The diffuse Galactic signal in RNO-G

Bachelor-Arbeit

zur Erlangung des Grades

Bachelor of Science (B.Sc.)

Physik

an der Naturwissenschaftlichen Fakultät der
Friedrich-Alexander-Universität Erlangen-Nürnberg

vorgelegt am 31.08.2023

von **Moritz Reichert**

Betreuerin: Prof. Dr. Anna Nelles

Contents

1 Introduction

Nowadays humankind is capable of measuring several different types of cosmic messengers. Messengers are for instance cosmic particles, radiation and gravitational waves ([1]). They are created in astronomical objects like galaxies or stars or in astronomical processes such as supernovae. The messengers are interesting because they contain information about their source and can therefore be used to check physical theories. For instance, a century after Albert Einstein's theory of gravity predicted gravitational waves, it was finally in 2015 possible for the first time in the Laser Interferometer Gravitational-Wave Observatory (LIGO) to detect gravitational waves from two merging black holes ([2]). Moreover, charged particles, which are part of cosmic particles, are the cosmic messengers with the highest energies ever detected ([1]). They are used to study astronomical particle accelerators. However, they have one big weakness: they are deflected by electric, magnetic and gravitational fields and they are absorbed by interstellar dust. This makes it impossible to trace their origin.

Neutrinos do not have this flaw.
Neutrinos are electrically neutral elementary particles with a rest energy less than 0.8 eV. Since they only interact through the weak force, they hardly interact with matter. This makes them difficult to measure. But this is also what makes neutrinos so interesting, because it means that they can mostly pass through dust and even planets without being absorbed. Thus, if you can measure neutrinos and trace their direction of travel, you will know where they came from. Neutrinos are created in various particle reactions. For example at acceleration sites of cosmic rays. There is an important creation process of neutrinos: the collision of cosmic ray particles with slower particles. In this interaction, new particles are created which can decay into neutrinos. For instance ([1]):

$$p + \gamma \longrightarrow n + \pi^+ \longrightarrow n + \mu^+ + \nu_\mu \longrightarrow n + e^+ + \nu_e + \bar{\nu}_\mu + \nu_\mu \tag{1}$$

where p is a proton, n a neutron, e^+ a positron, π^+ a pion, μ^+ a muon, ν a neutrino, $\bar{\nu}$ an anti-neutrino and γ is a photon.

There are already several detectors which are capable or plan to measure neutrinos. Optical neutrino detection is well established. The IceCube detector has recently observed the Milky Way in neutrinos and is finding first neutrino point sources. Moreover, there is a recently upcoming technique to measure neutrinos with radio waves generated via the Askaryan effect. This measurement process will be explained in the beginning of this thesis. For instance the Radio Neutrino Observatory Greenland (RNO-G) at Summit Station can measure ultra high energy neutrinos (above 10^{16} eV ([3])). To use the Askaryan effect, it is important to understand radio background signals since they would disturb the neutrino measurement. One aim of this thesis is to find and analyse regularly occurring background signals to assist the process of finding and eliminating background signal sources.

This thesis will further study two new measurement artefacts. Known issues like digitizer artefacts, battery charging noise or wind induced signals ([4]) will not be dealt with.

Besides anthropogenic background signals, there are cosmic radio signals. For example, the ARIANNA experiment, which is a neutrino detector in Antarctica, detects the Galactic signal in 85 - 130 MHz frequency range ([5]). RNO-G is sensitive in this frequency region too, therefore the main aim of this thesis is to investigate whether RNO-G detects the Galactic signal. The Galactic

signal is often used for detector calibration if it is the primary background signal. Especially since calibration with reference antennas has uncertainties that are difficult to determine. Whereas the uncertainties of Galactic calibration can be estimated ([6]). Thus, it is planned to use Galactic calibration at RNO-G.

2 Neutrino in-ice detection with RNO-G

As mentioned in the introduction, being able to measure neutrinos would be a huge step forward. The following sections give a brief overview of how neutrinos can be measured at RNO-G. The article [7] is used as a guideline.

2.1 Neutrino in-ice detection

The idea behind in-ice neutrino detection is to measure ultra high energy neutrinos (~10^{18} eV) indirectly via the Askaryan-effect. The count of interactions between neutrinos and matter is proportional to the count of neutrinos and the amount of matter available for interaction. The neutrino flux cannot be changed, so you want to increase the volume of the detector as much as possible. As soon as a high energy neutrino interacts with for example ice, a cascade of secondary particles is created. The creation and propagation of these secondary particles is often called particle shower. The particle shower accumulates electrical charge over time through various effects. Consequently, it emits detectable radiation in the radio frequency range called Askaryan radiation. Further investigations considering costs and other aspects show that ice is particularly suitable for the detector volume. It has been decided to construct the neutrino detector RNO-G at Summit station in Greenland where the ice sheet has a thickness of about 3 km ([3]).

