, or simply human activity during the summer season can affect the noise level. Also, a difference between the noise levels of different antennas, either arising from their hardware or from their location, is seen in the data. All these variations probably explain why, in general, the LLHs performs worse than their χ^2 equivalent with the prototype station, where the χ^2 methods samples directly the noise from the trace of the specific recording. These issues could probably be mitigated by doing an antenna-wise parametrization and by deriving the parametrization from the soft-triggered data of the day where the event is recorded and using a complete two-dimensional parametrization, to account for the bi-modal distribution.

Having said that, the noise level in the real data could be accounted for by including a noise factor directly in the parameterizations, which might be an easier approach than the previous suggestions.

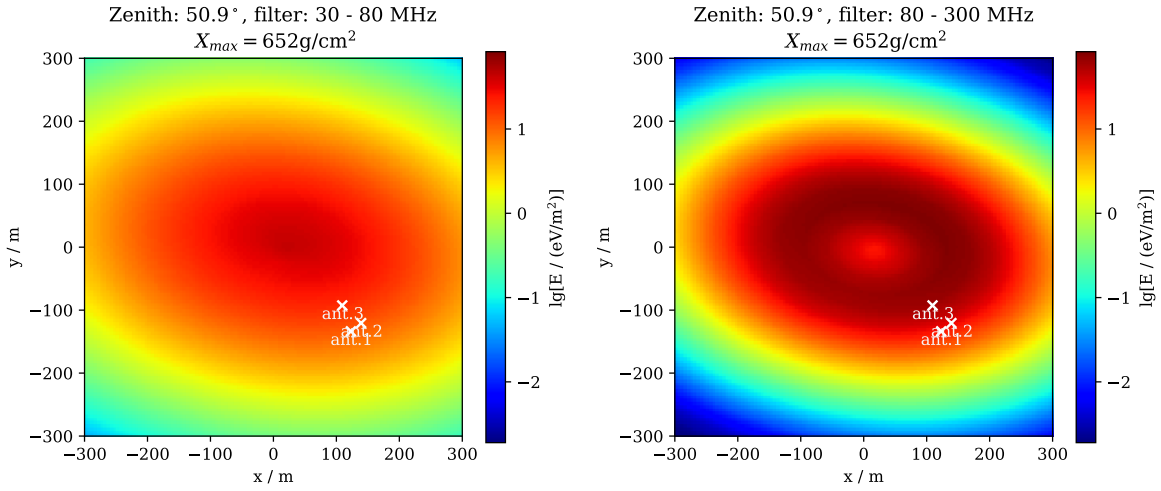


Figure 10.22: Radio emission received at the ground for a simulation based on the example event (runid 134630, eventid 53455709). On the left, the electric field is filtered between 30-80 MHz, and on the right, between 80-300 MHz.

This noise factor can be measured for each trace using the noise estimation of the previous chapter, σ_{trace}, then instead of $Q'(f, Q_{\text{MC}})$, the term would become $Q'(f, Q_{\text{MC}}, \sigma_{\text{trace}})$, similarly $\sigma'_Q(f, Q_{\text{MC}})$, would become $\sigma'_Q(f, Q_{\text{MC}}, \sigma_{\text{trace}})$. Where Q represents any measurable observed.

Nevertheless, effort should be put into reducing the noise level, because the reconstruction quality will generally depend on the noise level, no matter the method utilized. That can be seen in the noticeable difference in resolutions between the modeled noise and the measured noise. Once more, since this is a prototype station, active improvements are ongoing. Concerning the noise level, work on baseline adjustments that seem very promising in reducing the noise is on the way. Also, a more intricate filter, that takes the frequency spectra of the background and of an average air shower into account, is under development. Finally, all the characterizations of the noise have to be repeated with the new TAXI v3.2, as the behavior is different from the previous TAXI v3.0 system on which this work is based.

10.9.5 Advantage of the frequency band

As discussed earlier, the prototype station uses a higher frequency band compared to the usual 30-80 MHz bandwidth used in the other experiments. To estimate the effects of the higher band, one can look at the radio footprint of the example shower.

Figure 10.22 illustrates the radio radiation energy on the ground for the 30-80 MHz frequency band on the left, and for the 80-300 MHz band on the right. The footprint is obtained from one MC simulation, using the star-pattern interpolation to acquire the values over the whole area. These measurables are calculated from the electric field directly and no noise or electronic response is added. These pictures already illustrate a qualitative difference in the footprint. In the lower frequency band, the emission is smooth, with the highest energy located in the center, at the shower axis. In the higher frequency band, the emission is sharper, thus the gradient of the emission is higher. The presence of the Čerenkov-like ring can be seen close to the positions of the antennas, and the strongest emission is at this position. From the presence of the Čerenkov-like ring and from the stronger gradient, intuitively, one would expect that the 80-300 MHz footprint is more sensitive to the $X_{\max}$ position.

If the $\chi^2_{\max}$ reconstruction method is applied to that example event with a 30 to 80 MHz filter (Figure 10.23 and compared to the previous Figure 10.9 where the filter is set from 80 to 300 MHz, the initial observation is that the uncertainty band increase from 88 g/cm^2 to about 150 g/cm^2 . Nonetheless, the characteristic parabolic shape is still present.

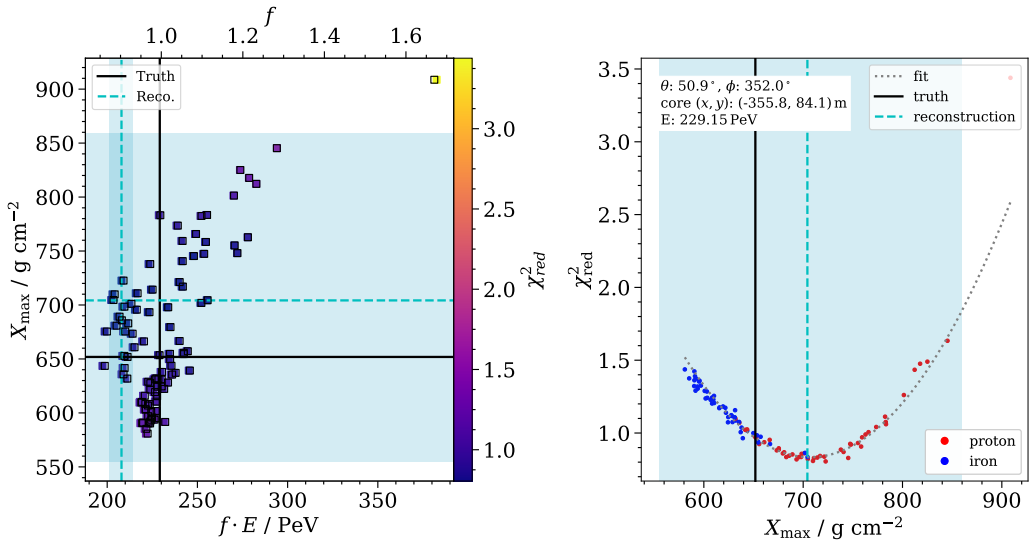


Figure 10.23: Reconstruction using the χ^2_{max} method. The mock data is with a proton as primary and modeled noise injected. (runid 134630, eventid 53455709)

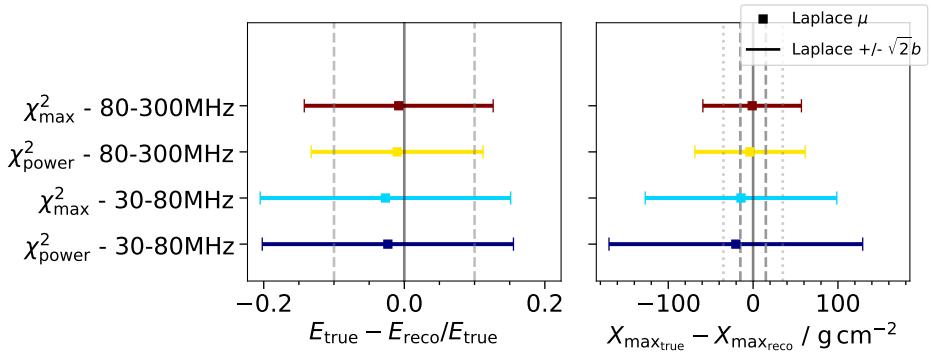


Figure 10.24: Bias and precision for energy and X_{max} for the 30-80 MHz versus the 80-300 MHz frequency bands with modelled noise injected. No parametrisation of the noise is used here, because they have not derived been for the lower bandwidth.

It could be that the uncertainties are over-estimated. Therefore, the reconstruction is run using a MC-to-MC comparison for all events. The results only account for the reconstructed value versus the true value. The resulting accuracy of the χ^2 power method and the χ^2 maximum of the Hilbert envelope method with the two different bands is shown in Figure 10.24. No LLH are tested here because it would involve a complete parametrization of the noise for the lower band. A net decrease of the resolution in energy as well as in X_{max} is noted, for both methods, in the 30-80 MHz bandwidth. The higher bandwidth of the SAE does appear to improve notably the resolution on X_{max} , but also on energy. Lower frequencies (around 100 MHz) are, however, still important, because they carry more energy from the air shower, due to the coherence phenomenon discussed in Chapter 3. Furthermore, the larger bandwidth of the prototype station also helps to capture more energy from the radio pulse in the trace.

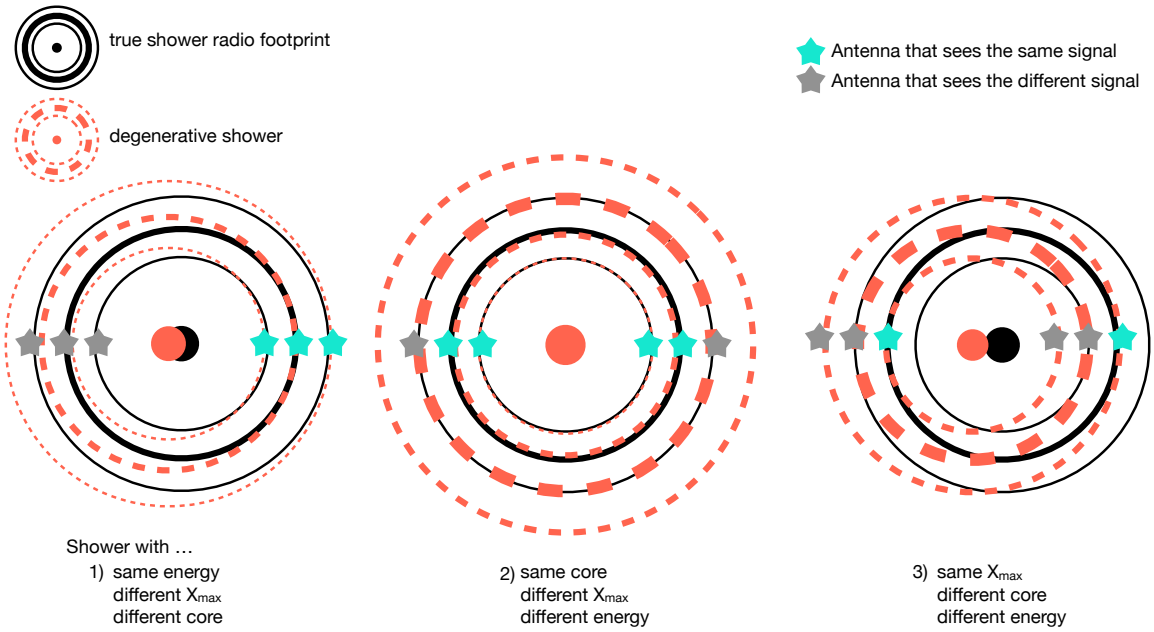


Figure 10.25: Schematics of the shower degeneracies.

10.9.6 Degeneracy of the reconstruction – Fixing the core

Mathematically, it is only possible to resolve an equation with the number of parameters not exceeding the number of degree-of-freedom (DOF)−1. For a fitting procedure, the more points, the better. With the prototype station, there are 6 DOF, assuming that there is a signal in each polarisation of each antenna. Although, from these six DOF, only three are at a different location. Theoretically, it is possible to fit in addition to the energy (via f), the core position of the air shower (x_c, y_c).

Besides, there are other constraints rising from the geometrical layout of the array. Figure 10.25 schematically represents the possible degeneracies of the reconstruction. The thickness of the lines represents the strength of the radio emission, the ring is the characteristic Čerenkov-like ring of the radio footprint, and the stars are the antennas. If the star is colored light blue, the antenna would see the same signal for the "true shower" and the "degenerative shower". The three antennas of the prototype station are clustered together, which can be approximately represented by the three antennas on the right side of each ring. In situation 1, the prototype station antennas see the same signal for both the true shower and a shower with a specific different core, hence making it impossible to distinguish between both situations, i.e. where the core location is and what the X_{max} value is. It would be possible to distinguish between the two showers if at least one antenna would be located on the other side of the core. In situation 2, applied in this chapter, the core is fixed while the energy and the X_{max} are varied. It is then possible for the three antennas on the right to distinguish between the true shower and the degenerative shower because one of the antennas records a different signal. In situation 3, assuming that the X_{max} is known, it is possible with the three antennas of the prototype station to reconstruct the core position and the energy of the shower.

Extending this line of thought, it would be possible to fit the core and the X_{max} with only three antennas in the case where the core would be located between the antennas. In the same vein, it would still be impossible to separate the core from the X_{max} effects on the footprint with hundreds of antennas located in a dense cluster on one side of the shower footprint.

10.9.7 Application on measurements

From the results of the χ^2 method, it can be seen that the technique works for the real measurements since several showers show the characteristic parabola and a few others are close to it.

The reason why the χ^2_{fwhm} method does not work well on measurements is probably the same as why the 10 ns window does not work either; the phase calibration of the TAXI v3.0 and of the radioTad v1 is not accurate enough to reconstruct the pulse shape properly. This aspect is already known from the calibration presented in Chapter 7, and is improved with the new DAQ system. The χ^2_{fwhm} method is more sensitive to the phase as it fits the envelope of the pulse to the simulations. A difference in simulation and measurement will thus be more noticeable in this method than for an integrative method or the maximum of the Hilbert envelope, both more robust against mis-calibration of the phase, and group delay. Similarly, for the 10 ns window, since only a small part of the pulse is taken, if a shift of 1 or 2 ns in the pulse window occurs, a different section of the pulse is considered in the measurements versus in the comparison of the simulation. Whereas with a 50 ns the pulses are (for most of the pulses) contained within the pulse window, thus, if a shift in the window happens, the extraction of the pulse quantity is not affected.

It is interesting that all the working methods result all either in a "good reconstruction" or all in a "bad reconstruction" for a given event. For the "outside reconstructions", the ones located just outside the parabola, the two simplest reasons are: first, the initial energy values are too far from reality, and then the wrong X_{max} phase-space is created, or, second, the core simulated is too far away from the real core, due to an erroneous core reconstruction.

10.9.8 Which energy to reconstruct

The radiation energy from proton and iron-initiated showers differs, with around 10% less EM energy in iron-induced air showers than for proton-induced air showers, which can explain the two populations visible in the left figure of Figure 10.9. The uncertainty of the energy reconstruction could be reduced by using the radiated energy instead of the primary energy. Furthermore, the radiation energy is an important variable as it relates directly to the EM energy carried by the air shower, and combined with other observables, e.g., the number of muons, gives separation power for mass composition.

However, since the array is located at a high altitude of around 3 km, which is equivalent to an atmospheric depth of around 700 g/cm^2, it is possible and even frequent that the shower development continues inside the ice shelf. Since, as discussed in Chapter 3, the majority of the radiation is emitted close to the shower maximum, it is unclear at the moment how the radiation energy can be calculated accurately without introducing a bias. This is why here the primary energy is used. A deeper study should be done in this area, but the implementation will be straightforward. The scaling parameter f would be simply related to the radiation energy instead of the primary via

$$E^{\text{reco}}_{\text{rad}} = f\, E^{\text{sim}}_{\text{rad}}. \tag{10.16}$$

Finally, one might wonder why this study was not performed with a large MC simulation library like they are used for most of the reconstruction methods under development. The reason is, that this method requires a lot (~ 100) MC simulations for every shower geometry, making it computationally impossible to create a sufficiently large library of events. A very interesting approach around that could be a machine learning algorithm that could re-create the radio footprint, with the air shower geometry as input. Machine learning approaches are generally performing well for average quantities, but struggle with fluctuations, which are crucial in air shower simulations. A possible way to circumvent that could be the introduction of the X_{max} of the desired footprint as input.

In conclusion, all discussed methods deliver a comparable reconstruction accuracy for the three antennas and work with the novel higher bandwidth that will be used for the SAE. The higher band leads to a better resolution for energy as well as for X_{max}, than the widely used 30-80 MHz bandwidth. Furthermore, once the array is completely deployed and consists of several antennas, the LLH method is promising for handling the small and the medium size pulses. With the noise in the waveforms

reduced to only Galactic and thermal noise, which is likely doable at the South Pole, the future SAE will have a great handle on X_{max} from its frequency band, showing an uncertainty due to the method as low as 14g/cm^2 without any quality cuts on the events reconstructed.

The method, however, does not account, as of yet, for a core displacement, which can be implemented in the method by fitting the core position once more antennas are deployed in the field. The main drawback of the method at the moment arises from the clustered and limited number of antennas on the field, which forces the core to be fixed. Another important challenge is related to the noise of the prototype station, but as with any new detectors, active improvements, and understanding are ongoing.

Finally, multiple different methods were implemented, and tested with mock data using modeled noise and measured noise for 16 events of recorded by the prototype station. The workings of those methods were confirmed with their application on measured data. More importantly, all the tools necessary for this analysis are implemented in the IceCube software and use the radcube module, making it easy to be extended and/or applied to new data or more antennas.

CHAPTER 11

Conclusion

In conclusion, multiple different subjects related to the radio antennas of the IceCube's surface array enhancement have been exposed in this work, from hardware to analysis, encompassing commissioning and production. Overall, the prototype station works, apart from a few shortcomings that were discussed. The detection of cosmic rays was proven possible with the prototype station, and the first energy and X_{max} analysis was implemented and tested on measured data.

To achieve this, a structure to keep the antennas above the snow, and to facilitate the deployment and maintenance was designed and built. The first version of this structure has been exposed to the harsh weather conditions of the South Pole and successfully resisted them for three consecutive years. The second version, an improved design over the first version, was produced for more than seven stations and is ready for the coming deployment. Both versions of the mounting structures, primarily composed of wooden dowels, reduced the cost of the structure by one order of magnitude compared to the previous version in [102]. In addition, the mount v2 is optimized to be simple to assemble at the site. In fact, the assembly of the antenna and the mount takes two people less than 30 minutes and requires only 8 bolts to be tightened for a full assembly in the field. Furthermore, the raising of the structure is only limited by the cable slack of the antennas, and its height of 1,20 m above the snow warrants necessary maintenance only every 5 to 6 years. A system of shims enables the leveling of the antennas to 0.5°, which translates to a maximal difference in gain of 0.21 dB and an average uncertainty of 0.031 dB for air showers with a zenith of less than 60°. Concerning the azimuthal positioning of the antenna, a differential GPS survey shows the effect to be negligible on the measurement uncertainty of the station.

Regarding the electronics of the radio signal chain, their functioning, design, and calibration were detailed. The radio electronics front-board, for pre-processing the radio signal by filtering, amplifying, and splitting the signal in "almost"-identical copies was redesigned from its predecessor. Multiple tests were done to improve the response of the board. In this work, the version of the radioTad v1 that was used in the prototype station, and the newer version that was produced more massively, radioTad v2 were explained. It was shown that the channel-to-channel variations, in gain as well as in phase, greatly improved from version 1 to version 2, reducing the amplitude uncertainty from 7% to 1.7%.

The calibration of the radio signal chain is of crucial importance for the radio array to be able to reconstruct the energy of an air shower within 10% uncertainty. Hence, at most a 10% amplitude uncertainty is required for the total radio signal chain. The temperature dependence of the radioTad v2 is mostly negligible, of the order of 0.6%, for the nominal frequency band at operating temperature. While the calibration of the LNA shows a small dependency on the temperature, inducing an uncertainty of about 2%, the variation between LNAs is of the same order with a standard deviation of 2.8%. The influence on the variation of gain in the cable is negligible as well, with a marginal 0.3%. The 60 produced radioTads v2 were tested for malfunctions, and are undergoing a complete calibration in conjunction with the TAXI v3.2 board.

All of these uncertainties combined together result in an estimated 3.9% partial uncertainty on the

amplitude of the radio signal. This calculation, however, does not include the uncertainties arising from the antennas. The radioTad–TAXI calibration is similarly not included. The method of calibrating the antennas is under investigation, whether through anechoic chamber measurements, the galactic background radio emission, a calibration with a pulser on-site, or a combination of them. In the situation where the variation between the different TAXIs or antennas would be larger than 10%, a station-wise electronic response can be implemented in the analysis. An open question, however, is the absolute calibration of the LNA. As was shown, the calibration deviates largely between different measurements done in different contexts up to a few dB.

After the first iteration in design, a prototype station was deployed at the South Pole in January 2020. The commissioning of the data from the prototype station has been successful and 121 air showers have been identified during the two and a half years of operation, where, the system was operational far less than 100% of the time due to active development of the firmware onboard the TAXI and various tests. Of these 121 air showers, 5 have an in-ice counterpart. Albeit some features in the radio data, e.g., bin spikes, possible ROI corruption, a not well-implemented capacitor baseline correction, and a higher noise level than expected, the radio data is relatively well-behaved.

In addition to the hardware and commissioning of the radio array within the hybrid prototype station, a reconstruction method for the radio data was developed, based on the state-of-the-art reconstruction method in the field of radio arrays, and applied to only three antennas. So far all experiments required at least 5 antennas for X_{max} reconstruction. Moreover, the used frequency band of 70-350 MHz is novel. The technique consists in comparing measured data to a set of Monte Carlo simulations created based on an initial reconstruction of the event. A χ^2 method is used where a scaling parameter related to the energy is minimized. Lastly, the χ^2 values are plotted against the X_{max} of the simulation set, and a parabola is obtained. The minimum of this parabola represents the reconstructed X_{max} of the measured data. In order to test the technique and estimate the reconstruction accuracy, mock data is used. The technique was developed with 16 air showers with energies between 50 and 400 PeV for different measurable quantities from a radio waveform.The results for the three antennas using mock data with modeled noise injected are mostly equivalent for all quantities with a resolution in the energy of around 13% and about 60 g/cm^2 in X_{max}. This was compared with the widely used lower frequency band of 30 to 80 MHz, with otherwise the same parameters. This resulted in a significantly worse reconstruction for the lower frequency band of about uncertainties 20% in energy and about 140 g/cm^2 in X_{max}. This confirms the potential of the radio array within the surface array enhancement (SAE) to deliver an accurate X_{max} reconstruction once more antennas are deployed.

Furthermore, the potential uncertainties on energy and X_{max} reconstruction, with a log-likelihood (LLH) method reach less than 10% and around 14g/cm^2 respectively, for the planned complete array. In both cases with negligible bias. This corresponds to 5g/cm^2 less than its equivalent using a χ^2 method, assuming that the noise is well-behaved and the only contribution is from the galactic background and the thermal noise of the LNA. Although the χ^2 method may be further improved regarding the correction of the measured radio pulses for noise. Nevertheless, these are encouraging results for the capability of the planned SAE to measure the mass composition of cosmic rays. Furthermore, these results do not yet take any quality cuts on the data into account.

When the reconstruction methods developed are applied using measured noise, the resolution and bias worsen significantly, reaching around 40% in energy and 150 g/cm^2 in X_{max}. These results are not surprising when observing the noise level in the waveforms, where the measured noise is about twice as high as the modeled noise and presents a bimodal distribution. It is a clear indication that a reduction of the noise level from the prototype station is needed. A capacitor-baseline correction has shown, so far, great potential in that regard. Finally, the discussed reconstruction technique was applied to the data. A third of the events to which it was applied showed the characteristic X_{max} reconstruction parabola. The other two-thirds show either just a branch of the parabola or simply a line. Both effects are most likely attributed to two factors: first, an initial reconstruction too far from reality, and thus the measured data falls outside the given X_{max} phase space. And second, the fixed shower core position in the minimization biases the X_{max} reconstruction. Thus, the assumption that the initial reconstruction could be coarse does not hold with only three antennas, because the core cannot be moved and the

number of measured signals is so small that the only way to have an estimate on $X_{\max}$ would be to have a precise initial reconstruction and strong quality cuts. With an accurate initial reconstruction and strong quality cuts, as was shown with the MC-to-MC comparison, with modeled noise, is there a chance for an $X_{\max}$ accuracy below 35 g/cm^2. In the examined 16 air shower sample, only 5 events in the MC-to-MC comparison achieve such a resolution. It was ambitious to implement this technique with the prototype station as it has only 3 antennas, but nonetheless, the tools are now developed, work within the IceCube software, and can be easily expanded.

The implementation of the LLH required a complete study of how the noise affects the radio signal. This was done for a quantity of the envelope of the pulse, the maximum of the Hilbert envelope, and for an integrative quantity, the power in the pulse. The results are given for the modeled noise as well as for the noise measured by the prototype station, and constitute the base of the parametrizations used in the LLH method.

The prototype station well exceeded its initial mandate and performed well on the field tests with 121 recorded air showers in approximately 2 and a half years of partial operation. The future entails a larger array with more antennas, whose electronics improved from the prototype station. Production of the hardware for the first seven stations of the final design is already concluded. In the shorter term, the operational time of the array will reach close to 100% with the improvements on the firmware, making the set of detected air showers with radio signal substantially larger, and through that allowing more quality cuts. For the energy and $X_{\max}$ reconstruction, the parametrization of the noise should be repeated with the new electronics. Eventually, when more antennas are on the field, the technique should allow the core to be fitted as a free parameter, which is possible with the techniques of interpolation. On a broader outlook, the goal of this hybrid array should be the combined reconstruction of all the detector types the IceCube Neutrino Observatory houses. The potential is great, with radio antennas for the electromagnetic part of the air showers and $X_{\max}$, the scintillation panels and the IceTop tanks for the electromagnetic particles and low energy muons at the surface, and the in-ice detector for the high energy muons.

List of Abbreviations

TDE	tidal disruption event
CMB	cosmic microwave background
HE	high energy
VHE	very high energy
UHE	ultra high energy
UHECR	ultra high energy cosmic rays
CR	cosmic ray
GZK	Greisen, Zatsepin and Kuzmin
EAS	extensive air shower
EM	electromagnetic
IT	IceTop
SAE	surface array enhancement
DOM	digital optical module
PMT	photomultiplier tube
GRB	gamma-ray burst
SN	supernovae
GW	gravitational wave
AGN	active galactic nuclei
LDF	lateral distribution function
NIAC	non-imaging Cherenkov telescope
FD	fluorescence detector
TA	Telescope Array
LPDA	log-periodic dipole array
SD	surface detector
PMNS	Pontecorvo-Maki-Nakagawa-Sakata
CC	charged current
NC	neutral current
LLH	log-likelihood
COG	center of gravity
ICL	IceCube lab
RFI	radio frequency interference
DAQ	data acquisition
WR	white rabbit

FPGA field programmable gate array
DRS domino ring sampler
ADC analog-to-digital converter
uDAQ scintillator uDAQ v4.1
radioTad radio front-end board
VNA vector network analyser
LNA low noise amplifier
SiPM silicon photomultiplier
SerDes Serializer Deserializer
KIT Karlsruhe Institute of Technology
ROI Region-of-Interest
TAXI Transportable Array for extremely large area Instrumentation studies
DHCP dynamic host configuration protocol
PCB printed circuit board
DUT device under test
SKA Square Kilometer Array
SNR signal-to-noise ratio
SA spectrum analyser
RMS root mean square
CAD computer-aided design
FEA finite element analysis
TA Telescope Array
FFT fast Fourier transform
NOAA National Oceanic and Atmospheric Administration
FWHM full width at half maximum
MC-to-MC Monte Carlo to Monte Carlo
GDAS global data assimilation system
MC Monte Carlo
DOF degree-of-freedom

Appendices

APPENDIX A

Programming information

A.1 Laputop reconstruction configuration

```
lambda = 2.4
tray.Add('I3VEMConverter',
         PEPulses='IceTopPulses',
         VEMPulses='ITVEMPulses')

tray.Add(IceTopWaveformSplitter,
         Input='ITVEMPulses',
         OutputHLC='ITHLCVEMPulses',
         OutputSLC='ITSLCVEMPulses')

tray.Add('I3HLCTankPulseMerger',
         InputVEMPulses='ITHLCVEMPulses',
         OutputTankPulses='ITHLCTankPulses',
         ExcludedTanks='MergerExcludedHLCTanks')

tray.Add('Delete',
         Keys=['ClusterCleaningExcludedTanks', 'CleanedHLCTankPulses']
         )

tray.Add('I3TopHLCClusterCleaning',
         BadTankList='TankPulseMergerExcludedTanks',
         ExcludedTanks='ClusterCleaningExcludedTanks',
         InputPulses='ITHLCTankPulses',
         InterStationTimeTolerance=200.0 * icetray.I3Units.ns,
         IntraStationTimeTolerance=200.0 * icetray.I3Units.ns,
         OutputPulses='CleanedHLCTankPulses',
         SubEventStreamName='IceTopSplit',
         If=lambda frame: ('ITHLCTankPulses' in frame)
         )

tray.Add('I3TopRecoCore',
         DataReadout='CleanedHLCTankPulses',
         NTanks=7,
         ShowerCore='ShowerCOG',
         Verbose=False,
         Weighting_Power=0.5,
         If=lambda frame: ('CleanedHLCTankPulses' in frame)
         )

tray.Add('I3TopRecoPlane',
         DataReadout='CleanedHLCTankPulses',
         ShowerPlane='ShowerPlane',
         Trigger=3,
         Verbose=False,
```

```
        If=lambda frame: ('CleanedHLCTankPulses' in frame)
        )

tray.AddSegment(toprec.LaputopStandard, 'Laputop',
                ShowerCOGSeed='ShowerCOG',
                ShowerPlaneSeed='ShowerPlane',
                excluded='ClusterCleaningExcludedTanks',
                pulses='CleanedHLCTankPulses',
                snowfactor=snowfactor,
                If=lambda frame: ('CleanedHLCTankPulses' in frame)
                )

tray.AddSegment(toprec.LaputopSmallShower, 'LaputopSmall',
                ShowerCOGSeed='ShowerCOG',
                ShowerPlaneSeed='ShowerPlane',
                excluded='ClusterCleaningExcludedTanks',
                pulses='CleanedHLCTankPulses',
                snowfactor=snowfactor,
                If=lambda frame: ('CleanedHLCTankPulses' in frame)
                )
```

A.2 Important bash commands for FPGA configuration of the radio data taking

```
## Cascading options:
# Cascading. 4-channels => 0x11 -> 1000 1000
smcrw -w10ac -v11
## Non-cascading -> 0xff -> 1111 1111
smcrw -w10ac -vff
## Semi-cascading -> 1010 1010
smcrw -w10ac -vaa

## SerDes delay
# Non-cascaded -> Set setSerdesDelay to a value of app. 60
# Semi-cascaded -> Set setSerdesDelay to a value of app. ????
# Cascaded --> Set setSerdesDelay to a value of app. 250
/home/root/aw/test/icescint_config_20200701 --setSerdesDelay 60

## Power
## Enabling the power supply to the radio boards -> v1 == Off, v0 == On
smcrw -w10f4 -v1

## Soft trigger
# Turn off FPGA trigger
smcrw -w11ec -vFF
# Turn on the soft trigger and set the generator period (in clock cycles)
/home/root/aw/test/icescint_config_20200701 --setSoftTriggerGeneratorEnable 0
/home/root/aw/test/icescint_config_20200701 --setSoftTriggerGeneratorPeriod 4294967295
```

A.3 Injection of noise in the radcube processing of Monte Carlo simulation

without noise (WN)	**with ideal noise**	**with prototype noise**
import the electronic response and the antenna response		
read CORSIKA simulation pad E-field with zeros convolve E-field with antenna response re-sample for 1 ns bins add electronic response digitize waveform	read "WN" i3 file *n* times grab re-sampled frame add ideal noise add electronic response digitize waveform	read "WN" i3 file *n* times grab re-sampled frame add electronic response digitize add prototype noise
remove pedestal and convert to voltage filter between 80 and 300 MHz write i3 file extract pulses from i3 file analyse		

A.4 Complete steering file example for reconstruction method

RUNNR	0
EVTNR	1
SEED	30168043 0 0
SEED	31168043 0 0
SEED	32168044 0 0
NSHOW	1
ERANGE	135.01404724762747e+6 135.01404724762747e+6
ESLOPE	-1.0
PRMPAR	14
THETAP	33.0596 33.0596
PHIP	105.5091 105.5091
THIN	1e-06 135.01404724762747 0.0
THINH	2.00E+02 10.000000
ECUTS	0.02 0.01 4.0E-04 4.00E-04
CASCADE	F F F
ELMFLG	T T
OBSLEV	284000.0
ECTMAP	1.e11
SIBYLL	T 0
SIBSIG	T
FIXHEI	0. 0
HADFLG	0 1 0 1 0 2
STEPFC	1.0
MUMULT	T
MUADDI	T
MAXPRT	1
MAGNET	16.75 -51.96
LONGI	T 10. T T
RADNKG	2.E5
ATMFILE	/atmosphere/atmos_runId134826_eventId53455709.txt
DIRECT	/coreas/runID_134826_eventID_53455709/proton/000000/

APPENDIX B

Analysis supplements

B.1 Influence of the radio noise on the measured quantity P_{10}

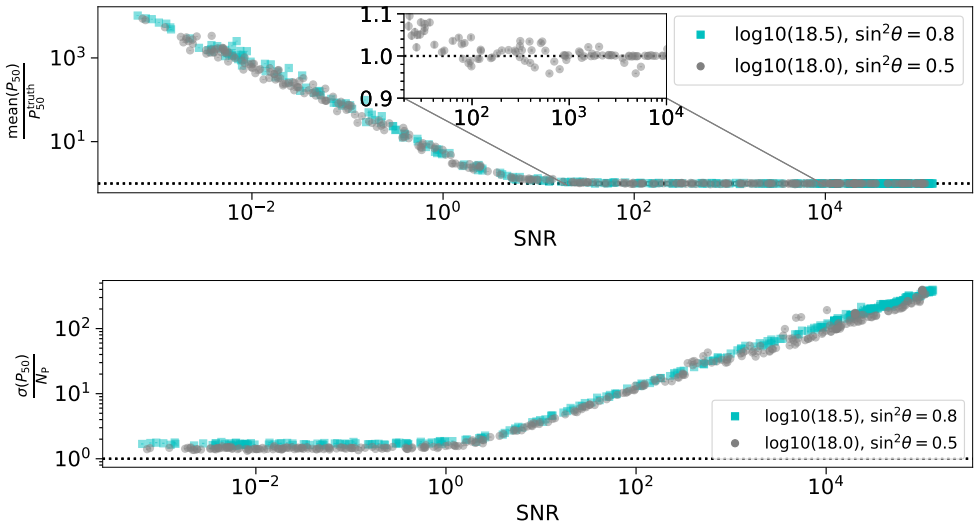


Figure B.1: Mean and standard deviation of the measured quantity P_{10} with modelled noise injected.

B.2 Parametrisation used in for the Log-Likelihood minimisation

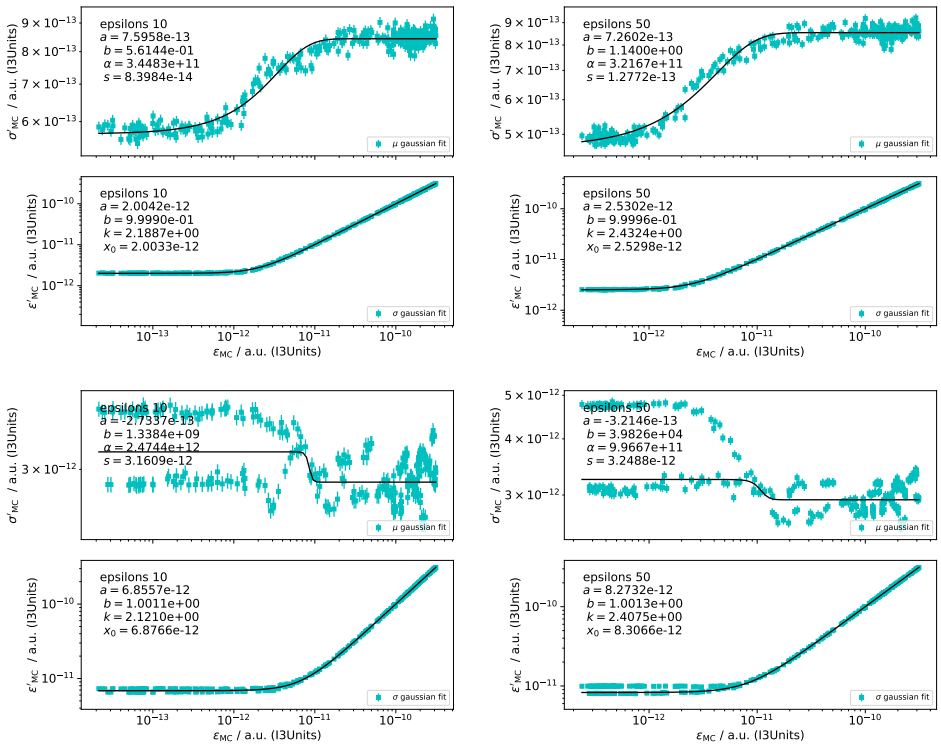


Figure B.2: Mean and sigma parametrisation of the Gaussian fits for the maximum of the Hilbert Envelope. Right column is for a 10 ns window and left for a 50 ns window. Top row, for the modelled noise, and bottom row for the measured noise

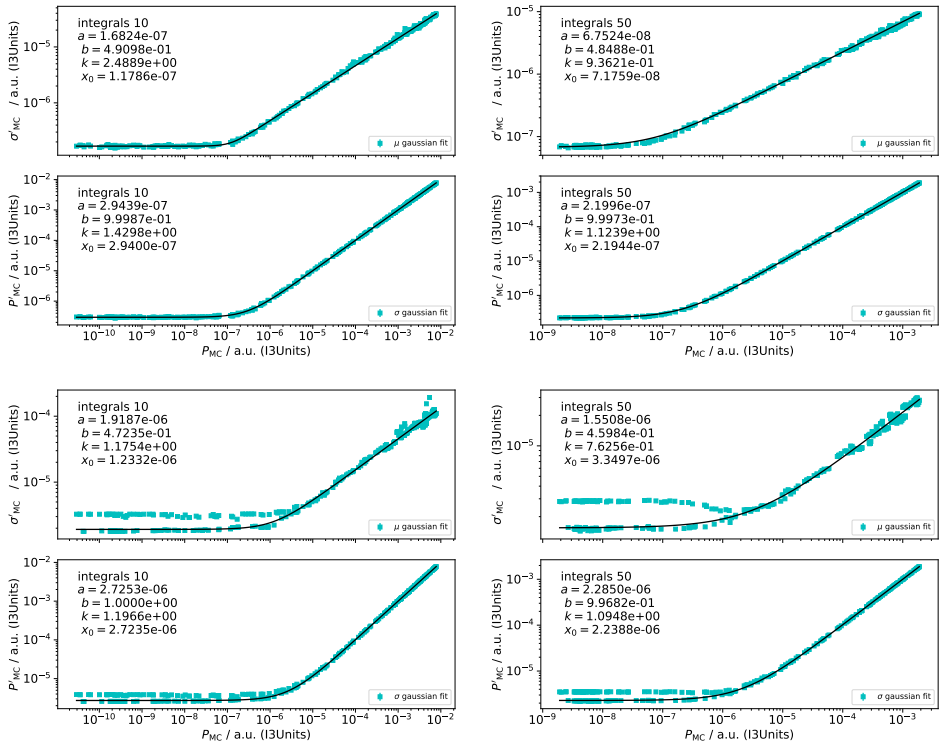


Figure B.3: Mean and sigma parametrization of the Gaussian fits for the power. The right column is for a 10 ns window, and the left one is for a 50 ns window. Top row, for the modeled noise, and the bottom row for the measured noise

B.3 IceTop reconstruction uncertainties

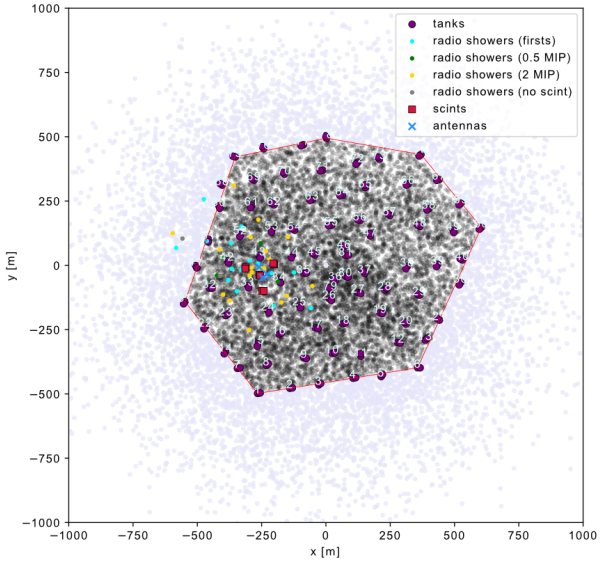


Figure B.4: All IceTop reconstructed core position. The simulations whose cores are in black are taken for the estimation of the uncertainty of the Laputop reconstruction.

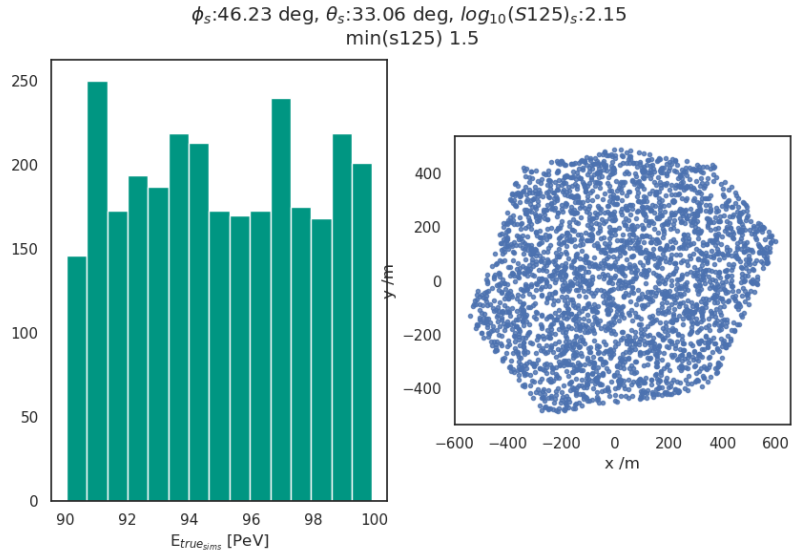


Figure B.5: Energy distribution and shower core positions for one event whose Laputop uncertainty is estimated.

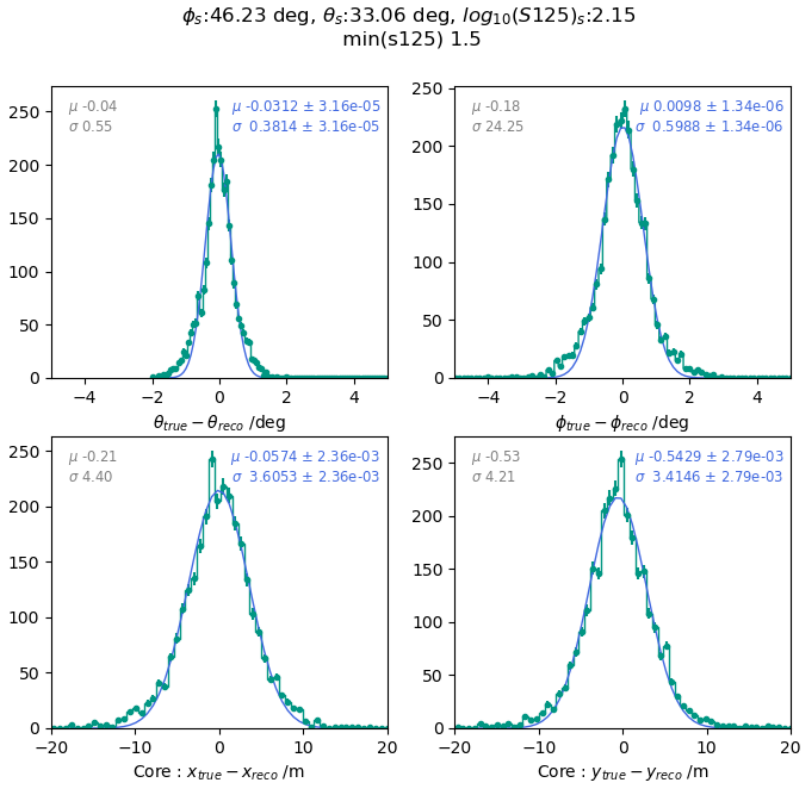


Figure B.6: Distribution of the shower reconstructed parameters with Laputop compared to their true value. Gaussian fitted to the distribution.

B.4 Supplementary plots of the Energy and X_{max} reconstruction

Table B.1: position of the plots for Figure B.7 and Figure B.8

χ^2 – maximum Hilbert Envelope	LLH – maximum Hilbert Envelope
χ^2 FWHM (50 bins)	LLH – FWHM (50 bins)
χ^2 – power (50 bins)	LLH – power (50 bins)

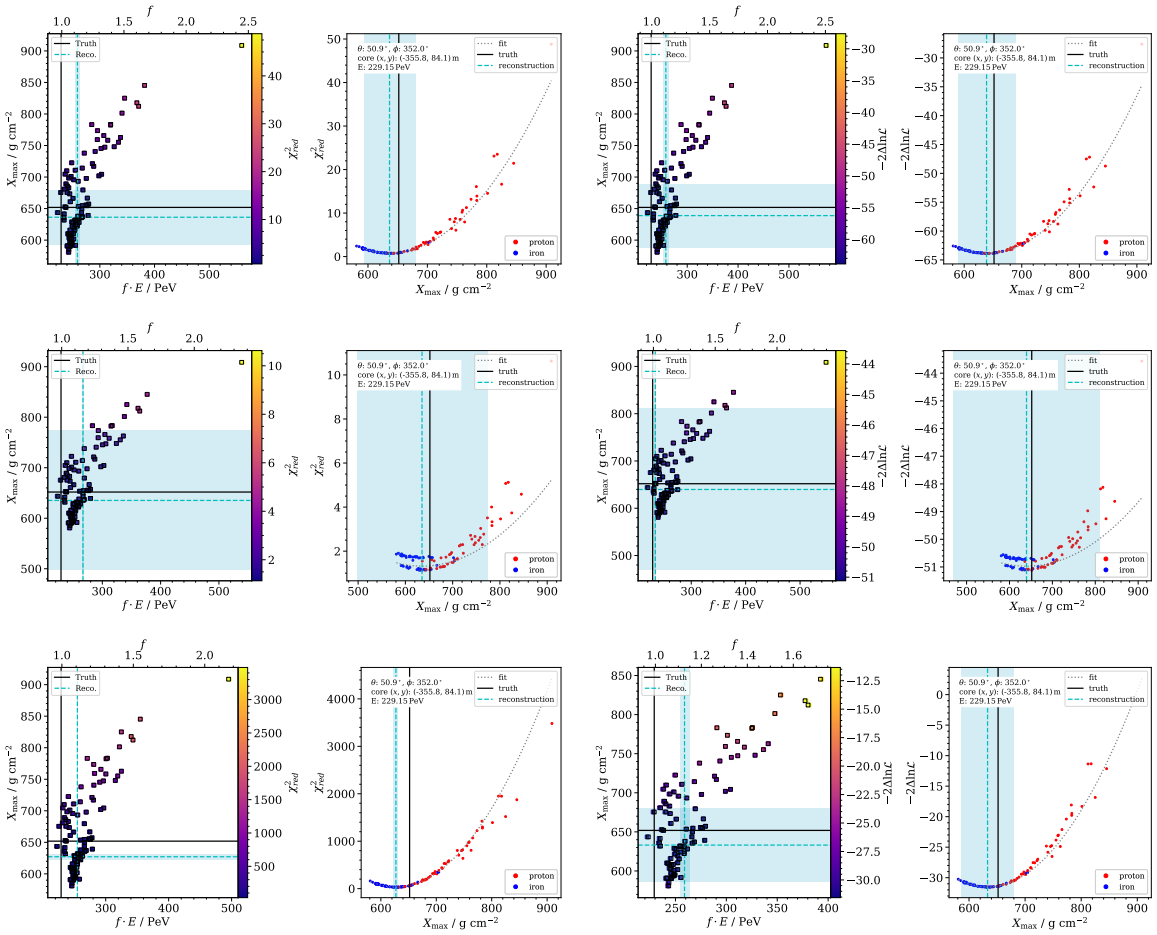


Figure B.7: Reconstruction of one mock event data of runid 134630 eventid 53455709 with different quantities and methods.