To measure the radio waves emitted by the Askaryan effect, several antennas have been deployed deep in the ice. However, there are many background signals in the radio frequency range which must be eliminated from measured data. Examples are Walkie-Talkies or air traffic communication. These noise signals generally travel from the surface downwards in contrast to neutrino radio signals which travel upwards since they are created deep in the ice. This is why additional antennas are placed just below the ice surface to detect background signals. The detected signals can then be removed from measured data. The latter antennas are used in this thesis for a background signal analysis.

2.2 Sky-temperature conversion to signal

The last section explained that one can detect neutrinos by measuring radio waves. Radio waves are a combination of electric and magnetic fields. If they hit a metallic antenna, they cause free electrical charges to oscillate in the antenna. The moving charges create a voltage across the antenna which can be measured. However this process does not have any spatial resolution besides the spatial dependency of the sensitivity of the detector/antenna. Therefore, the detector measures the sum of all radio signals arriving at its antenna (multiplied with the antenna sensitivity).
The issue is that, in addition to neutrinos, there are many radio signal emitters. For example, astronomical objects such as stars or galaxies are important emitters. The detectors receive signals from them all the time. The signal power from the sky can be calculated with the sky temperature T_{Sky} ([8]):

$$P_{\text{Sky}}(t,f) = \frac{Z_0}{Z_L}\frac{k_B}{c^2}\int_\Omega f^2\, T_{\text{Sky}}(t,f,\alpha,\delta)\cdot|H_i(f,\alpha,\delta)|^2\,\mathrm{d}\Omega \tag{2}$$

where t is the time and f the frequency to look at. k_B is the Boltzmann constant, c the speed of

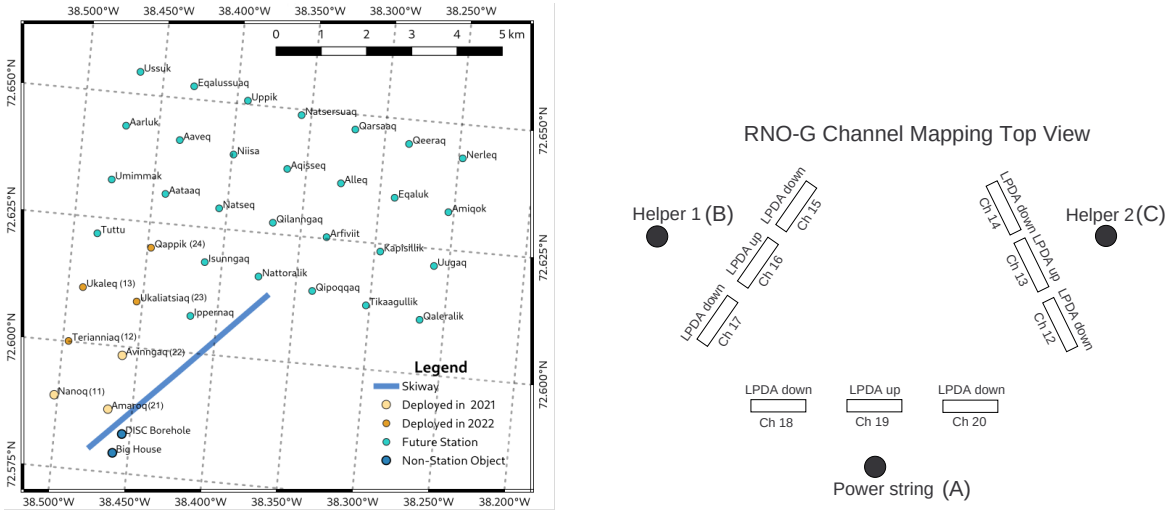

Figure 1: The left map shows the station positions. The number of each station is written after the station name. The right picture shows position of the upward (13, 16, 19) and downward (12, 14, 15, 17, 18, 20) facing channels. The images are taken from [9].

light. Z_L is the impedance of the detector and Z_0 the free space impedance. The antenna sensitivity is described by H_i which is dependent of frequency and direction.

With that formula one can simulate the radio signals received from the sky which is done in section 5.4.

2.3 Detector setup

RNO-G comprises multiple detector stations, each designated by a number. Each station has several antennas which are referred to by channel and are as well as designated by a number. All channels with the same number have the same function. This analysis uses the nine channels 12 - 20 of each station. There are two types of channels: upward and downward facing ones. Log Periodic Dipole Antennas (LPDAs) have been used for both types. Figure 1 shows the position of the stations and an overview of the channels. The upward antennas/channels 13, 16 and 19 are placed just below the ice surface. Their task is to measure background signals originating from the surface/atmosphere/space. Therefore, they are mainly sensitive in upward direction. Whereas downward facing channels 12, 14, 15, 17, 18 and 20 are not sensitive to those background signals but to upward moving radio waves. Thus, they are additional antennas for reconstructing and identifying neutrino candidates. Figure 2 shows the complete station design.