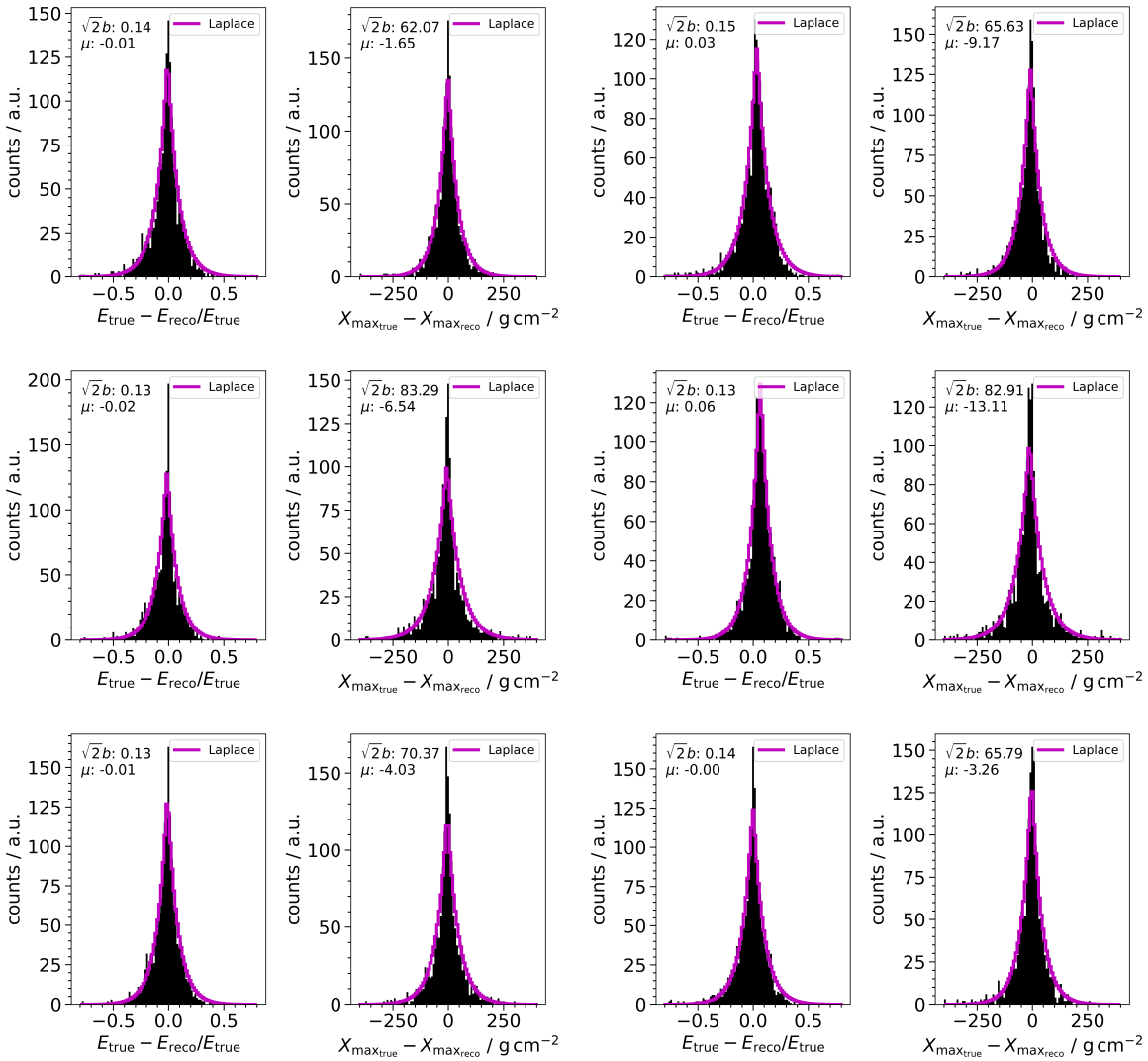


Figure B.8: Cumulative distributions of the 16 events with the Laplace fit using different quantities and methods with injected noise.

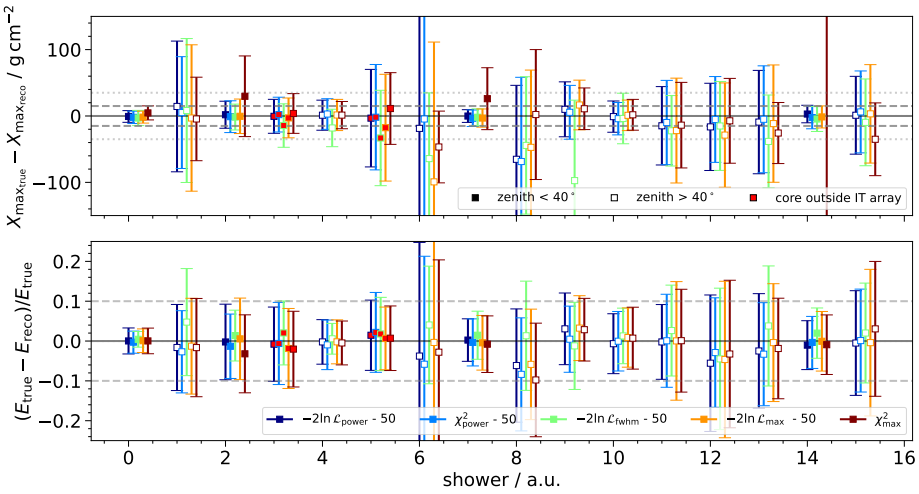


Figure B.9: Resolution and precision for energy and X_{max} of the five different methods on an event-per-event basis with simulated noise injected. Plotted shower per shower with a slight x-axis shift for better visualization.

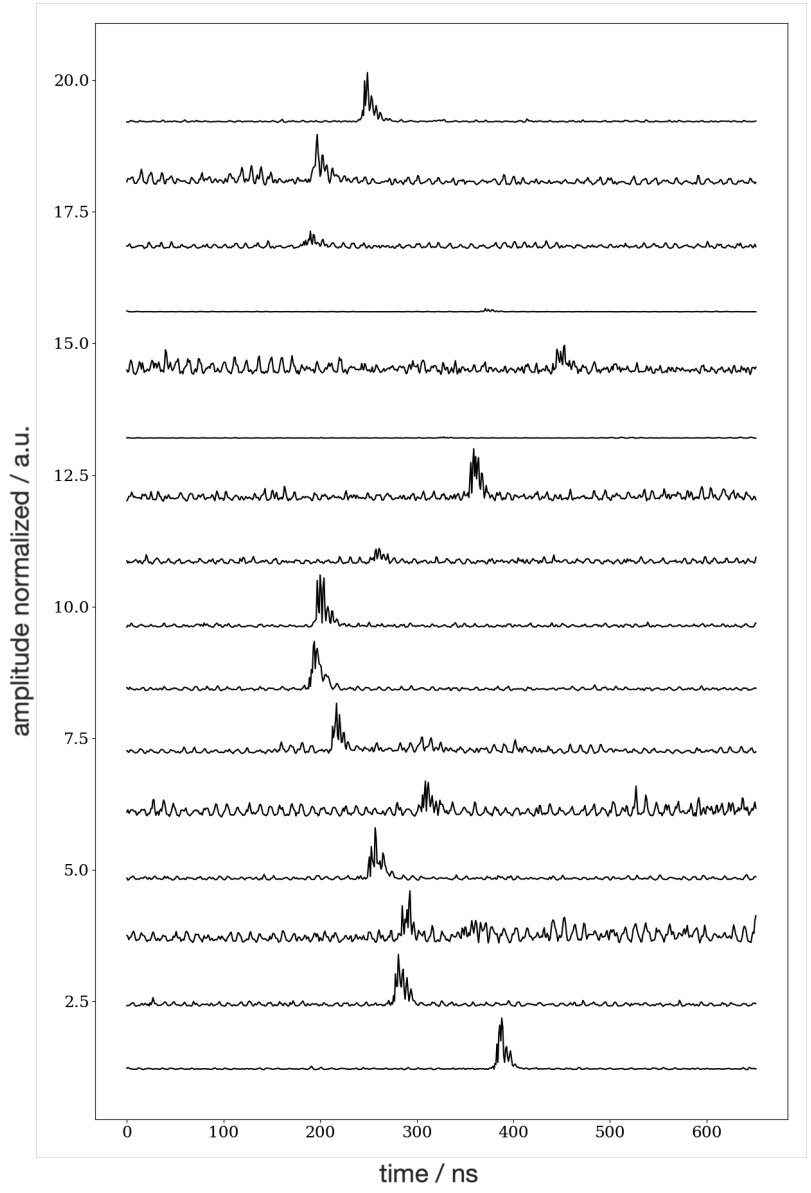


Figure B.10: Beamforming of all the channels normalized to the maximum on the measured 16 event waveforms.

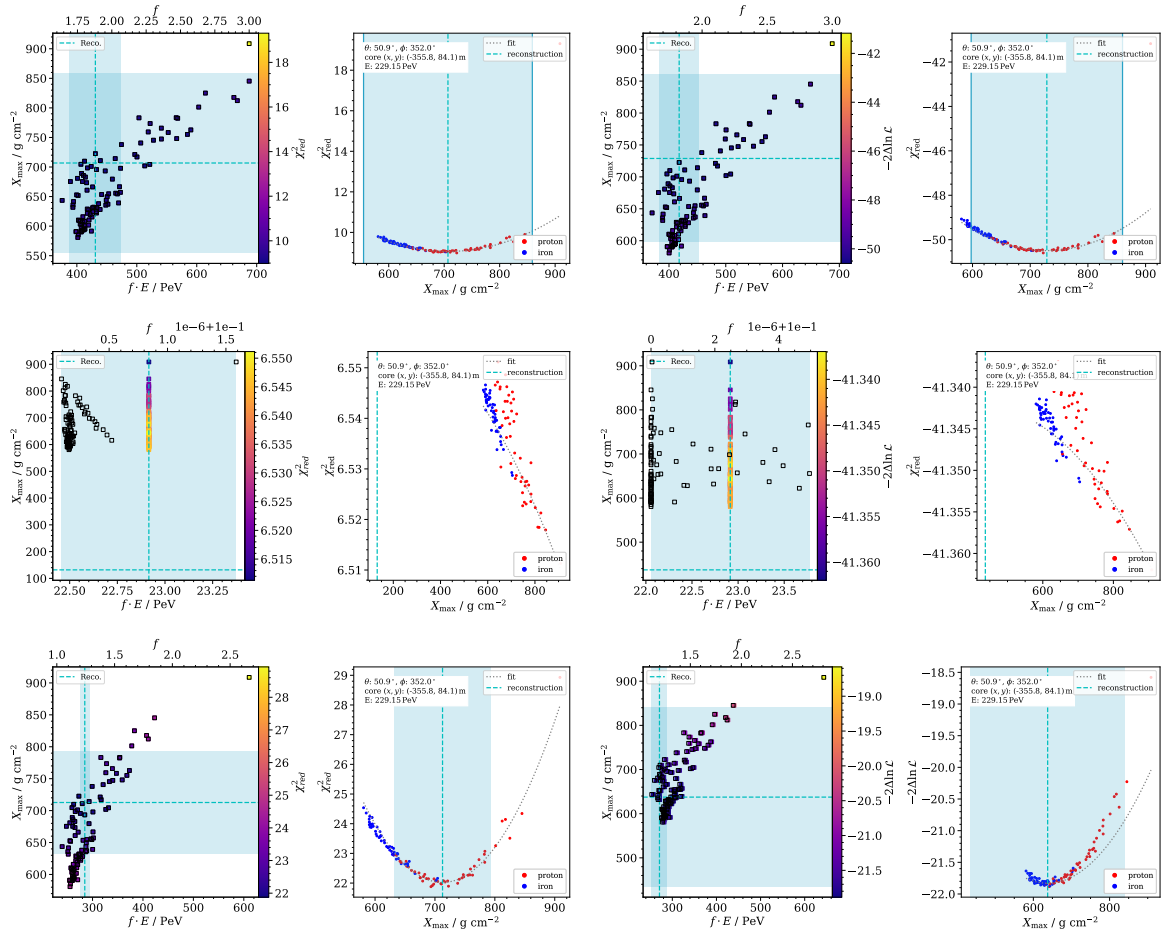


Figure B.11: Application of different methods and different quantities on the measured waveforms of runid 134630 eventid 53455709.

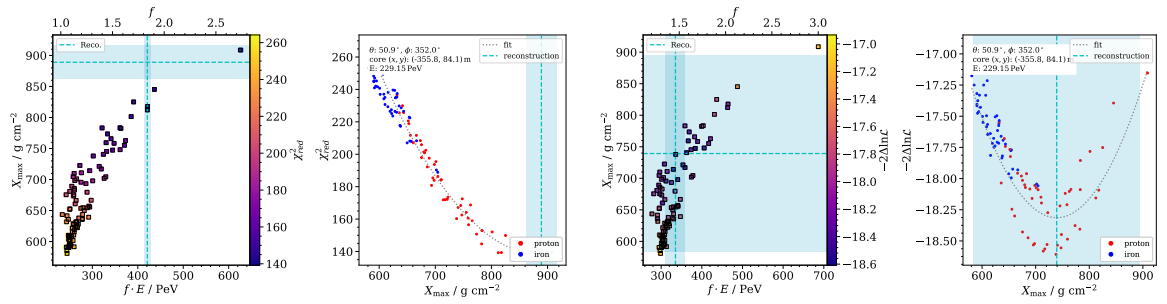


Figure B.12: Two examples of the 10 ns window on the measurement with χ^2 of power on the left and LLH of power on the right (runid 134630 eventid 53455709).

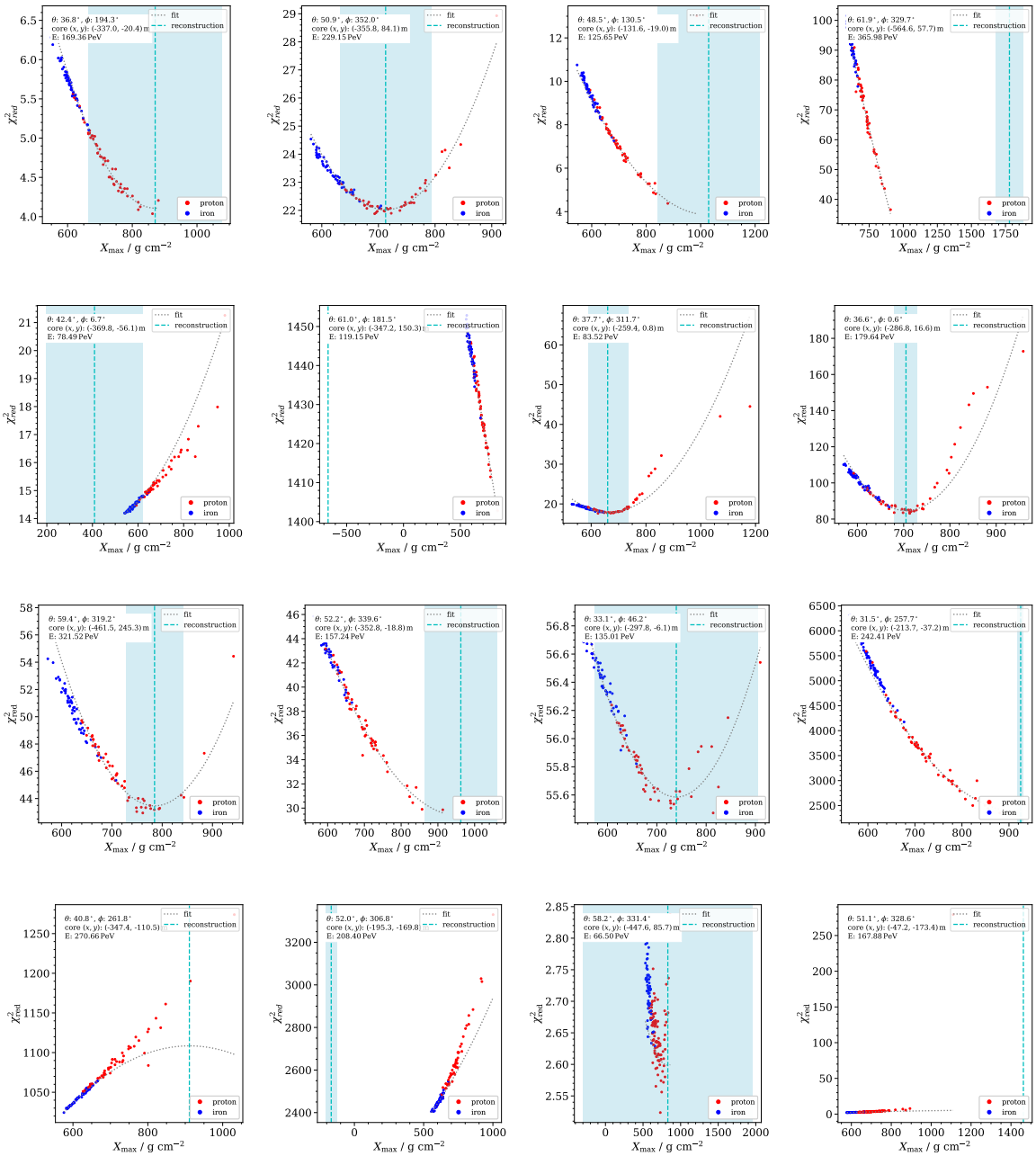


Figure B.13: χ^2 using the power for a 50 ns window applied on the 16 measured events in the set.

C APPENDIX

Measurements setups

C.1 Measurements setups

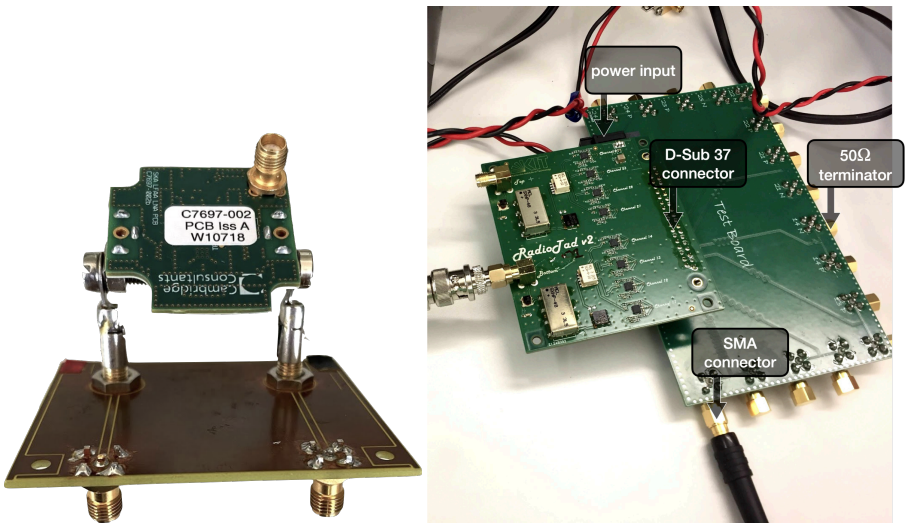


Figure C.1: Left: Connector board of the LNAs with an LNA installed. Right: Connector board of the radioTad.

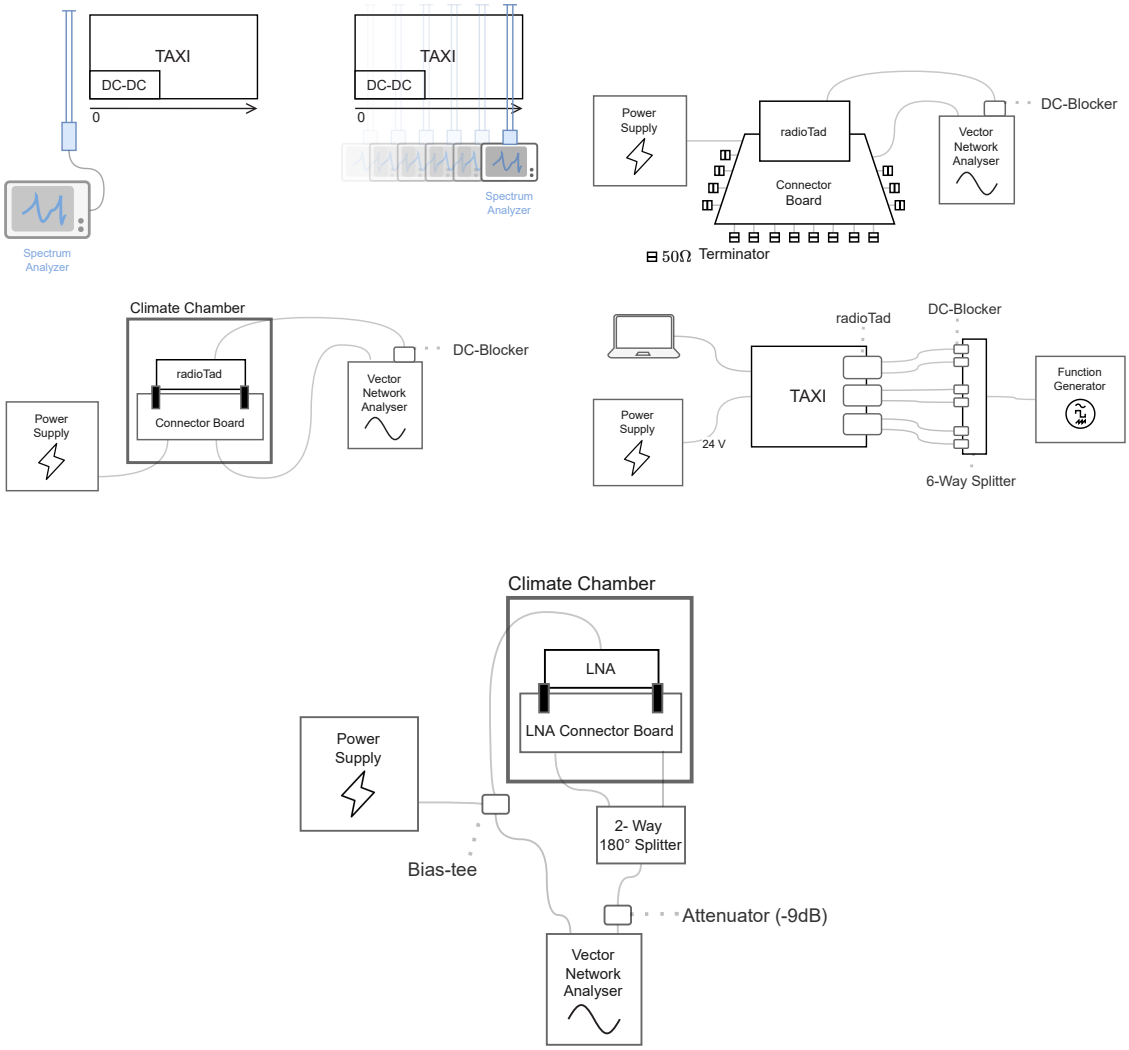


Figure C.2: Schematics of the diverse experimental setups used. From left to right, top to bottom: setup for the RFI measurements from the TAXI board, setup of the channels measurements of the radioTad, also used for the pre-calibration of the radioTad in production, setup of the temperature measurement of the radioTad, setup for the TAXI–radioTad complete calibration, and setup of the temperature measurement of the LNA.

C.2 Supplementary calibration information

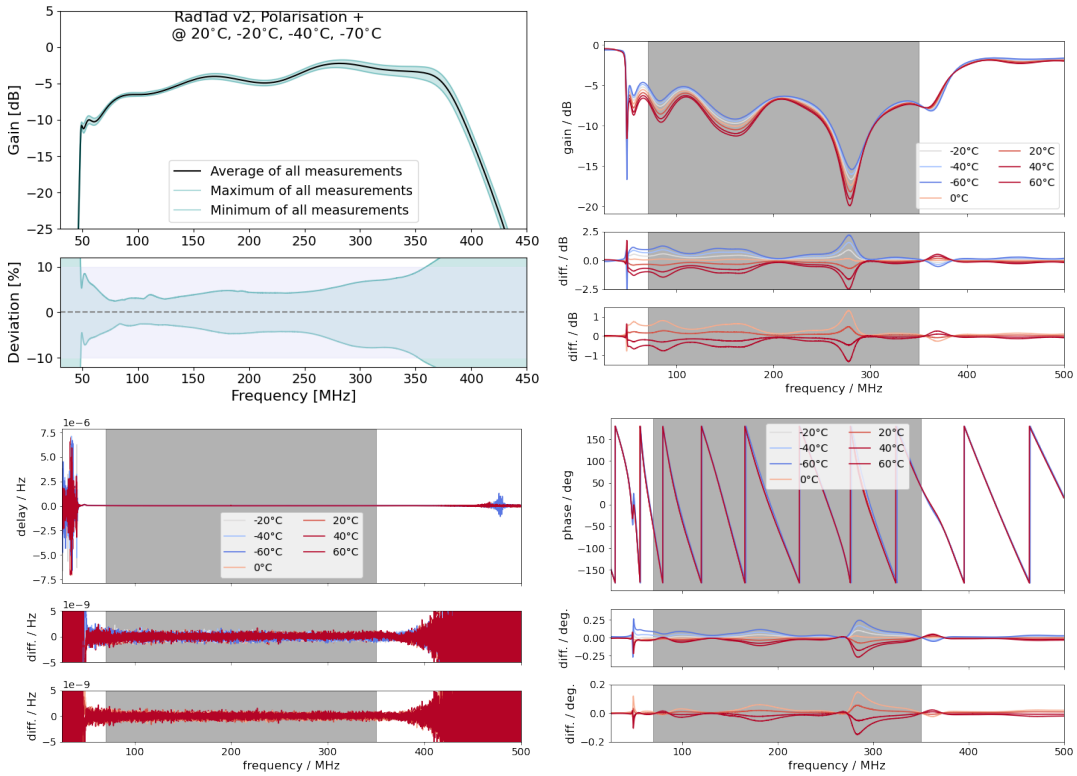


Figure C.3: RadioTad v2. Upper left: Variation in the gain of all channels combined at 4 different temperatures (the band represents the maximum and minimum values). Upper right: S_{11} variation of one channel (21N). Lower left: S_{21} group delay variation of one channel (21N). Lower right: S_{21} phase variation of one channel (21N).

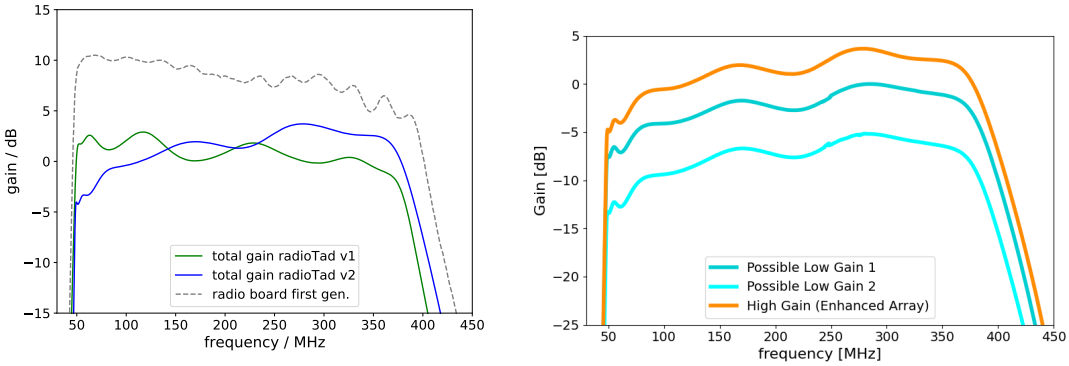


Figure C.4: Left: radioTad v1 and radioTad v2 compare to their predecessor. Right: Different gain tested, by modifying the amplification circuit, for the radioTad v2.

APPENDIX D

Mechanical drawing

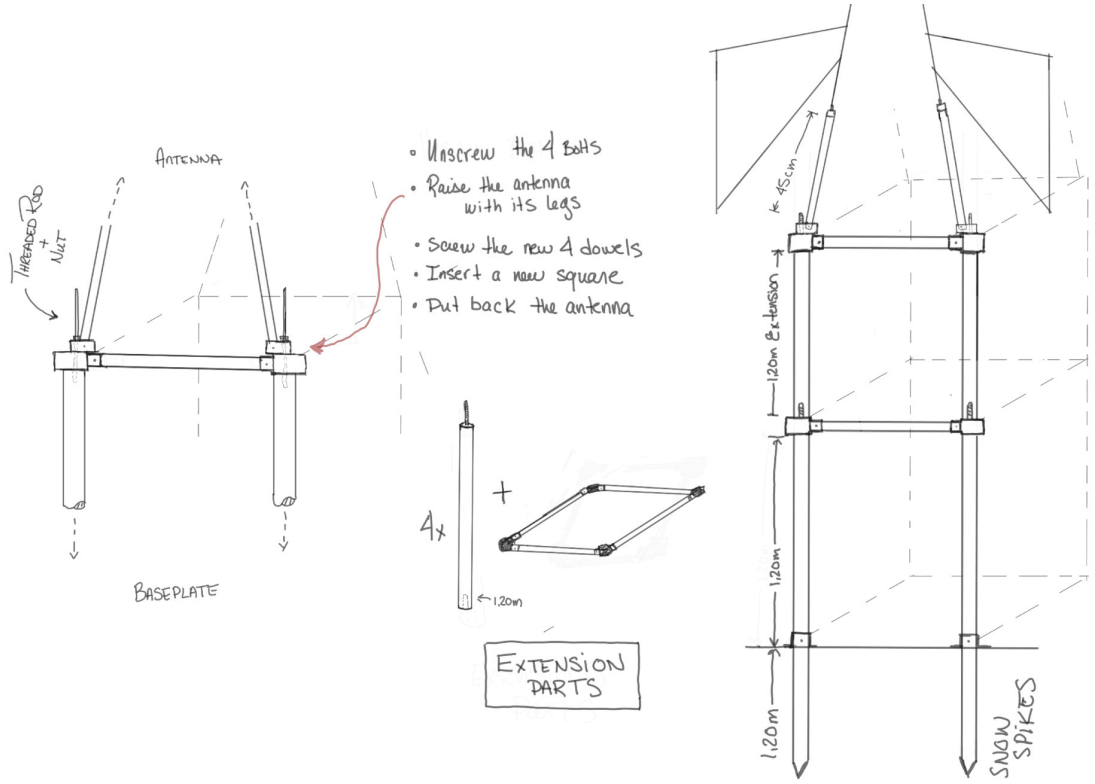
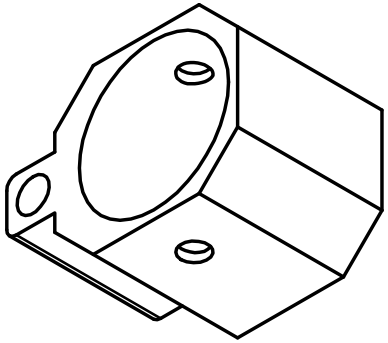
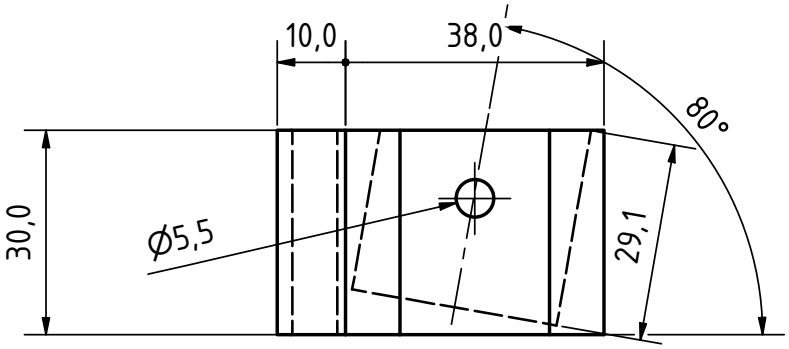
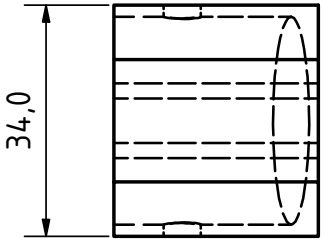
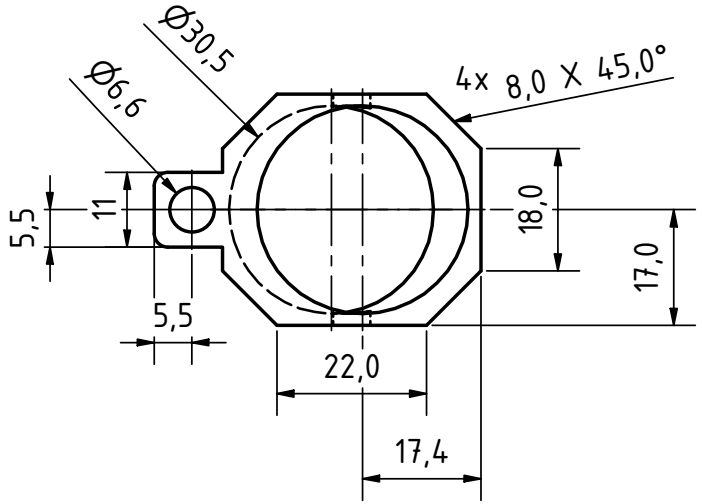


Figure D.1: Mount v2 extension



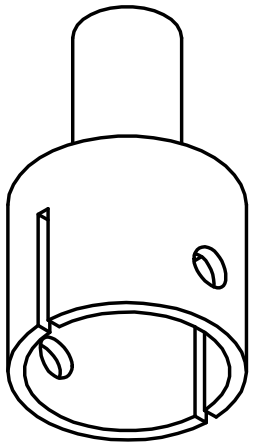


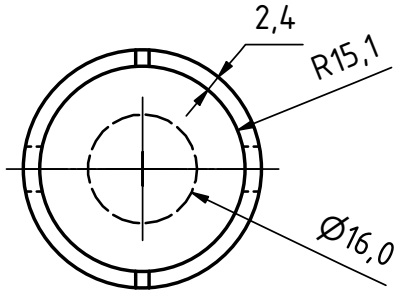


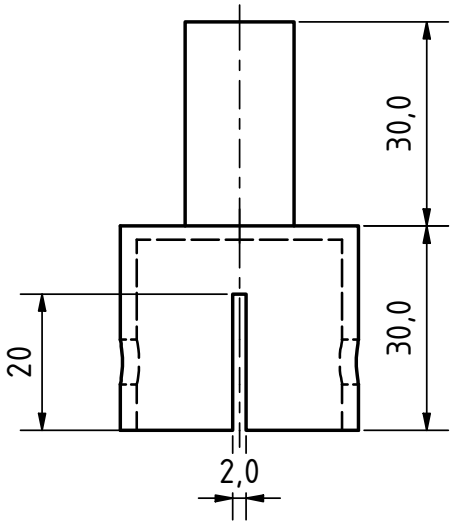


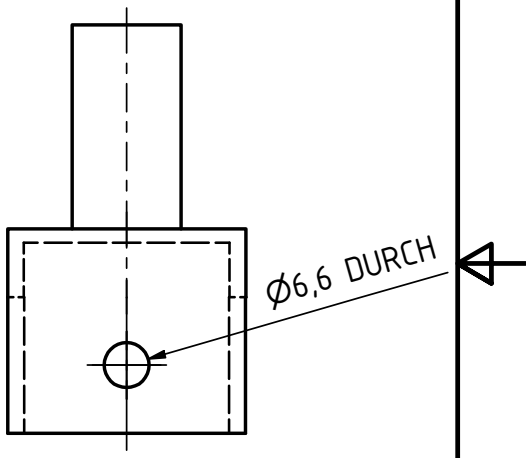
ALUMINUM

1:1						ROXANNE TURCOTTE
					Datum	Name
				Gezeichnet	08.11.2020	df3866
				Kontrolliert		
				Norm		
						angled_adapter_prod 1
Status	Änderungen	Datum	Name			A4



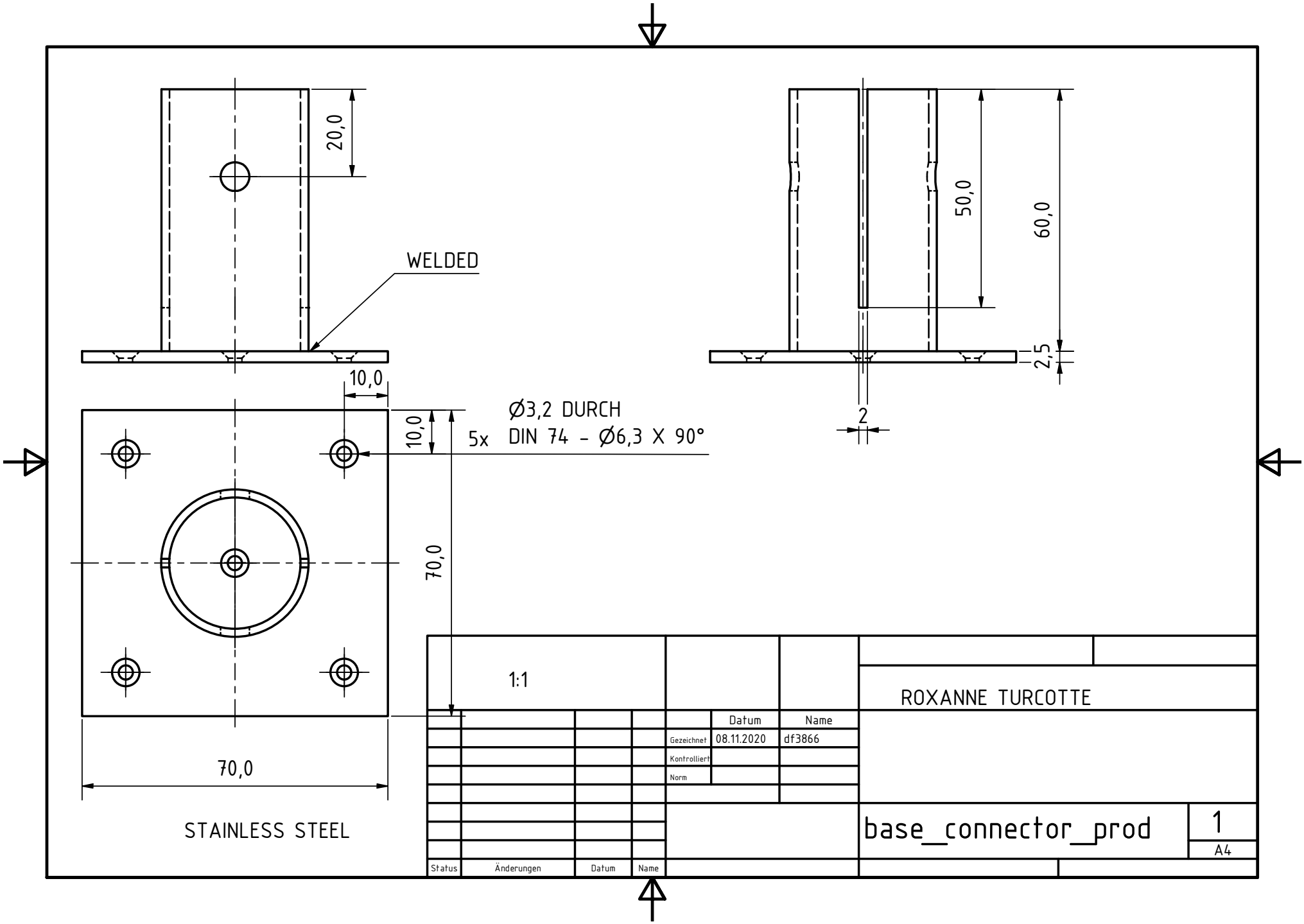


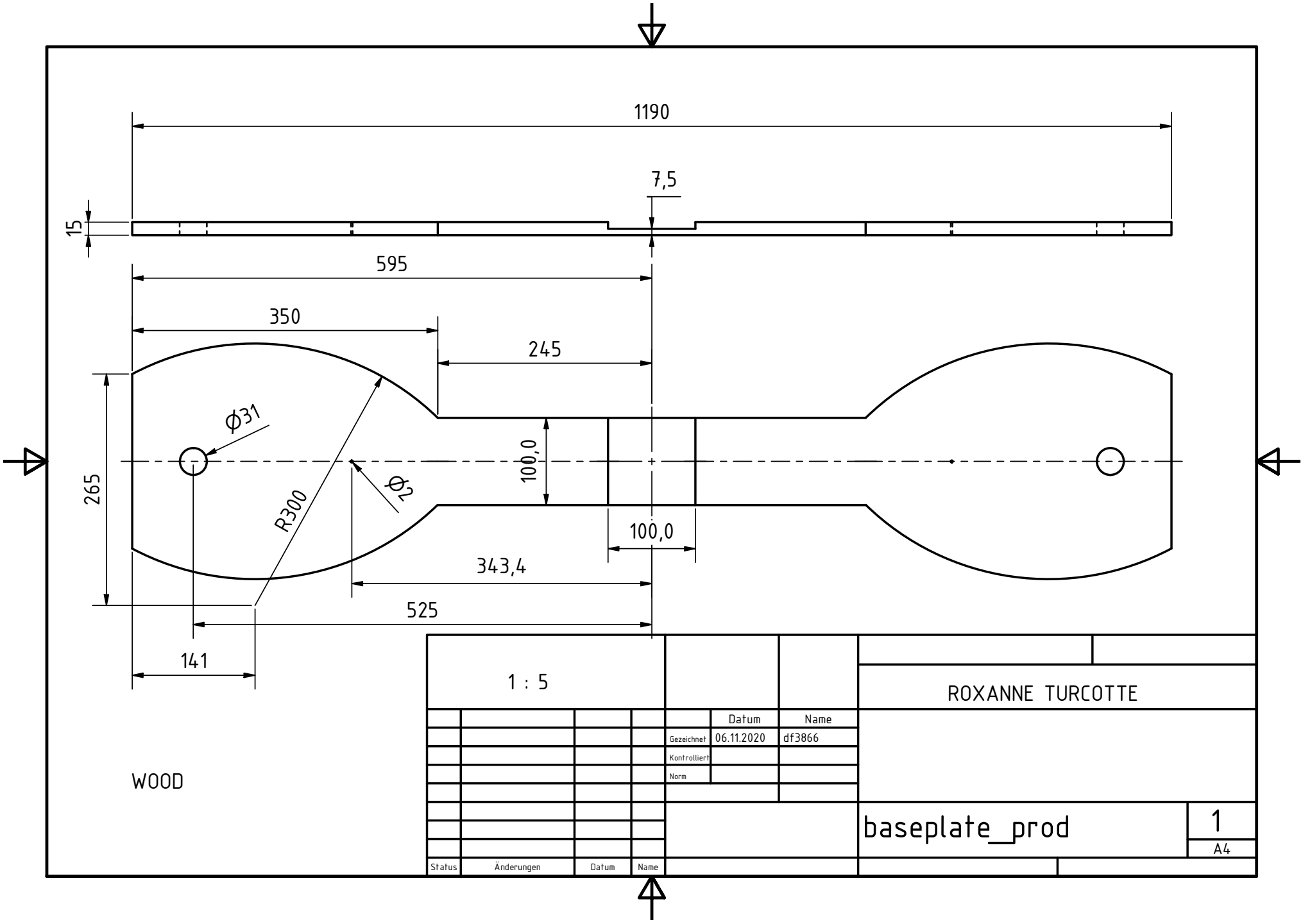


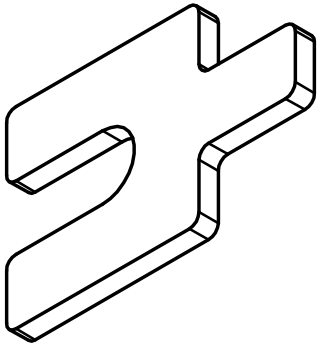


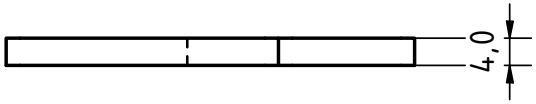
GLASS FIBER

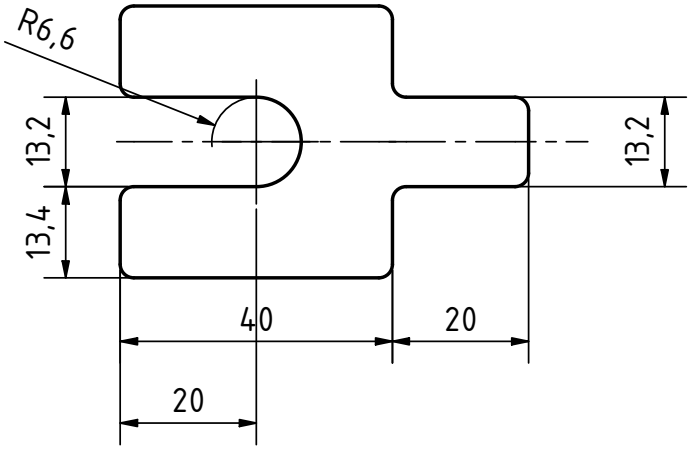
						ROXANNE TURCOTTE
					Datum	Name
				Gezeichnet	08.11.2020	df3866
				Kontrolliert		
				Norm		
						antenna_connector_prod 1 A4
Status	Änderungen	Datum	Name			


20,0
WELDED
50,0
60,0
2,5
10,0
10,0
5x Ø3,2 DURCH
DIN 74 - Ø6,3 X 90°
2
70,0
70,0
STAINLESS STEEL
1:1
ROXANNE TURCOTTE
Datum
Name
Gezeichnet
08.11.2020
df3866
Kontrolliert
Norm
base_connector_prod
1
A4
Status
Änderungen
Datum
Name


1190
7,5
15
595
350
245
ϕ31
265
R300
ϕ2
100,0
100,0
343,4
525
141
1 : 5
WOOD
ROXANNE TURCOTTE
Datum
Name
Gezeichnet
06.11.2020
df3866
Kontrolliert
Norm
baseplate_prod
1
A4
Status
Änderungen
Datum
Name

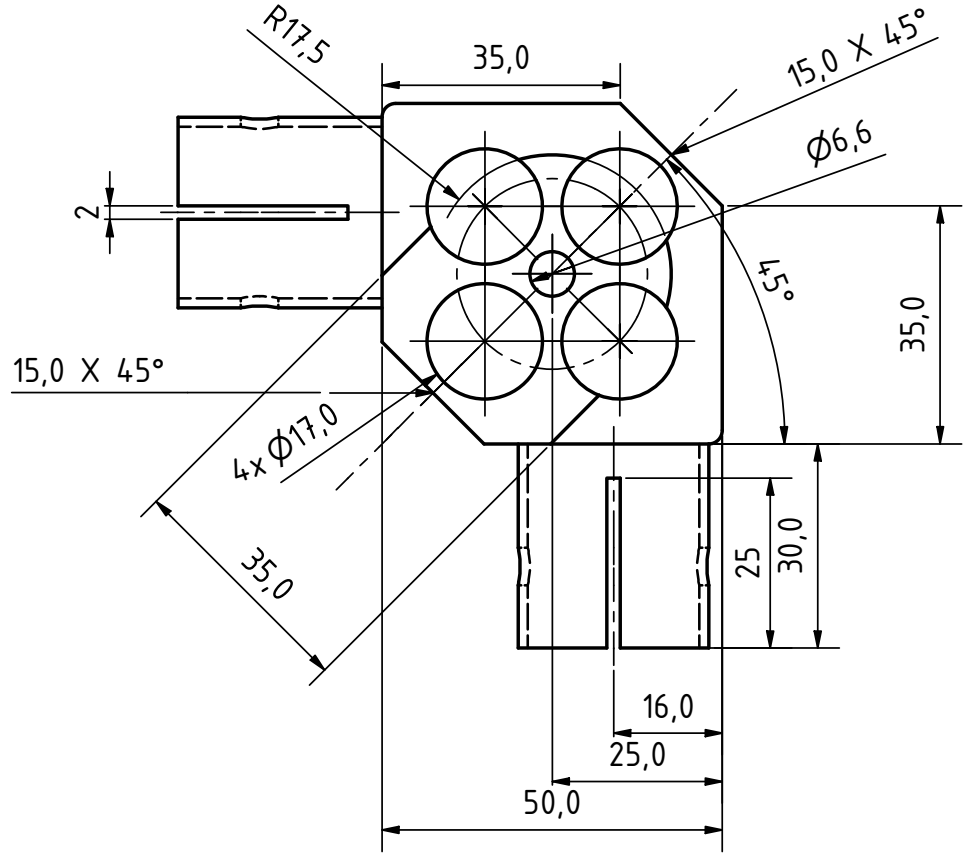


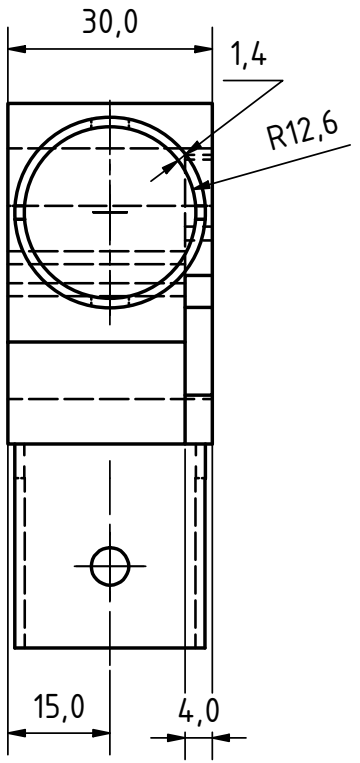




ALUMINUM

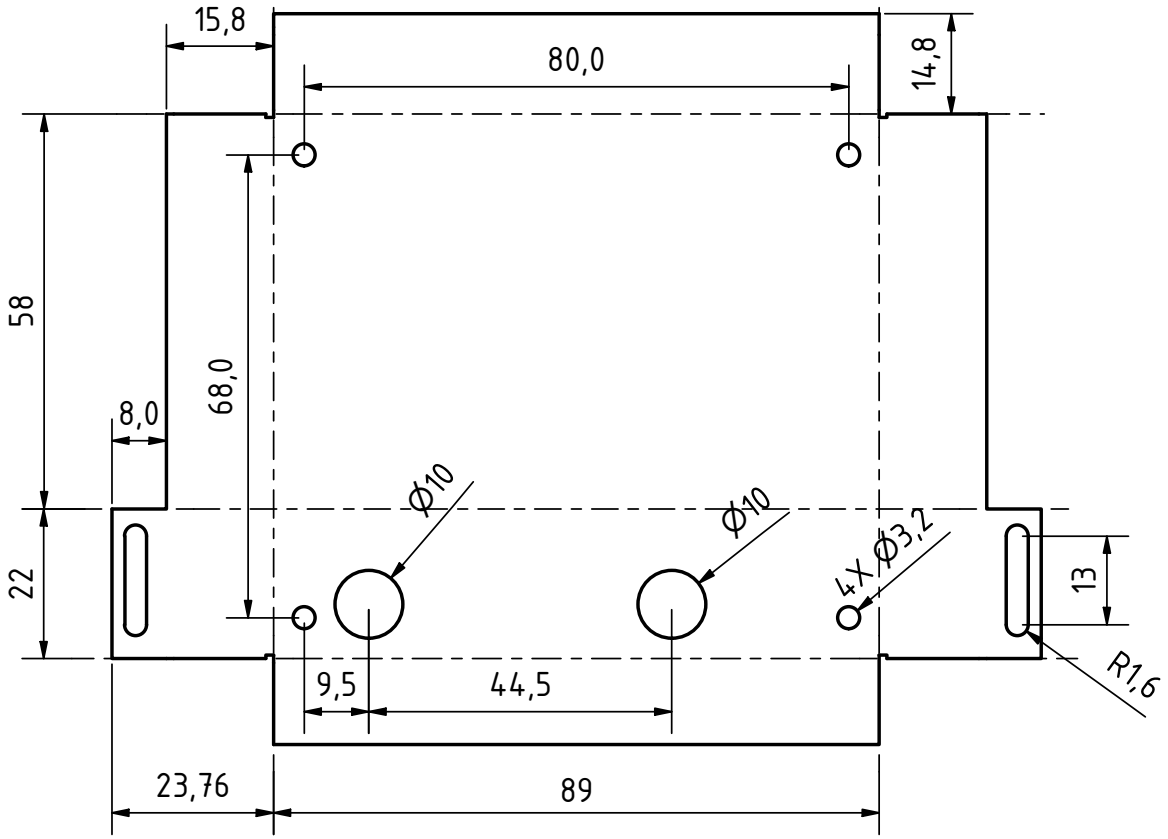
1:1						ROXANNE TURCOTTE
					Datum	Name
				Gezeichnet	08.11.2020	df3866
				Kontrolliert		
				Norm		
						shim_prod 1
						A4
Status	Änderungen	Datum	Name			





ALUMINUM

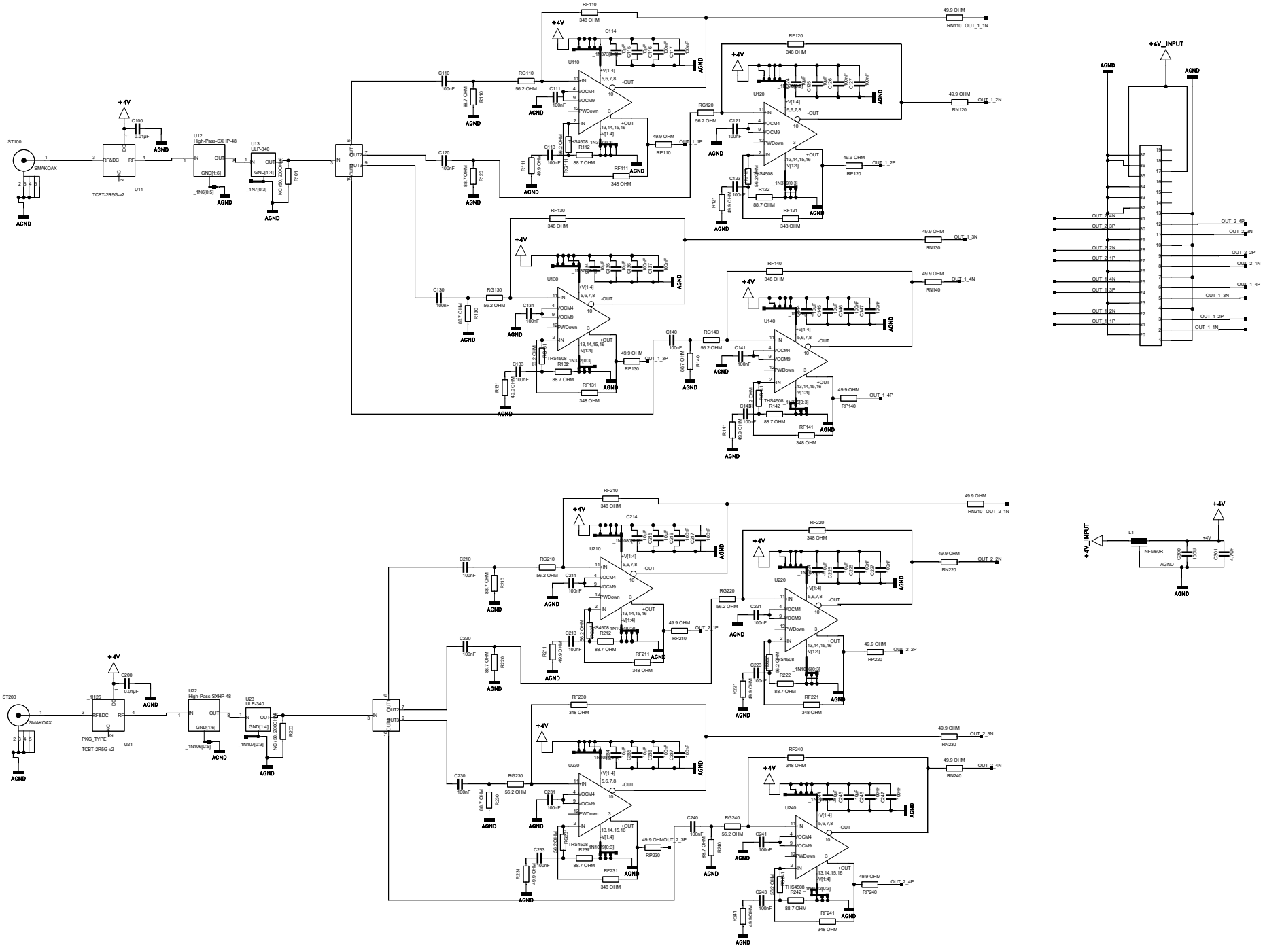
1:1						ROXANNE TURCOTTE
					Datum	Name
				Gezeichnet	08.11.2020	df3866
				Kontrolliert		
				Norm		
						square_connector_prod 1 A4
Status	Änderungen	Datum	Name			

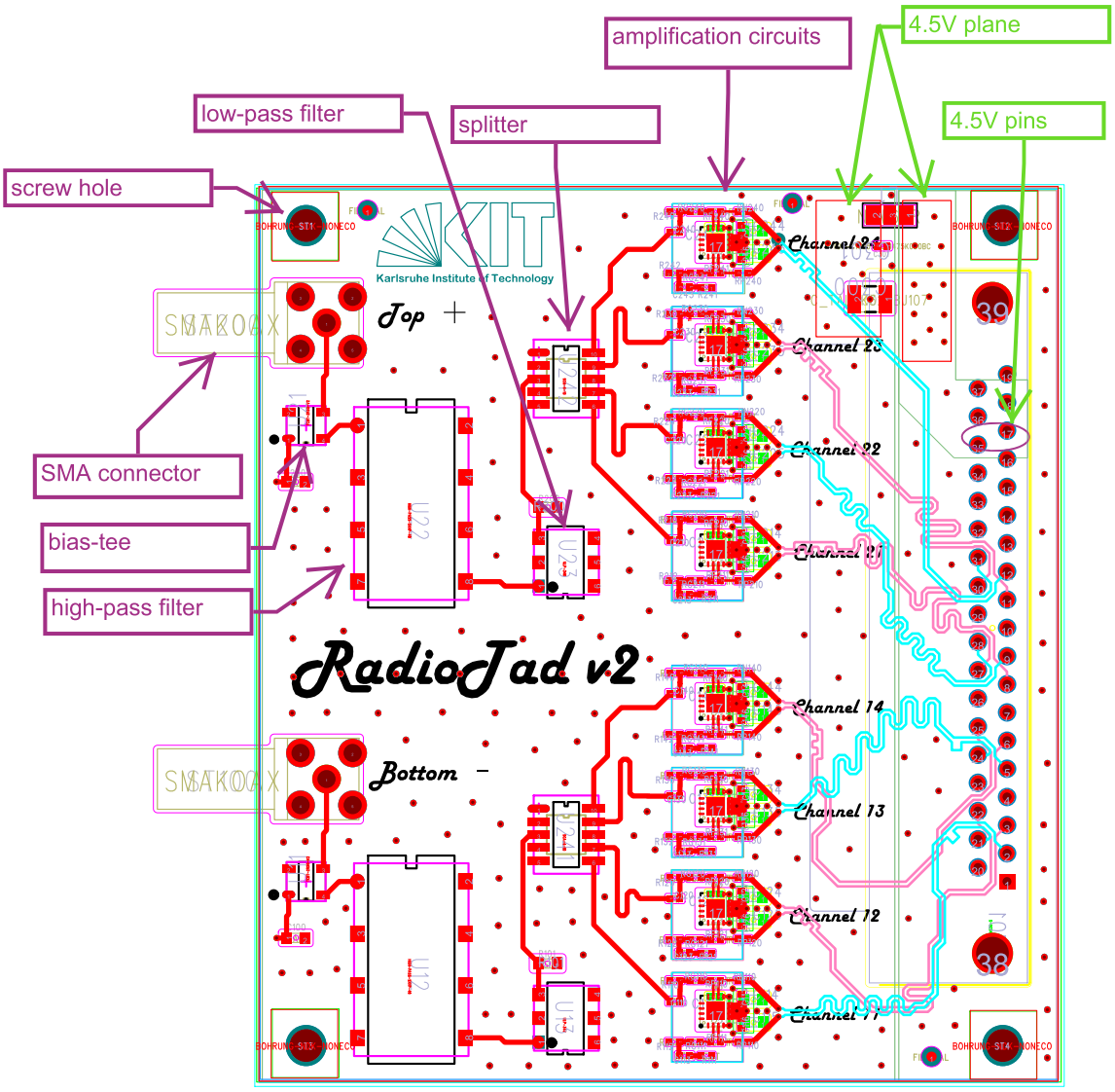


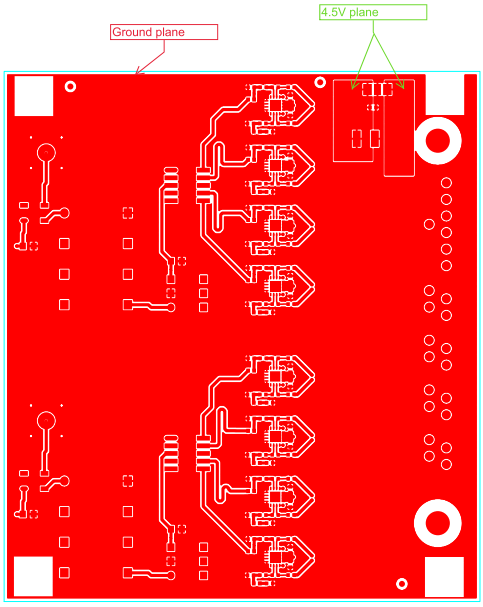
						ROXANNE TURCOTTE
					Datum	Name
				Gezeichnet	30.09.2020	df3866
				Kontrolliert		
				Norm		
						radio_shieldingv2_unfolded
Status	Änderungen	Datum	Name			A4

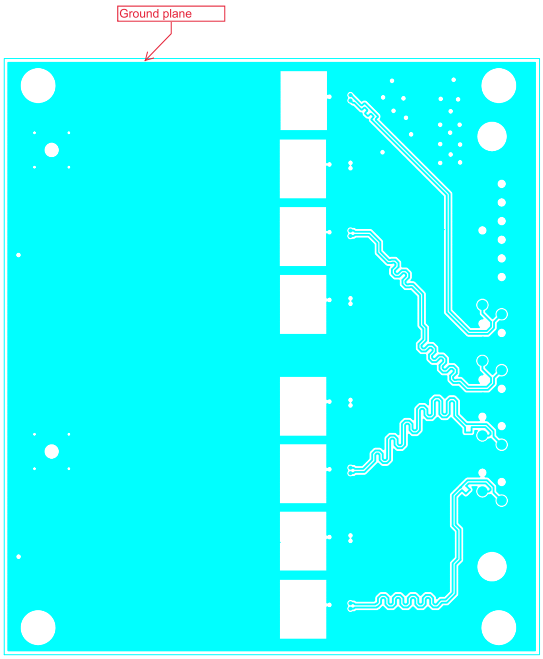
APPENDIX E

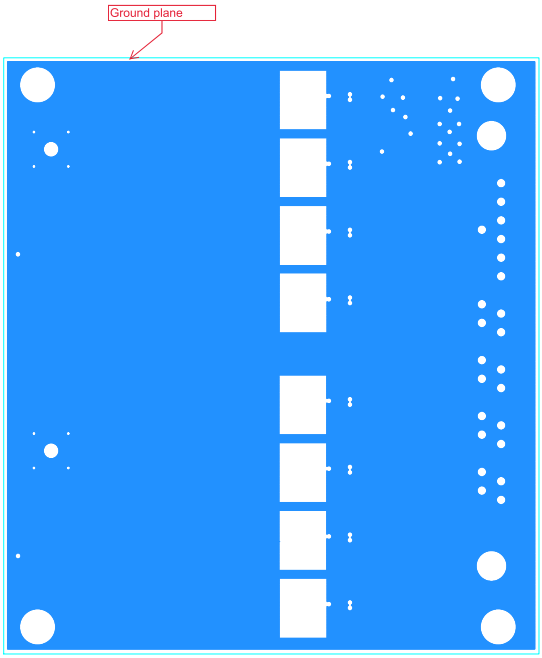
Electrical drawing

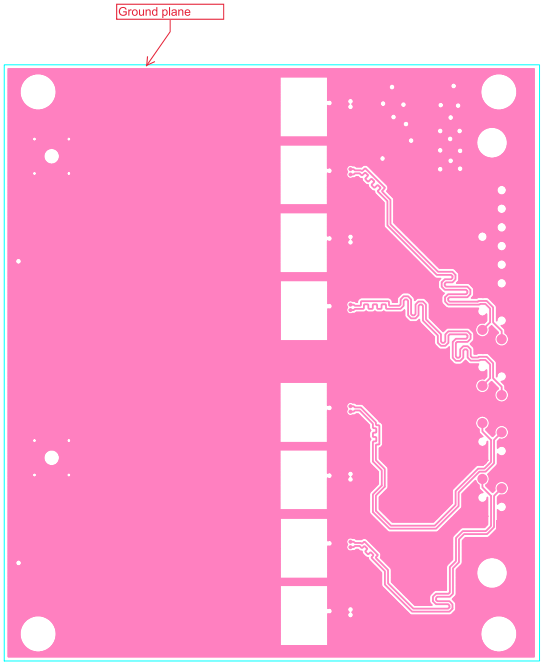

+4V_INPUT
AGND
OUT_2_4N
OUT_2_3P
OUT_2_3N
OUT_2_2N
OUT_2_2P
OUT_2_1P
OUT_2_1N
OUT_1_4N
OUT_1_4P
OUT_1_3N
OUT_1_3P
OUT_1_2N
OUT_1_2P
OUT_1_1P
OUT_1_1N
+4V
L1
NFM60R
C300
100pF
C301
4.7µF

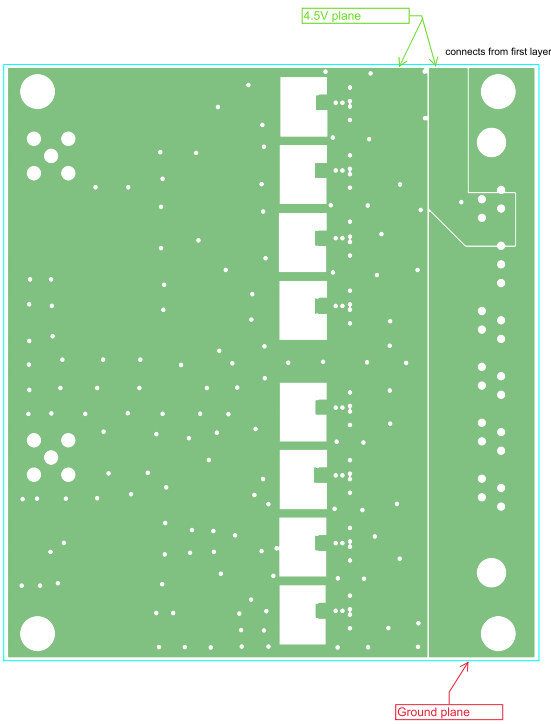

amplification circuits
4.5V plane
low-pass filter
splitter
4.5V pins
screw hole
SMA connector
bias-tee
high-pass filter
Karlsruhe Institute of Technology
Top
Bottom
RadioTad v2
Channel 24
Channel 23
Channel 22
Channel 21
Channel 14
Channel 13
Channel 12
Channel 11

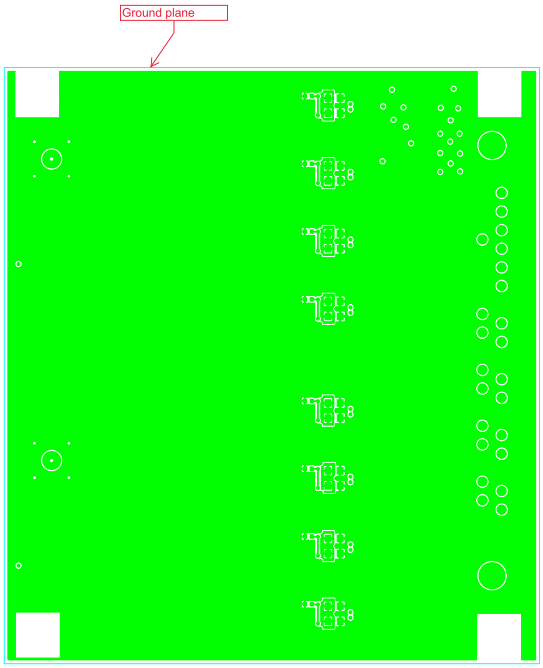

Ground plane
4.5V plane

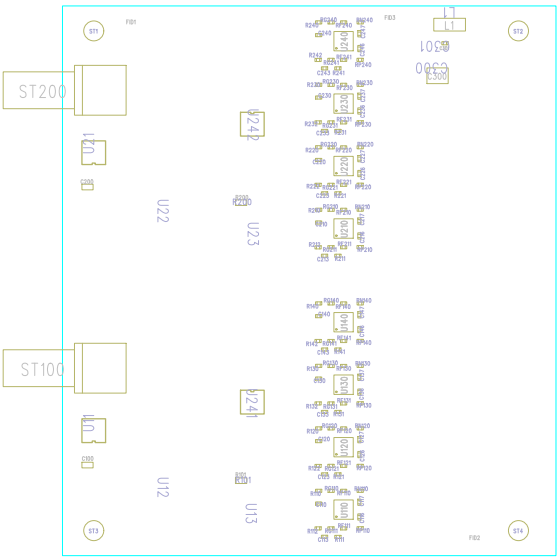

Ground plane

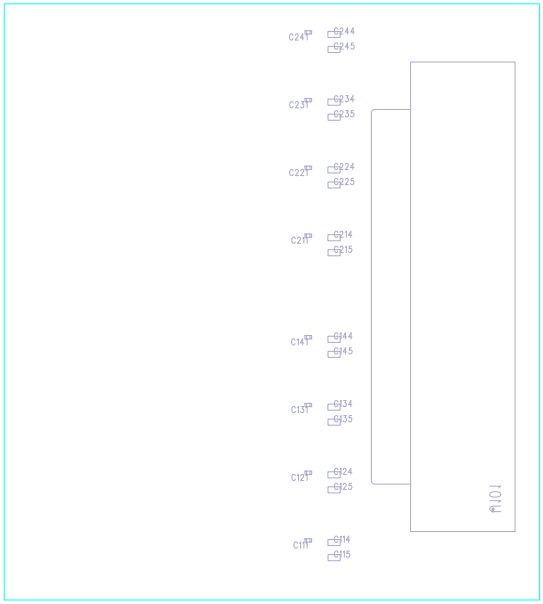

Ground plane

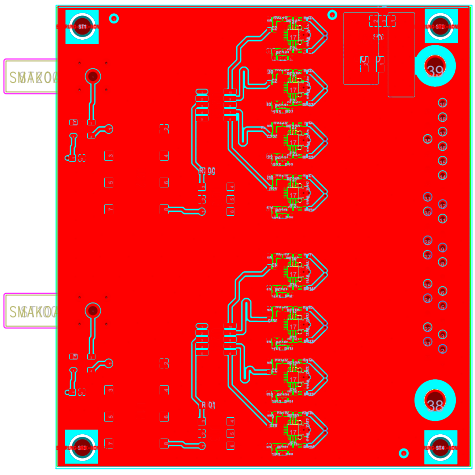

Ground plane


4.5V plane
connects from first layer
Ground plane


Ground plane


ST200
ST100
U22
U23
U12
U13


C24
C23
C22
C21
C14
C13
C12
C11
U101


SMAKON
SMAKON

APPENDIX F

In-Ice coincidences

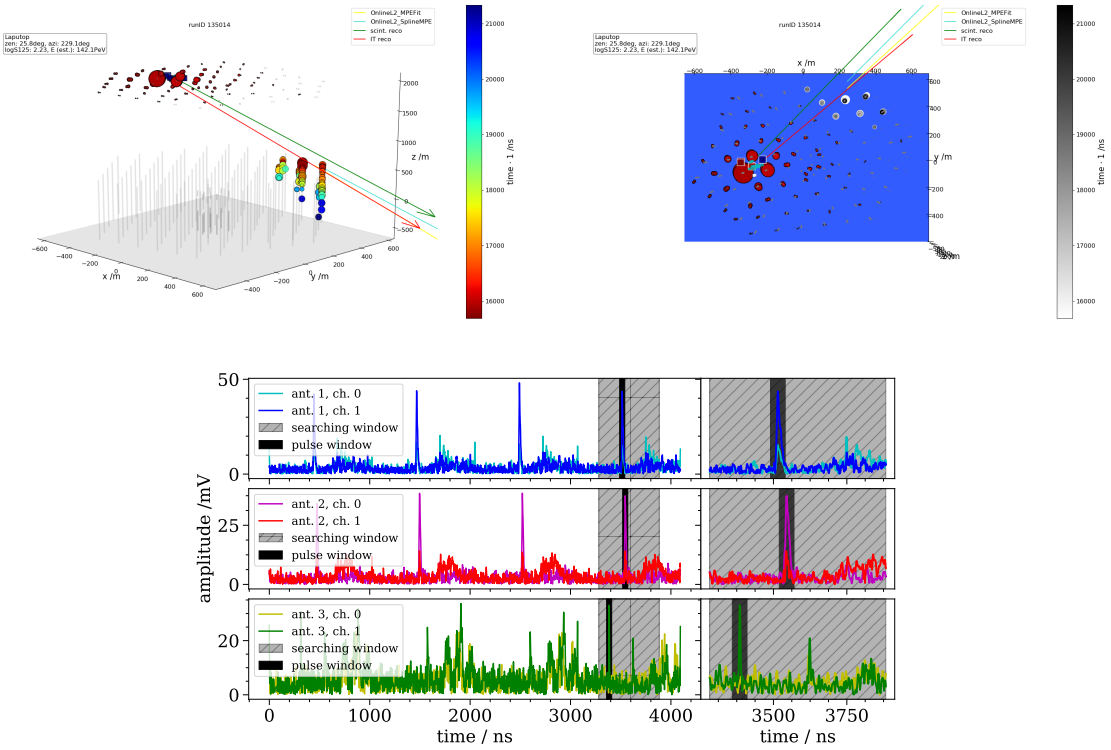


Figure F.1: runid 135014, eventid 4744026 (2021)

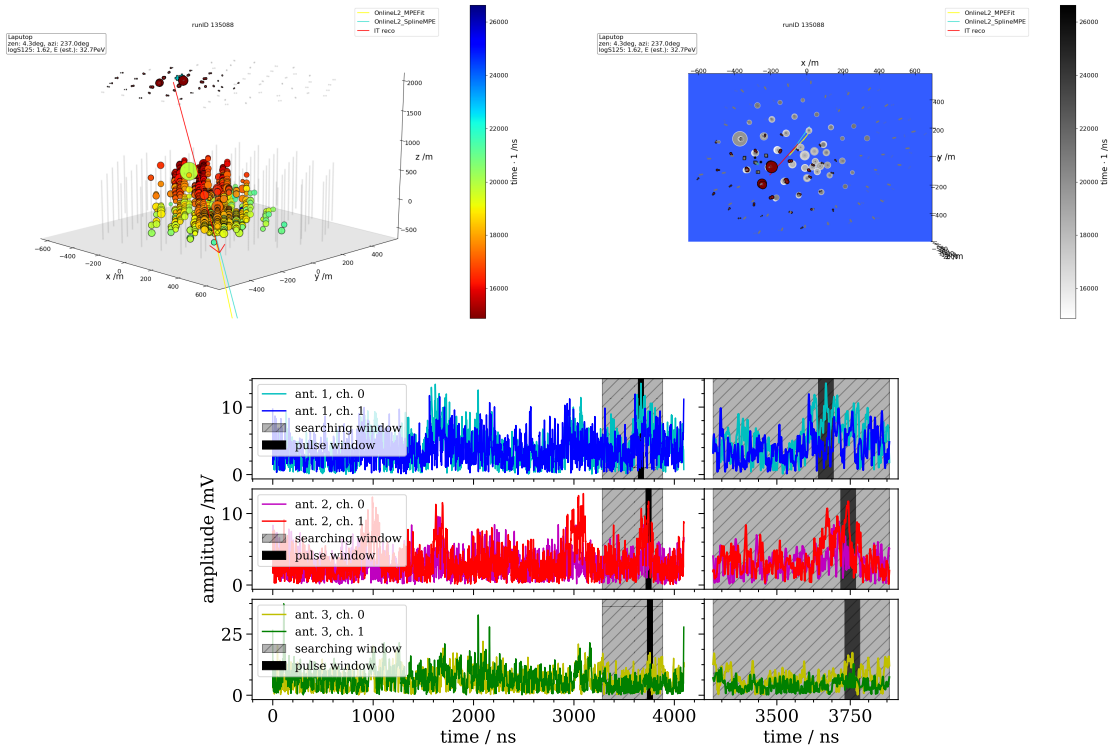


Figure F.2: runid 135088, eventid 27842235 (2021)

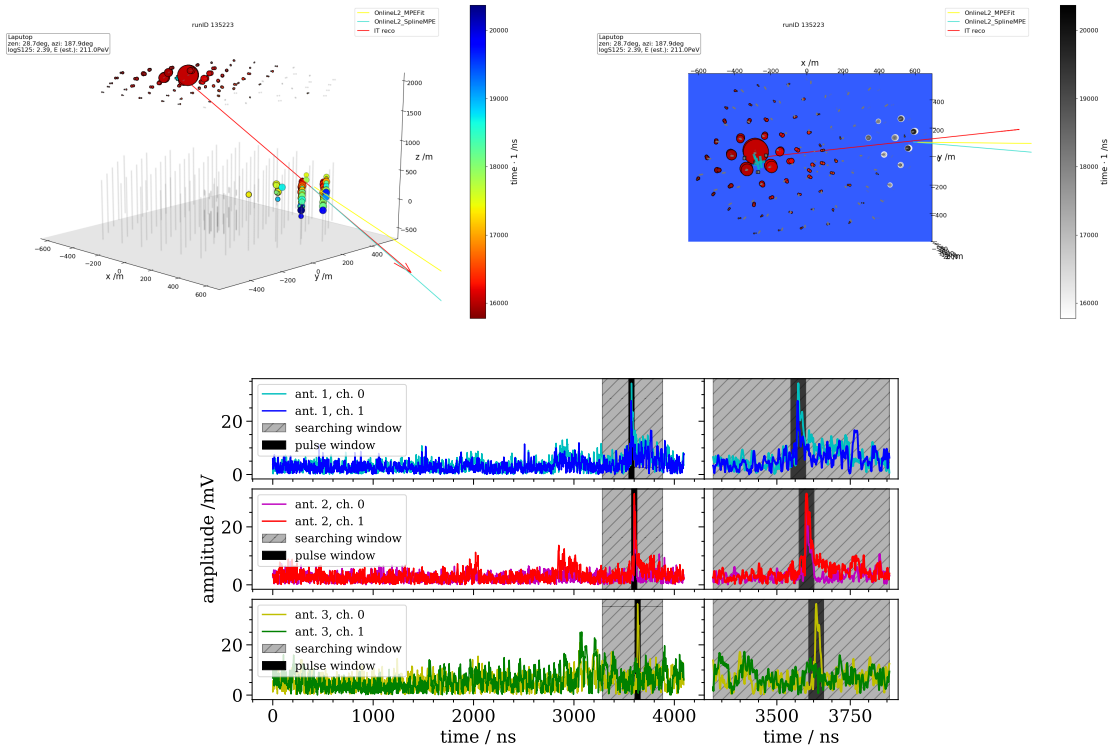


Figure F.3: runid 135223, eventid 18225942 (2021)

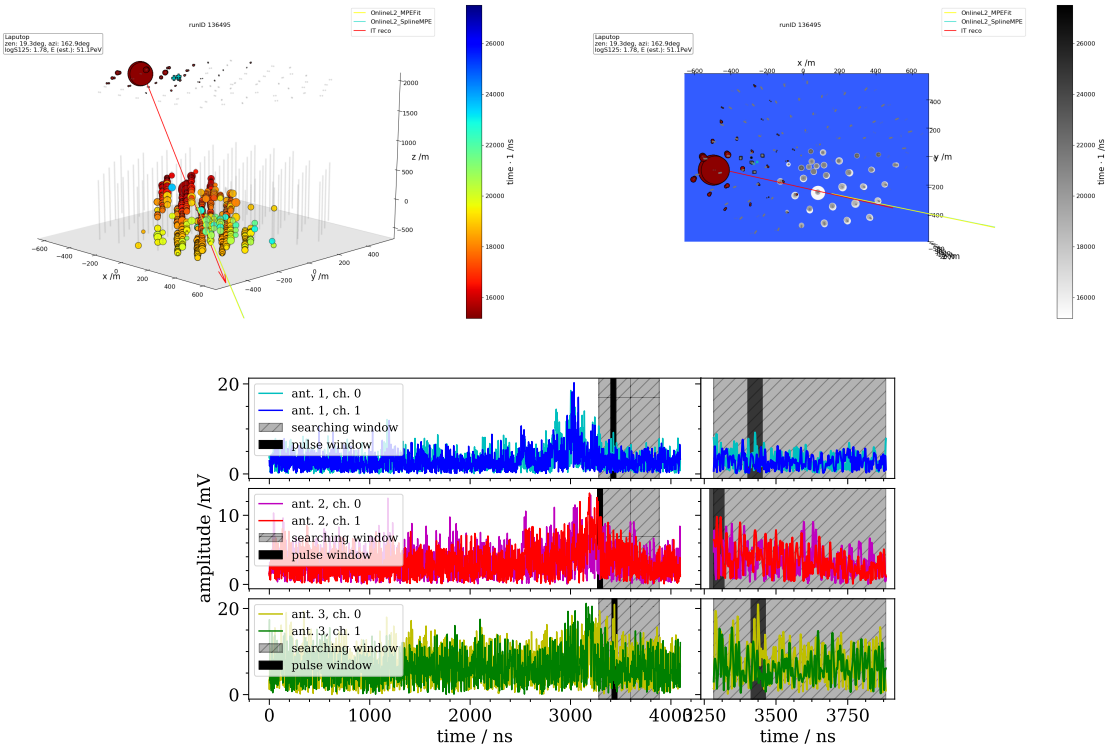


Figure F.4: runid 136495, eventid 79701847 (2022)

APPENDIX G

Production

G.1 The mount v1

Two units of mount v1 were created, and both were deployed with the prototype station in January 2020. These mounts demanded a substantial amount of manual labor and underwent several rounds of adjustments, carefully fine-tuned through trial and error at KIT, in order to be seamlessly deployed in the challenging terrain of the South Pole.

Figure G.1: One of the mount v1 at KIT.

G.2 The mount v2

The ordering of the wood and the production of all the connectors by the KIT workshop for 10 stations was finished in 2022. The pre-assembly of the parts included :

- square brace,
- painting antenna connector for UV resistance,
- antenna connector glued in antenna legs,
- antenna dowel screwed in the antenna connector with the angled connector screwed on the other side,
- threaded insert and rod installed in 1.20 m dowels,
- paddles of baseplate assembled together,
- baseplate connectors screws in the baseplate,

and was completed for 30 antennas. Three of them are installed at KIT since 2020. The antennas and cables are ordered for the complete deployment. Everything is shipping-ready on the hardware side of the antennas for at least the next 7 stations.

Figure G.2: Left: Pre-assembly of the Mount v2 at KIT. Right: Antenna array at KIT.

G.3 Production electronics

The 60 radioTad v2 were ordered and populated. All of them are checked for misfunctions and were assigned to a TAXIs board. The production and population of the TAXIs are also completed for seventeen boards. The calibration of the TAXI and the radio board is almost completed. A complete TAXI box with eight scintillation panels and three antennas was sent to the Auger observatory in Malargüe in the summer of 2022. Another complete station was sent to Telescope Array observatory in Utah. Its installation was planned to finish in autumn 2022.

APPENDIX H

Waveform artifacts remover (bin spikes)

H.1 Investigation on the binary behavior of "bin spikes"

Figure H.1 shows an example of several waveforms with multiple "bin spikes". The data was recorded with the TAXI v3.2 in the lab. The TAXI was heated in the climate chamber to 50ºC. The amplitude in each bin is associated with one of the four categories: baseline, completely down, mid-range down, and mid-range up. Each amplitude value in ADC counts is converted back to a binary number.

The classification of the amplitude in binary of each bin is shown in Figure H.2. The upper row represents the baseline amplitudes (well-behaved), the second row, the completely down (close to zero) amplitudes, the third row, the mid-range down, and the bottom row, the mid-range up. The number associated with each register of the 14-bits array is shown on the y-axis, where black is 0 and white is 1. On that axis, 0 corresponds to the least significant bit and 13 to the most significant bit.

For the well-behaved amplitudes (top row), the noise regime can be seen starting at 2^8 and thus the baseline regime stops a 2^9. This can be seen because of the higher fluctuation and lack of pattern in the noise regime. The baseline, being located about half the dynamic range, has always a pattern of either [10000XXXXXXXXX], or [01111XXXXXXXXX].

The second row represents the completely down bin spikes, and one can see that the bits 9-10-11-12-13 are all equal to zero ([00000XXXXXXXXX]), but the noise regime still starts at 2^8. This is concluded to be a bit flip of the most significant bit from 1 to 0.

For the second and third rows, the noise regime starts at 2^7, and the bits 8 to 12 look very similar to the baseline regime of the well-behaved amplitude. It is also interesting to note that in the waveform, the mid-range up is followed by a mid-range down and vice-versa, as shown in Figure H.3. All this together brings the hypothesis of a bit shift.

It is interesting to note that the bit flips are always from 1 to 0, completely high bin spikes ([11111XXXXXXXXX]) were not seen in the data so far. This might have to do with the preceding

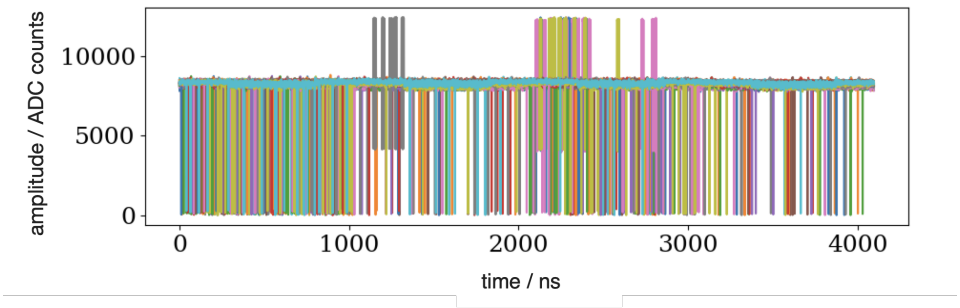


Figure H.1: Highly corrupted waveform from lab measurement with TAXI v3.2 at 50ºC.

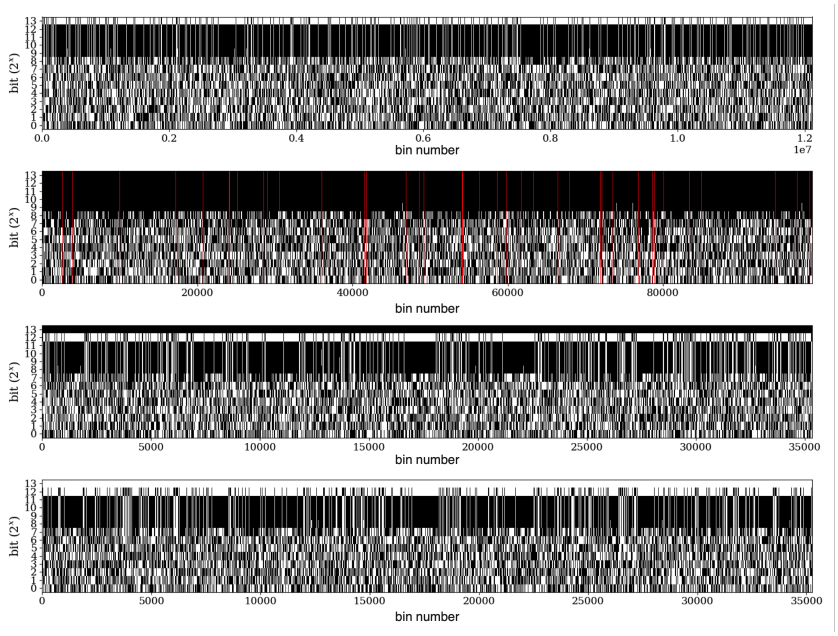


Figure H.2: Classification of the amplitude in a bin according to their values. Transformed into a binary number. Upper row: baseline amplitudes, second row: completely down, third row: mid-range down, and bottom row: mid-range up.

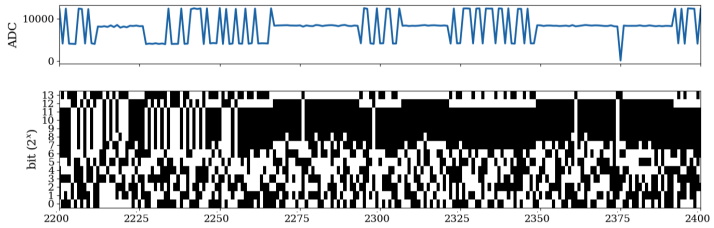


Figure H.3: Mid-ranges bin spikes.

packet in the file that always ends with 0.

H.2 Testing the hypothesis

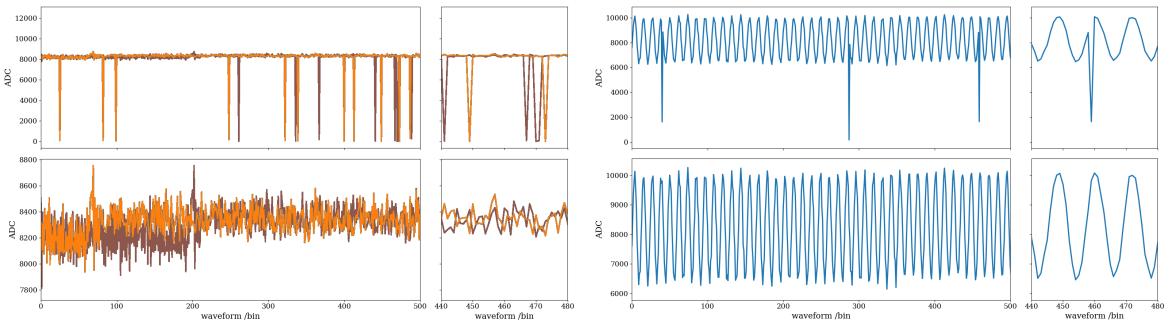


Figure H.4: Correction applied on two noise waveforms on the left, and applied on a sinusoidal signal on the right. In both cases, the expected shape of the signal is obtained after correction.

Those hypotheses were tested on a waveform with noise only, and with a waveform with a sinusoidal

signal, shown in Figure H.4 In the baseline waveform, one can see the disappearance of the bin spikes, but more convincing is the sinusoidal test. Where the sinusoidal shape is obtained after correction and no residual artifact is seen in the correction.

H.3 Implementation of the "binArtifactRemover" within radcube

From these results, a module was implemented within radcube, which corrects the bin spikes. The difficulty in correcting the bin spikes lies in correctly identifying them, and not, for example, mistakenly correcting an air shower pulse. The redundancy of the channels in the non-cascaded and semi-cascaded modes is exploited. The waveforms are compared bin-wise and if the difference between the same bin of the two (or four) channels is over, a user-defined, threshold, the bin is flagged and corrected. The default value is 2^{10}. For the fully cascaded mode, the identification of bin spikes is slightly more prone to errors. Each bin is compared to its closest neighbors and if the difference is again over a threshold, the bin is corrected. This works relatively well in the limit where the corruption is not too severe, otherwise, the flagging of faulty bins starts to break down. The results of the bin correction on many waveforms using the developed module are shown in Figure H.5.

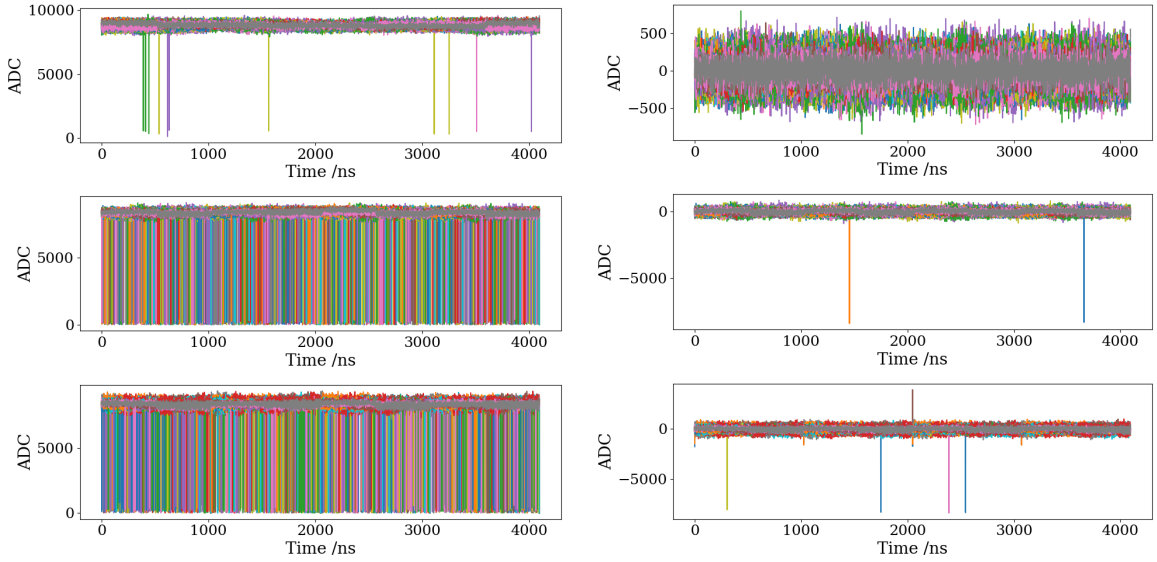


Figure H.5: Correction of a few waveforms with the "binArtifactRemover" module. On the right, the uncorrected noise waveforms from the laboratory measurement, and on the left, the same waveforms after correction.

Bibliography

[1] Aartsen, M. G. *et al.*, "Multimessenger observations of a flaring blazar coincident with high-energy neutrino IceCube-170922A," *Science*, vol. 361, no. 6398, eaat1378, 2018. DOI: 10.1126/science.aat1378. arXiv: 1807.08816 [astro-ph.HE] (cit. on p. 3).

[2] Stein, R. *et al.*, "A tidal disruption event coincident with a high-energy neutrino," *Nature Astron.*, vol. 5, no. 5, pp. 510–518, 2021. DOI: 10.1038/s41550-020-01295-8. arXiv: 2005.05340 [astro-ph.HE] (cit. on pp. 3, 9).

[3] Aartsen, M. G. *et al.*, "Neutrino emission from the direction of the blazar TXS 0506+056 prior to the IceCube-170922A alert," *Science*, vol. 361, no. 6398, pp. 147–151, 2018. DOI: 10.1126/science.aat2890. arXiv: 1807.08794 [astro-ph.HE] (cit. on p. 3).

[4] Coleman, A. *et al.*, "Ultra high energy cosmic rays The intersection of the Cosmic and Energy Frontiers," *Astropart. Phys.*, vol. 149, p. 102 819, 2023. DOI: 10.1016/j.astropartphys.2023.102819. arXiv: 2205.05845 [astro-ph.HE] (cit. on pp. 4, 7, 24).

[5] Recchia, S., Gabici, S., Aharonian, F., and Vink, J., "A local fading accelerator and the origin of TeV cosmic ray electrons," *PoS*, vol. ICRC2021, p. 168, 2021. DOI: 10.22323/1.395.0168 (cit. on p. 4).

[6] Adamo, M., Pietroni, S., and Spurio, M., "Astrophysical sources and acceleration mechanisms," *PoS*, vol. CORFU2021, p. 318, 2022. DOI: 10.22323/1.406.0318. arXiv: 2202.09170 [astro-ph.HE] (cit. on p. 5).

[7] Abbott, R. *et al.*, "GW190521: A Binary Black Hole Merger with a Total Mass of $150M_{\odot}$," *Phys. Rev. Lett.*, vol. 125, no. 10, p. 101 102, 2020. DOI: 10.1103/PhysRevLett.125.101102. arXiv: 2009.01075 [gr-qc] (cit. on p. 5).

[8] Abreu, P. *et al.*, "The energy spectrum of cosmic rays beyond the turn-down around 10^{17} eV as measured with the surface detector of the Pierre Auger Observatory," *Eur. Phys. J. C*, vol. 81, no. 11, p. 966, 2021. DOI: 10.1140/epjc/s10052-021-09700-w. arXiv: 2109.13400 [astro-ph.HE] (cit. on p. 5).

[9] Aab, A. *et al.*, "Features of the Energy Spectrum of Cosmic Rays above 2.5×10^{18} eV Using the Pierre Auger Observatory," *Phys. Rev. Lett.*, vol. 125, no. 12, p. 121 106, 2020. DOI: 10.1103/PhysRevLett.125.121106. arXiv: 2008.06488 [astro-ph.HE] (cit. on p. 5).

[10] Aab, A. *et al.*, "Measurement of the cosmic-ray energy spectrum above 2.5×10^{18} eV using the Pierre Auger Observatory," *Phys. Rev. D*, vol. 102, no. 6, p. 062 005, 2020. DOI: 10.1103/PhysRevD.102.062005. arXiv: 2008.06486 [astro-ph.HE] (cit. on p. 5).

[11] Hillas, A. M., "The Origin of Ultrahigh-Energy Cosmic Rays," *Ann. Rev. Astron. Astrophys.*, vol. 22, pp. 425–444, 1984. DOI: 10.1146/annurev.aa.22.090184.002233 (cit. on pp. 5, 8).

[12] Blasi, P., "Cosmic Ray Acceleration in Supernova Remnants," in *ICATPP Conference on Cosmic Rays for Particle and Astroparticle Physics*, 2011, pp. 493–506. DOI: 10.1142/9789814329033_0061. arXiv: 1012.5005 [astro-ph.HE] (cit. on p. 5).

[13] Aartsen, M. G. *et al.*, "Observation of Cosmic Ray Anisotropy with the IceTop Air Shower Array," *Astrophys. J.*, vol. 765, p. 55, 2013. DOI: 10.1088/0004-637X/765/1/55. arXiv: 1210.5278 [astro-ph.HE] (cit. on p. 6).

[14] Peters, B., "Primary cosmic radiation and extensive air showers," *Il Nuovo Cimento (1955-1965)*, vol. 22, no. 4, pp. 800–819, 1961 (cit. on p. 6).

[15] Apel, W. D. *et al.*, "Kneelike structure in the spectrum of the heavy component of cosmic rays observed with KASCADE-Grande," *Phys. Rev. Lett.*, vol. 107, p. 171 104, 2011. DOI: 10.1103/PhysRevLett.107.171104. arXiv: 1107.5885 [astro-ph.HE] (cit. on p. 6).

[16] Greisen, K., "End to the cosmic ray spectrum?" *Phys. Rev. Lett.*, vol. 16, pp. 748–750, 1966. DOI: 10.1103/PhysRevLett.16.748 (cit. on p. 6).

[17] Zatsepin, G. T. and Kuzmin, V. A., "Upper limit of the spectrum of cosmic rays," *JETP Lett.*, vol. 4, pp. 78–80, 1966 (cit. on p. 6).

[18] Aloisio, R., Berezinsky, V., and Gazizov, A., "Ultra High Energy Cosmic Rays: The disappointing model," *Astropart. Phys.*, vol. 34, pp. 620–626, 2011. DOI: 10.1016/j.astropartphys.2010.12.008. arXiv: 0907.5194 [astro-ph.HE] (cit. on p. 6).

[19] Aab, A. *et al.*, "The Pierre Auger Observatory Upgrade - Preliminary Design Report," Apr. 2016. arXiv: 1604.03637 [astro-ph.IM] (cit. on pp. 6, 7, 16).

[20] Alves Batista, R. *et al.*, "Open Questions in Cosmic-Ray Research at Ultrahigh Energies," *Front. Astron. Space Sci.*, vol. 6, p. 23, 2019. DOI: 10.3389/fspas.2019.00023. arXiv: 1903.06714 [astro-ph.HE] (cit. on pp. 7, 9).

[21] Schröder, F. G. *et al.*, "High-Energy Galactic Cosmic Rays (Astro2020 Science White Paper)," *Bull. Am. Astron. Soc.*, vol. 51, p. 131, Mar. 2019. arXiv: 1903.07713 [astro-ph.HE] (cit. on p. 8).

[22] Almeida, R. de *et al.*, "Large-scale and multipolar anisotropies of cosmic rays detected at the Pierre Auger Observatory with energies above 4 EeV," *PoS*, vol. ICRC2021, p. 335, 2021. DOI: 10.22323/1.395.0335 (cit. on p. 8).

[23] Biehl, D., Boncioli, D., Lunardini, C., and Winter, W., "Tidally disrupted stars as a possible origin of both cosmic rays and neutrinos at the highest energies," *Sci. Rep.*, vol. 8, no. 1, p. 10 828, 2018. DOI: 10.1038/s41598-018-29022-4. arXiv: 1711.03555 [astro-ph.HE] (cit. on p. 9).

[24] Aab, A. *et al.*, "An Indication of anisotropy in arrival directions of ultra-high-energy cosmic rays through comparison to the flux pattern of extragalactic gamma-ray sources," *Astrophys. J. Lett.*, vol. 853, no. 2, p. L29, 2018. DOI: 10.3847/2041-8213/aaa66d. arXiv: 1801.06160 [astro-ph.HE] (cit. on p. 9).

[25] Anchordoqui, L. A., "Acceleration of ultrahigh-energy cosmic rays in starburst superwinds," *Phys. Rev. D*, vol. 97, no. 6, p. 063 010, 2018. DOI: 10.1103/PhysRevD.97.063010. arXiv: 1801.07170 [astro-ph.HE] (cit. on p. 9).

[26] Hinton, J. A., "The Status of the H.E.S.S. project," *New Astron. Rev.*, vol. 48, pp. 331–337, 2004. DOI: 10.1016/j.newar.2003.12.004. arXiv: astro-ph/0403052 (cit. on p. 10).

[27] Aleksic, J. *et al.*, "Performance of the MAGIC stereo system obtained with Crab Nebula data," *Astropart. Phys.*, vol. 35, pp. 435–448, 2012. DOI: 10.1016/j.astropartphys.2011.11.007. arXiv: 1108.1477 [astro-ph.IM] (cit. on p. 10).

[28] Smith, A. J., "HAWC: Design, Operation, Reconstruction and Analysis," *PoS*, vol. ICRC2015, p. 966, 2016. DOI: 10.22323/1.236.0966. arXiv: 1508.05826 [astro-ph.IM] (cit. on p. 10).

[29] Cao, Z., "A future project at Tibet: The large high altitude air shower observatory (LHAASO)," *Chin. Phys. C*, vol. 34, pp. 249–252, 2010. DOI: 10.1088/1674-1137/34/2/018 (cit. on p. 10).

[30] Gould, R. and Schréder, G., "Opacity of the Universe to High-Energy Photons," *Phys. Rev. Lett.*, vol. 16, no. 6, pp. 252–254, 1966. DOI: 10.1103/PhysRevLett.16.252 (cit. on p. 10).

[31] Desiati, P., "IceCube: Toward a km^3 Neutrino Telescope," in *20th European Cosmic Ray Symposium*, Nov. 2006, S5–73. arXiv: astro-ph/0611603 (cit. on p. 10).

[32] Cao, Z., Aharonian, F., An, Q., *et al.*, "Ultrahigh-energy photons up to 1.4 petaelectronvolts from 12 γ-ray galactic sources," *Nature*, vol. 594, no. 7861, pp. 33–36, 2021. DOI: 10.1038/s41586-021-03498-z (cit. on p. 10).

[33] Katz, U. F. and Spiering, C., "High-Energy Neutrino Astrophysics: Status and Perspectives," *Prog. Part. Nucl. Phys.*, vol. 67, pp. 651–704, 2012. DOI: 10.1016/j.ppnp.2011.12.001. arXiv: 1111.0507 [astro-ph.HE] (cit. on p. 11).

[34] Aartsen, M. G. *et al.*, "First observation of PeV-energy neutrinos with IceCube," *Phys. Rev. Lett.*, vol. 111, p. 021 103, 2013. DOI: 10.1103/PhysRevLett.111.021103. arXiv: 1304.5356 [astro-ph.HE] (cit. on p. 11).

[35] Aartsen, M. G. *et al.*, "Detection of a particle shower at the Glashow resonance with IceCube," *Nature*, vol. 591, no. 7849, pp. 220–224, 2021, [Erratum: Nature 592, E11 (2021)]. DOI: 10.1038/s41586-021-03256-1. arXiv: 2110.15051 [hep-ex] (cit. on p. 11).

[36] Aartsen, M. G. *et al.*, "Constraints on Ultrahigh-Energy Cosmic-Ray Sources from a Search for Neutrinos above 10 PeV with IceCube," *Phys. Rev. Lett.*, vol. 117, no. 24, p. 241 101, 2016, [Erratum: Phys.Rev.Lett. 119, 259902 (2017)]. DOI: 10.1103/PhysRevLett.117.241101. arXiv: 1607.05886 [astro-ph.HE] (cit. on p. 11).

[37] Abbott, B. P. *et al.*, "Observation of Gravitational Waves from a Binary Black Hole Merger," *Phys. Rev. Lett.*, vol. 116, no. 6, p. 061 102, 2016. DOI: 10.1103/PhysRevLett.116.061102. arXiv: 1602.03837 [gr-qc] (cit. on p. 12).

[38] Abbott, B. P. *et al.*, "Multi-messenger Observations of a Binary Neutron Star Merger," *Astrophys. J. Lett.*, vol. 848, no. 2, p. L12, 2017. DOI: 10.3847/2041-8213/aa91c9. arXiv: 1710.05833 [astro-ph.HE] (cit. on p. 12).

[39] Auger, P., Ehrenfest, P., Maze, R., Daudin, J., and Robley, A. F., "Extensive cosmic ray showers," *Rev. Mod. Phys.*, vol. 11, pp. 288–291, 1939. DOI: 10.1103/RevModPhys.11.288 (cit. on p. 13).

[40] Aab, A. *et al.*, "The Pierre Auger Cosmic Ray Observatory," *Nucl. Instrum. Meth. A*, vol. 798, pp. 172–213, 2015. DOI: 10.1016/j.nima.2015.06.058. arXiv: 1502.01323 [astro-ph.IM] (cit. on p. 13).

[41] Abu-Zayyad, T. *et al.*, "The surface detector array of the Telescope Array experiment," *Nucl. Instrum. Meth. A*, vol. 689, pp. 87–97, 2013. DOI: 10.1016/j.nima.2012.05.079. arXiv: 1201.4964 [astro-ph.IM] (cit. on p. 13).

[42] Aartsen, M. G. *et al.*, "The IceCube Neutrino Observatory: Instrumentation and Online Systems," *JINST*, vol. 12, no. 03, P03012, 2017. DOI: 10.1088/1748-0221/12/03/P03012. arXiv: 1612.05093 [astro-ph.IM] (cit. on pp. 13, 27, 30).

[43] Zyla, P. A. *et al.*, "Review of Particle Physics," *PTEP*, vol. 2020, no. 8, p. 083C01, 2020. DOI: 10.1093/ptep/ptaa104 (cit. on pp. 13–15).

[44] Heitler, W., *The quantum theory of radiation*, ser. International Series of Monographs on Physics. Oxford: Oxford University Press, 1936, vol. 5 (cit. on p. 14).

[45] Gaisser, T. K., Engel, R., and Resconi, E., *Cosmic Rays and Particle Physics: 2nd Edition*. Cambridge University Press, Jun. 2016, ISBN: 978-0-521-01646-9 (cit. on p. 14).

[46] Engel, R., Heck, D., and Pierog, T., "Extensive air showers and hadronic interactions at high energy," *Ann. Rev. Nucl. Part. Sci.*, vol. 61, pp. 467–489, 2011. DOI: 10.1146/annurev.nucl.012809.104544 (cit. on p. 15).

[47] Matthews, J., "A Heitler model of extensive air showers," *Astropart. Phys.*, vol. 22, pp. 387–397, 2005. DOI: 10.1016/j.astropartphys.2004.09.003 (cit. on p. 15).

[48] Aab, A. *et al.*, "Extraction of the muon signals recorded with the surface detector of the Pierre Auger Observatory using recurrent neural networks," *JINST*, vol. 16, no. 07, P07016, 2021. DOI: 10.1088/1748-0221/16/07/P07016. arXiv: 2103.11983 [hep-ex] (cit. on p. 16).

[49] Abbasi, R. *et al.*, "Discrimination of Muons for Mass Composition Studies of Inclined Air Showers Detected with IceTop," *PoS*, vol. ICRC2021, p. 212, 2021. DOI: 10.22323/1.395.0212. arXiv: 2107.11293 [astro-ph.HE] (cit. on p. 16).

[50] Albrecht, J. *et al.*, "The Muon Puzzle in cosmic-ray induced air showers and its connection to the Large Hadron Collider," *Astrophys. Space Sci.*, vol. 367, no. 3, p. 27, 2022. DOI: 10.1007/s10509-022-04054-5. arXiv: 2105.06148 [astro-ph.HE] (cit. on p. 16).

[51] Feng, J. L. *et al.*, "The Forward Physics Facility at the High-Luminosity LHC," *J. Phys. G*, vol. 50, no. 3, p. 030 501, 2023. DOI: 10.1088/1361-6471/ac865e. arXiv: 2203.05090 [hep-ex] (cit. on p. 16).

[52] Huege, T., "Radio detection of cosmic ray air showers in the digital era," *Phys. Rept.*, vol. 620, pp. 1–52, 2016. DOI: 10.1016/j.physrep.2016.02.001. arXiv: 1601.07426 [astro-ph.IM] (cit. on p. 16).

[53] Schröder, F. G., "Radio detection of Cosmic-Ray Air Showers and High-Energy Neutrinos," *Prog. Part. Nucl. Phys.*, vol. 93, pp. 1–68, 2017. DOI: 10.1016/j.ppnp.2016.12.002. arXiv: 1607.08781 [astro-ph.IM] (cit. on p. 16).

[54] "P series radiowave propagation," in *Recommendation ITU-R P.372-15*, Union, I. T., Ed., Electronic Publication, 2021, pp. 0372–03 (cit. on pp. 16, 23).

[55] Falcke, H. *et al.*, "Detection and imaging of atmospheric radio flashes from cosmic ray air showers," *Nature*, vol. 435, pp. 313–316, 2005. DOI: 10.1038/nature03614. arXiv: astro-ph/0505383 (cit. on p. 16).

[56] Bezyazeekov, P. A. *et al.*, "Measurement of cosmic-ray air showers with the Tunka Radio Extension (Tunka-Rex)," *Nucl. Instrum. Meth. A*, vol. 802, pp. 89–96, 2015. DOI: 10.1016/j.nima.2015.08.061. arXiv: 1509.08624 [astro-ph.IM] (cit. on p. 16).

[57] Haarlem, M. P. van *et al.*, "LOFAR: The LOw-Frequency ARray," *Astron. Astrophys.*, vol. 556, A2, 2013. DOI: 10.1051/0004-6361/201220873. arXiv: 1305.3550 [astro-ph.IM] (cit. on p. 16).

[58] Abreu, P. *et al.*, "Antennas for the Detection of Radio Emission Pulses from Cosmic-Ray," *JINST*, vol. 7, P10011, 2012. DOI: 10.1088/1748-0221/7/10/P10011. arXiv: 1209.3840 [astro-ph.IM] (cit. on p. 16).

[59] Carilli, C. L. and Rawlings, S., "Science with the Square Kilometer Array: Motivation, key science projects, standards and assumptions," *New Astron. Rev.*, vol. 48, p. 979, 2004. DOI: 10.1016/j.newar.2004.09.001. arXiv: astro-ph/0409274 (cit. on pp. 16, 35).

[60] Schröder, F. G., "Status of the radio technique for cosmic-ray induced air showers," *Nucl. Part. Phys. Proc.*, vol. 279-281, pp. 190–197, 2016. DOI: 10.1016/j.nuclphysbps.2016.10.027 (cit. on p. 17).

[61] Allan, H., Neat, K., and Jones, J., "Mechanism of radio emission from extensive air showers," *Nature*, vol. 215, no. 5098, pp. 267–268, 1967 (cit. on p. 16).

[62] Askaryan, G. A., "Excess negative charge of an electron-photon shower and its coherent radio emission," *Zh. Eksp. Teor. Fiz.*, vol. 41, pp. 616–618, 1961 (cit. on p. 17).

[63] Askaryan, G. A., "Coherent radio-emission from cosmic ray showers in the air and in dense media," 1965 (cit. on p. 17).

[64] Alvarez-Muniz, J., Romero-Wolf, A., and Zas, E., "Practical and accurate calculations of Askaryan radiation," *Phys. Rev. D*, vol. 84, p. 103 003, 2011. DOI: 10.1103/PhysRevD.84.103003. arXiv: 1106.6283 [astro-ph.HE] (cit. on p. 17).

[65] Paudel, E. N., Coleman, A., and Schroeder, F. G., "Simulation study of the relative Askaryan fraction at the south pole," *Phys. Rev. D*, vol. 105, no. 10, p. 103 006, 2022. DOI: 10.1103/PhysRevD.105.103006. arXiv: 2201.03405 [hep-th] (cit. on p. 18).

[66] Glaser, C., Erdmann, M., Hörandel, J. R., Huege, T., and Schulz, J., "Simulation of Radiation Energy Release in Air Showers," *JCAP*, vol. 09, p. 024, 2016. DOI: 10.1088/1475-7516/2016/09/024. arXiv: 1606.01641 [astro-ph.HE] (cit. on pp. 18–20, 106).

[67] James, C. W., "Nature of radio-wave radiation from particle cascades," *Phys. Rev. D*, vol. 105, no. 2, p. 023 014, 2022. DOI: 10.1103/PhysRevD.105.023014. arXiv: 2201.01298 [astro-ph.HE] (cit. on p. 18).

[68] Lenz, E., "Ueber die bestimmung der richtung der durch elektrodynamische vertheilung erregten galvanischen ströme," *Annalen der Physik*, vol. 107, no. 31, pp. 483–494, 1834 (cit. on p. 18).

[69] Alvarez-Muniz, J., Carvalho Jr., W. R., and Zas, E., "Monte Carlo simulations of radio pulses in atmospheric showers using ZHAireS," *Astropart. Phys.*, vol. 35, pp. 325–341, 2012. DOI: 10.1016/j.astropartphys.2011.10.005. arXiv: 1107.1189 [astro-ph.HE] (cit. on p. 18).

[70] Vries, K. D. de, Berg, A. M. van den, Scholten, O., and Werner, K., "Coherent Cherenkov Radiation from Cosmic-Ray-Induced Air Showers," *Phys. Rev. Lett.*, vol. 107, p. 061 101, 2011. DOI: 10.1103/PhysRevLett.107.061101. arXiv: 1107.0665 [astro-ph.HE] (cit. on p. 18).

[71] Vries, K. D. de, Scholten, O., and Werner, K., "The air shower maximum probed by Cherenkov effects from radio emission," *Astropart. Phys.*, vol. 45, pp. 23–27, 2013. DOI: 10.1016/j.astropartphys.2013.02.003. arXiv: 1304.1321 [astro-ph.HE] (cit. on p. 18).

[72] Alvarez-Muniz, J., Carvalho, W. R., Romero-Wolf, A., Tueros, M., and Zas, E., "Coherent Radiation from Extensive Air Showers in the Ultra-High Frequency Band," *Phys. Rev. D*, vol. 86, p. 123 007, 2012. DOI: 10.1103/PhysRevD.86.123007. arXiv: 1208.0951 [astro-ph.HE] (cit. on pp. 18, 19).

[73] Balagopal V., A., Haungs, A., Huege, T., and Schroeder, F. G., "Search for PeVatrons at the Galactic Center using a radio air-shower array at the South Pole," *Eur. Phys. J. C*, vol. 78, no. 2, p. 111, 2018, [Erratum: Eur.Phys.J.C 78, 1017 (2018), Erratum: Eur.Phys.J.C 81, 483 (2021)]. DOI: 10.1140/epjc/s10052-018-5537-2. arXiv: 1712.09042 [astro-ph.IM] (cit. on pp. 19, 49).

[74] Mulrey, K. *et al.*, "The energy scale of cosmic rays detected with LOFAR," *PoS*, vol. ICRC2019, p. 362, 2020. DOI: 10.22323/1.358.0362 (cit. on p. 20).

[75] Gottowik, M., "Measurements of Inclined Air Showers with the Auger Engineering Radio Array at the Pierre Auger Observatory," *PoS*, vol. ICRC2019, p. 274, 2021. DOI: 10.22323/1.358.0274 (cit. on pp. 20, 126).

[76] Schlüter, F., Gottowik, M., Huege, T., and Rautenberg, J., "Refractive displacement of the radio-emission footprint of inclined air showers simulated with CoREAS," *Eur. Phys. J. C*, vol. 80, no. 7, p. 643, 2020. DOI: 10.1140/epjc/s10052-020-8216-z. arXiv: 2005.06775 [astro-ph.IM] (cit. on pp. 20, 126).

[77] Briechle, F. L., "Measurement of the radiation energy release using radio emission from extensive air showers at the Pierre Auger Observatory," Ph.D. dissertation, RWTH Aachen U., 2021. DOI: 10.18154/RWTH-2022-01209 (cit. on p. 20).

[78] Glaser, C., Jong, S. de, Erdmann, M., and Hörandel, J. R., "An analytic description of the radio emission of air showers based on its emission mechanisms," *Astropart. Phys.*, vol. 104, pp. 64–77, 2019. DOI: 10.1016/j.astropartphys.2018.08.004. arXiv: 1806.03620 [astro-ph.HE] (cit. on p. 22).

[79] Huege, T., Brenk, L., and Schlüter, F., "A Rotationally Symmetric Lateral Distribution Function for Radio Emission from Inclined Air Showers," *EPJ Web Conf.*, vol. 216, Riccobene, G., Biagi, S., Capone, A., Distefano, C., and Piattelli, P., Eds., p. 03 009, 2019. DOI: 10.1051/epjconf/201921603009. arXiv: 1808.00729 [astro-ph.IM] (cit. on p. 22).

[80] Huege, T., Schlueter, F., and Brenk, L., "Symmetrizing the signal distribution of radio emission from inclined air showers," *PoS*, vol. ICRC2019, p. 294, 2021. DOI: 10.22323/1.358.0294. arXiv: 1908.07840 [astro-ph.IM] (cit. on p. 22).

[81] Apel, W. D. *et al.*, "Lateral Distribution of the Radio Signal in Extensive Air Showers Measured with LOPES," *Astropart. Phys.*, vol. 32, pp. 294–303, 2010. DOI: 10.1016/j.astropartphys.2009.09.007. arXiv: 0910.4866 [astro-ph.HE] (cit. on p. 22).

[82] Kostunin, D., Bezyazeekov, P. A., Hiller, R., Schröder, F. G., Lenok, V., and Levinson, E., "Reconstruction of air-shower parameters for large-scale radio detectors using the lateral distribution," *Astropart. Phys.*, vol. 74, pp. 79–86, 2016. DOI: 10.1016/j.astropartphys.2015.10.004. arXiv: 1504.05083 [astro-ph.HE] (cit. on p. 22).

[83] Schlüter, F. and Huege, T., "Signal model and event reconstruction for the radio detection of inclined air showers," Mar. 2022. arXiv: 2203.04364 [astro-ph.HE] (cit. on p. 22).

[84] Aab, A. *et al.*, "Energy Estimation of Cosmic Rays with the Engineering Radio Array of the Pierre Auger Observatory," *Phys. Rev. D*, vol. 93, no. 12, p. 122 005, 2016. DOI: 10.1103/PhysRevD.93.122005. arXiv: 1508.04267 [astro-ph.HE] (cit. on p. 22).

[85] Buitink, S. *et al.*, "Method for high precision reconstruction of air shower X_{max} using two-dimensional radio intensity profiles," *Phys. Rev. D*, vol. 90, no. 8, p. 082 003, 2014. DOI: 10.1103/PhysRevD.90.082003. arXiv: 1408.7001 [astro-ph.IM] (cit. on pp. 22, 107, 125).

[86] Krause, R., "A new method to determine the energy scale for high-energy cosmic rays using radio measurements at the Pierre Auger Observatory," *PoS*, vol. ICRC2017, p. 528, 2018. DOI: 10.22323/1.301.0528 (cit. on p. 23).

[87] Aartsen, M. G. *et al.*, "Design and Performance of the first IceAct Demonstrator at the South Pole," *JINST*, vol. 15, no. 02, T02002, 2020. DOI: 10.1088/1748-0221/15/02/T02002. arXiv: 1910.06945 [astro-ph.IM] (cit. on p. 24).

[88] Mulrey, K. *et al.*, "On the cosmic-ray energy scale of the LOFAR radio telescope," *JCAP*, vol. 11, p. 017, 2020. DOI: 10.1088/1475-7516/2020/11/017. arXiv: 2005.13441 [astro-ph.HE] (cit. on pp. 25, 107).

[89] Abbasi, R. *et al.*, "The Design and Performance of IceCube DeepCore," *Astropart. Phys.*, vol. 35, pp. 615–624, 2012. DOI: 10.1016/j.astropartphys.2012.01.004. arXiv: 1109.6096 [astro-ph.IM] (cit. on p. 28).

[90] Ishihara, A., "The IceCube Upgrade - Design and Science Goals," *PoS*, vol. ICRC2019, p. 1031, 2021. DOI: 10.22323/1.358.1031. arXiv: 1908.09441 [astro-ph.HE] (cit. on p. 28).

[91] Abbasi, R. *et al.*, "Measurement of Astrophysical Tau Neutrinos in IceCube's High-Energy Starting Events," Nov. 2020. arXiv: 2011.03561 [hep-ex] (cit. on p. 28).

[92] Aartsen, M. G. *et al.*, "IceCube-Gen2: the window to the extreme Universe," *J. Phys. G*, vol. 48, no. 6, p. 060 501, 2021. DOI: 10.1088/1361-6471/abbd48. arXiv: 2008.04323 [astro-ph.HE] (cit. on pp. 28, 29, 31).

[93] Abbasi, R. *et al.*, "IceTop: The surface component of IceCube," *Nucl. Instrum. Meth. A*, vol. 700, pp. 188–220, 2013. DOI: 10.1016/j.nima.2012.10.067. arXiv: 1207.6326 [astro-ph.IM] (cit. on pp. 28–30).

[94] Amin, N. M. B., "Implementation of IceTop data in the IceCube Realtime Alert System," *J. Phys. Conf. Ser.*, vol. 2156, p. 012 217, 2021. DOI: 10.1088/1742-6596/2156/1/012217 (cit. on p. 29).

[95] Aartsen, M. G. *et al.*, "Cosmic ray spectrum from 250 TeV to 10 PeV using IceTop," *Phys. Rev. D*, vol. 102, p. 122 001, 2020. DOI: 10.1103/PhysRevD.102.122001. arXiv: 2006.05215 [astro-ph.HE] (cit. on p. 30).

[96] Klepser, S., "Reconstruction of extensive air showers and measurement of the cosmic ray energy spectrum in the range of 1 - 80 pev at the south pole," Ph.D. dissertation, Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät I, 2008. DOI: http://dx.doi.org/10.18452/15790 (cit. on p. 30).

[97] Aartsen, M. G. *et al.*, "Measurement of the cosmic ray energy spectrum with IceTop-73," *Phys. Rev. D*, vol. 88, no. 4, p. 042 004, 2013. DOI: 10.1103/PhysRevD.88.042004. arXiv: 1307.3795 [astro-ph.HE] (cit. on p. 30).

[98] DeYoung, T., "IceTray: A software framework for IceCube," in *14th International Conference on Computing in High-Energy and Nuclear Physics*, 2005, pp. 463–466 (cit. on pp. 30, 55, 76, 92).

[99] Aguilar, J. A. *et al.*, "Design and Sensitivity of the Radio Neutrino Observatory in Greenland (RNO-G)," *JINST*, vol. 16, no. 03, P03025, 2021. DOI: 10.1088/1748-0221/16/03/P03025. arXiv: 2010.12279 [astro-ph.IM] (cit. on p. 31).

[100] Leszczyńska, A. S., "Potential of the IceTop Enhancement with a Scintillation Detector Array," Ph.D. dissertation, KIT, Karlsruhe, 2021. DOI: 10.5445/IR/1000131245 (cit. on p. 32).

[101] Huber, T., "IceScint : A Scintillation Detector Array for the IceCube IceTop Enhancement," Ph.D. dissertation, KIT, Karlsruhe, 2021. DOI: 10.5445/IR/1000131545 (cit. on p. 33).

[102] Renschler, M., "A Prototype Radio Detector for the IceCube Surface Enhancement," Ph.D. dissertation, KIT, Karlsruhe, 2020. DOI: 10.5445/IR/1000104529 (cit. on pp. 33, 37, 38, 63, 68, 133).

[103] Frank. G. Schroeder, *Radio Background between 25 and 500 MHz at IceTop at the South Pole*, Internal Report, Feb. 2020 (cit. on pp. 33, 52, 83).

[104] Lera Acedo, E. de, Razavi-Ghods, N., Troop, N., Drought, N., and Faulkner, A. J., "Skala, a log-periodic array antenna for the ska-low instrument: Design, simulations, tests and system considerations," *Experimental Astronomy*, vol. 39, no. 3, pp. 567–594, Jul. 2015, ISSN: 1572-9508. DOI: 10.1007/s10686-015-9439-0 (cit. on p. 35).

[105] Kauer, M., Huber, T., Tosi, D., and Wendt, C., "The Scintillator Upgrade of IceTop: Performance of the prototype array," *PoS*, vol. ICRC2019, p. 309, 2021. DOI: 10.22323/1.358.0309. arXiv: 1908.09860 [astro-ph.HE] (cit. on p. 35).

[106] Aswathi, B. V., Haungs, A., Huege, T., and Schröder, F. G., "PeVatron Search Using Radio Measurements of Extensive Air Showers at the South Pole," *J. Phys. Conf. Ser.*, vol. 1342, no. 1, Clark, K., Jillings, C., Kraus, C., Saffin, J., and Scorza, S., Eds., p. 012 006, 2020. DOI: 10.1088/1742-6596/1342/1/012006 (cit. on p. 35).

[107] United States Antarctic Program, *Field manual - comms*, https://www.usap.gov/USAPgov/travelAndDeployment/documents/FieldManual-Comms.pdf (cit. on pp. 35, 49).

[108] Carvill, J., *Mechanical engineer's data handbook*. Butterworth-Heinemann, 1994 (cit. on p. 38).

[109] Hansson, S. O., "Risk and safety in technology," in *Philosophy of Technology and Engineering Sciences*, ser. Handbook of the Philosophy of Science, Meijers, A., Ed., Amsterdam: North-Holland, 2009, pp. 1069–1102. DOI: https://doi.org/10.1016/B978-0-444-51667-1.50043-4 (cit. on p. 41).

[110] Knoll, F., "Commentary on the basic philosophy and recent development of safety margins," *Canadian Journal of Civil Engineering*, vol. 3, no. 3, pp. 409–416, 1976 (cit. on p. 41).

[111] Moses, F., "Problems and prospects of reliability-based optimization," *Engineering Structures*, vol. 19, no. 4, pp. 293–301, 1997 (cit. on p. 41).

[112] *Laminate vetronit egs 102*, WH QGELA, Von Roll Deutschland GmbH, Apr. 2008 (cit. on p. 43).

[113] *Ep21nd technical data sheet*, H20C15, Master Bond (cit. on p. 43).

[114] Oehler, M., "The prototype station for the IceCube surface array enhancement," Ph.D. dissertation, KIT, Karlsruhe, 2022. DOI: 10.5445/IR/1000142813 (cit. on pp. 49, 83, 88).

[115] Schlüter, F. and Huege, T., "Expected performance of air-shower measurements with the radio-interferometric technique," *JINST*, vol. 16, no. 07, P07048, 2021. DOI: 10.1088/1748-0221/16/07/P07048. arXiv: 2102.13577 [astro-ph.IM] (cit. on p. 49).

[116] Schoorlemmer, H. and Carvalho, W. R., "Radio interferometry applied to the observation of cosmic-ray induced extensive air showers," *Eur. Phys. J. C*, vol. 81, no. 12, p. 1120, 2021. DOI: 10.1140/epjc/s10052-021-09925-9. arXiv: 2006.10348 [astro-ph.HE] (cit. on p. 49).

[117] Karg, T., Haungs, A., Kleifges, M., Nahnhauer, R., and Sulanke, K. H., "Introducing taxi: A transportable array for extremely large area instrumentation studies," 2014. arXiv: 1410.4685 [astro-ph.IM] (cit. on p. 49).

[118] Jansweijer, P. P. M., Peek, H. Z., and De Wolf, E., "White Rabbit: Sub-nanosecond timing over Ethernet," *Nucl. Instrum. Meth. A*, vol. 725, Anton, G. *et al.*, Eds., pp. 187–190, 2013. DOI: 10.1016/j.nima.2012.12.096 (cit. on p. 50).

[119] Abbasi, R. *et al.*, "Development of a scintillation and radio hybrid detector array at the South Pole," *PoS*, vol. ICRC2021, p. 225, 2021. DOI: 10.22323/1.395.0225. arXiv: 2107.09983 [astro-ph.HE] (cit. on pp. 50, 83, 86).

[120] PSI, *Drs4*, https://www.psi.ch/sites/default/files/2020-08/DRS4_rev09_2.pdf, Accessed: 2018-12-06, 2015 (cit. on pp. 50, 54).

[121] Abbasi, R. *et al.*, "Simulation Study of the Observed Radio Emission of Air Showers by the IceTop Surface Extension," *PoS*, vol. ICRC2021, p. 317, 2021. DOI: 10.22323/1.395.0317. arXiv: 2107.09666 [astro-ph.IM] (cit. on p. 52).

[122] Harris, C. R., Millman, K. J., Walt, S. J. van der, *et al.*, "Array programming with NumPy," *Nature*, vol. 585, no. 7825, pp. 357–362, Sep. 2020. DOI: 10.1038/s41586-020-2649-2 (cit. on pp. 55, 99).

[123] Abbasi, R. *et al.*, "First air-shower measurements with the prototype station of the IceCube surface enhancement," *PoS*, vol. ICRC2021, p. 314, 2021. DOI: 10.22323/1.395.0314. arXiv: 2107.08750 [astro-ph.HE] (cit. on pp. 56, 86, 88).

[124] de Lera Acedo, E., Drought, N., Wakley, B., and Faulkner, A., "Evolution of skala (skala-2), the log-periodic array antenna for the ska-low instrument," *2015 International Conference on Electromagnetics in Advanced Applications (ICEAA)*, pp. 839–843, Sep. 2015. DOI: 10.1109/ICEAA.2015.7297231 (cit. on pp. 57, 94).

[125] Renschler, M., "First measurements with prototype radio antennas for the IceTop detector array," *PoS*, vol. ICRC2019, p. 401, 2021. DOI: 10.22323/1.358.0401. arXiv: 1908.10565 [astro-ph.IM] (cit. on p. 58).

[126] *I-tera® mt rf very low-loss laminate material*, 10/14 DSI-TeraMTRFC, Isola group, 2014 (cit. on p. 60).

[127] SYSTEMS, T. M., *Lmr®-240 flexible low loss communications coax*, https://www.koax24.de/storage/datasheet/de/050152_Datenblatt_LMR-240_LMR-240.pdf (cit. on p. 60).

[128] SYSTEMS, T. M., *Lmr®-400 flexible low loss communications coax*, https://www.koax24.de/storage/datasheet/de/050152_Datenblatt_LMR-240_LMR-240.pdf (cit. on p. 60).

[129] *Ths4508 - wideband, low noise, low distortion fully differential amplifier*, SLA459E, Rev. Sept. 2008, Texas Instruments, Sep. 2005 (cit. on pp. 64, 73).

[130] Rech, D., "Characterization of the data acquisition system for the radio antennas of the icecube surface array enhancement," B.S. Thesis, Karlsruhe Institute of Technology, Jul. 2021 (cit. on pp. 64, 65, 77, 84).

[131] Steinmüller, P., "Development of radio front-end electronics for a prototype radio station at the south pole," M.S. thesis, Karlsruhe Institute of Technology, 2019 (cit. on p. 68).

[132] Nelles, A. *et al.*, "Calibrating the absolute amplitude scale for air showers measured at LOFAR," *JINST*, vol. 10, no. 11, P11005, 2015. DOI: 10.1088/1748-0221/10/11/P11005. arXiv: 1507.08932 [astro-ph.IM] (cit. on p. 73).

[133] Briechle, F., "In-situ absolute calibration of electric-field amplitude measurements with the LPDA radio detector stations of the Pierre Auger Observatory," *EPJ Web Conf.*, vol. 135, Buitnik, S., Hörandel, J. R., Jong, S. de, Lahmann, R., Nahnhauer, R., and Scholten, O., Eds., p. 01 014, 2017. DOI: 10.1051/epjconf/201713501014. arXiv: 1609.01511 [astro-ph.IM] (cit. on p. 73).

[134] Aab, A. *et al.*, "Calibration of the logarithmic-periodic dipole antenna (LPDA) radio stations at the Pierre Auger Observatory using an octocopter," *JINST*, vol. 12, no. 10, T10005, 2017. DOI: 10.1088/1748-0221/12/10/T10005. arXiv: 1702.01392 [astro-ph.IM] (cit. on p. 73).

[135] Oliveira-Costa, A. de, Tegmark, M., Gaensler, B. M., Jonas, J., Landecker, T. L., and Reich, P., "A model of diffuse Galactic Radio Emission from 10 MHz to 100 GHz," *Mon. Not. Roy. Astron. Soc.*, vol. 388, p. 247, 2008. DOI: 10.1111/j.1365-2966.2008.13376.x. arXiv: 0802.1525 [astro-ph] (cit. on p. 80).

[136] Abbasi, R. *et al.*, "Framework and tools for the simulation and analysis of the radio emission from air showers at IceCube," *JINST*, vol. 17, no. 06, P06026, 2022. DOI: 10.1088/1748-0221/17/06/P06026. arXiv: 2205.02258 [astro-ph.HE] (cit. on pp. 82, 87, 91, 92, 112).

[137] Cane, H., "Spectra of the non-thermal radio radiation from the galactic polar regions," *Monthly Notices of the Royal Astronomical Society*, vol. 189, pp. 465–478, Dec. 1979. DOI: 10.1093/mnras/189.3.465 (cit. on pp. 82, 94).

[138] NOAA, *Esrl global monitoring laboratory*, https://gml.noaa.gov/dv/data/index.php, Aug. 2022 (cit. on p. 84).

[139] Heck, D., Knapp, J., Capdevielle, J. N., Schatz, G., and Thouw, T., "CORSIKA: A Monte Carlo code to simulate extensive air showers," Feb. 1998 (cit. on pp. 91, 112).

[140] Huege, T., Ludwig, M., and James, C. W., "Simulating radio emission from air showers with CoREAS," *AIP Conf. Proc.*, vol. 1535, no. 1, Lahmann, R., Eberl, T., Graf, K., *et al.*, Eds., p. 128, 2013. DOI: 10.1063/1.4807534. arXiv: 1301.2132 [astro-ph.HE] (cit. on pp. 91, 112).

[141] Riehn, F., Engel, R., Fedynitch, A., Gaisser, T. K., and Stanev, T., "Hadronic interaction model Sibyll 2.3d and extensive air showers," *Phys. Rev. D*, vol. 102, no. 6, p. 063 002, 2020. DOI: 10.1103/PhysRevD.102.063002. arXiv: 1912.03300 [hep-ph] (cit. on pp. 91, 112).

[142] Ferrari, A., Sala, P. R., Fasso, A., and Ranft, J., "FLUKA: A multi-particle transport code (Program version 2005)," Oct. 2005. DOI: 10.2172/877507 (cit. on pp. 91, 112).

[143] Nelson, W. R., Hirayama, H., and Rogers, D. W. O., "The Egs4 Code System," Dec. 1985, SLAC-0265, SLAC-265, SLAC-R-0265, SLAC-R-265 (cit. on p. 91).

[144] Abdul Rehman, Frank G. Schroeder, Alan Coleman, *Radio signal interpolation using star-shape pattern*, Internal Report, 202007002, Jul. 2020 (cit. on pp. 91, 112).

[145] Heck, D. and Pierog, T., "Extensive air shower simulation with corsika: A user's guide," *Forschungszentrum Karlsruhe, Institut für Kernphysik*, 2000 (cit. on p. 92).

[146] James, F. and Roos, M., "Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations," *Comput. Phys. Commun.*, vol. 10, pp. 343–367, 1975. DOI: 10.1016/0010-4655(75)90039-9 (cit. on p. 99).

[147] Schoorlemmer, H. *et al.*, "Energy and Flux Measurements of Ultra-High Energy Cosmic Rays Observed During the First ANITA Flight," *Astropart. Phys.*, vol. 77, pp. 32–43, 2016. DOI: 10.1016/j.astropartphys.2016.01.001. arXiv: 1506.05396 [astro-ph.HE] (cit. on p. 103).

[148] Schröder, F. G., "Instruments and Methods for the Radio Detection of High Energy Cosmic Rays," Ph.D. dissertation, Karlsruhe U., 2010. DOI: 10.1007/978-3-642-33660-7 (cit. on p. 106).

[149] Corstanje, A. *et al.*, "Depth of shower maximum and mass composition of cosmic rays from 50 PeV to 2 EeV measured with the LOFAR radio telescope," *Phys. Rev. D*, vol. 103, no. 10, p. 102 006, 2021. DOI: 10.1103/PhysRevD.103.102006. arXiv: 2103.12549 [astro-ph.HE] (cit. on pp. 107, 118).

[150] Bezyazeekov, P. A. *et al.*, "Reconstruction of cosmic ray air showers with Tunka-Rex data using template fitting of radio pulses," *Phys. Rev. D*, vol. 97, no. 12, p. 122 004, 2018. DOI: 10.1103/PhysRevD.97.122004. arXiv: 1803.06862 [astro-ph.IM] (cit. on pp. 107, 118).

[151] Abreu, P. *et al.*, "The depth of the shower maximum of air showers measured with AERA," *PoS*, vol. ICRC2021, p. 387, 2021. DOI: 10.22323/1.395.0387 (cit. on pp. 107, 118).

[152] Aartsen, M. G. *et al.*, "Cosmic ray spectrum and composition from PeV to EeV using 3 years of data from IceTop and IceCube," *Phys. Rev. D*, vol. 100, no. 8, p. 082 002, 2019. DOI: 10.1103/PhysRevD.100.082002. arXiv: 1906.04317 [astro-ph.HE] (cit. on p. 109).

[153] NOAA, *Global data assimilation system*, https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/global-data-assimilation-system-gdas (cit. on p. 112).

[154] Mitra, P. *et al.*, "Reconstructing air shower parameters with LOFAR using event specific GDAS atmosphere," *Astropart. Phys.*, vol. 123, p. 102 470, 2020. DOI: 10.1016/j.astropartphys.2020.102470. arXiv: 2006.02228 [astro-ph.HE] (cit. on p. 112).

[155] Horneffer, A. *et al.*, "LOPES: Detecting radio emission from cosmic ray air showers," *Proc. SPIE Int. Soc. Opt. Eng.*, vol. 5500, Hough, J. and Sanders, G. H., Eds., p. 129, 2004. DOI: 10.1117/12.551466. arXiv: astro-ph/0409641 (cit. on p. 122).

[156] Ahn, E.-J., Engel, R., Gaisser, T. K., Lipari, P., and Stanev, T., "Cosmic ray interaction event generator SIBYLL 2.1," *Phys. Rev. D*, vol. 80, p. 094 003, 2009. DOI: 10.1103/PhysRevD.80.094003. arXiv: 0906.4113 [hep-ph] (cit. on p. 126).

[157] Corstanje, A., "Cosmic rays at LOFAR: precision measurements with a radio telescope," Ph.D. dissertation, Nijmegen U., 2019 (cit. on p. 127).