Each channel records the voltage applied to its belonging antenna at a sampling rate of 3.2 GHz. The voltage is directly processed by the electronic signal chain which comprises for instance the amplifier and an ADC (analog-digital converter). The result are ADC-counts which are proportional to the voltage applied to the antenna. Therefore, voltage is often used as a synonym for ADC-counts. 128 voltage samples are saved in digitizer blocks. 16 blocks fill a trace which consequently consists of 2048 consecutive voltage samples measured over a time span of 640 ns. A trace can be used for data analysis. Consecutive forced trigger traces of about two hours are packed in one data file which is called run.

Since saving voltage data throughout the whole time takes up too much hard disk space and is not

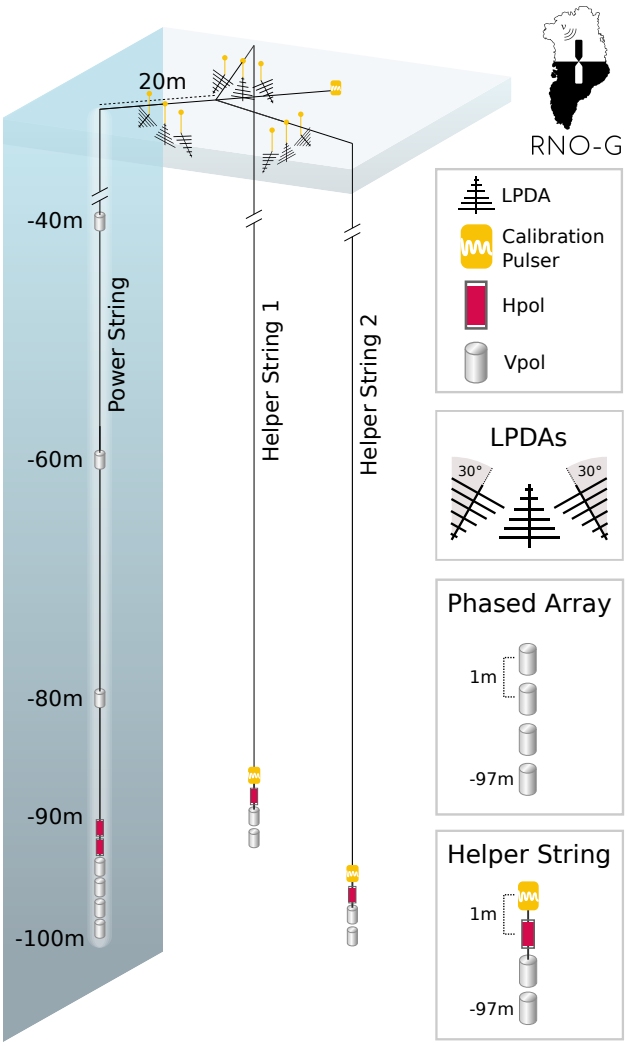

Figure 2: The image shows the station design of RNO-G. At the top are upward and downward facing channels which are used in this thesis. It is taken from [9].

really necessary, triggers have been introduced. A trigger determines while measuring whether a measured trace should be kept and saved or dismissed. This analysis uses data saved by forced triggers. A forced trigger saves every ten seconds the current trace. Thus, neutrino signals are not expected in these traces, as the signals are too short and rare to be measured by chance in a forced trigger event. Instead background signals are measured.

3 Methods

As explained in the previous chapter the detectors at RNO-G detect radio waves by measuring voltages. This chapter explains the process of converting measured voltages into a spectrum. Starting with a description of the data which is used in this thesis. Furthermore, a special type of time binning used for spectrum analysis is presented.

3.1 Available data

This thesis uses data from 2021, 2022 and 2023 and only considers forced trigger events. The data consists of "digitized" waveforms given in ADC counts. The gain has not been calibrated, thus a data normalisation procedure is performed, which is explained in the next section.
At RNO-G the stations 11, 21 and 22 have been running since 2021. In 2022 the stations 12, 13, 23 and 24 were added. The analysis is done on a private laptop with limited hardware resources, thus only data from three specific time scopes is used for this analysis: 09.2021, 07.2022, 05.2023. Each one lasts for a minimum of two weeks and up to nine weeks. The measurement time scopes differ slightly between stations. The exact dates can be seen in table 1.

Station	2021	2022	2023
11	03.09 - 21.09	21.07 - 15.08	18.04 - 06.06
12	-	22.06 - 07.08	-
13	-	06.07 - 15.08	24.04 - 06.06
21	03.09 - 21.09	14.06 - 14.08	31.03 - 06.06
22	03.09 - 21.09	25.06 - 15.08	-
23	-	27.06 - 15.08	24.04 - 06.06
24	-	13.07 - 15.08	21.04 - 06.06

Table 1: Time durations from which data is taken for this analysis.

The stations did not measure over the entire measurement time scope. There are measurement pauses ranging from minutes to several hours at which the detectors did not take any data.
All of the data from 2023 has been transmitted by satellite. Because of limited uplink bandwidth of the satellite, only 3 - 10 % of the measured data is available for this analysis. Due to the small amount of accessible data, the normalisation procedure explained below does not work so well for that data. Therefore, the normalised data from 2023 is significantly more noisy than expected and is consequently hardly used.
In addition, a laboratory measurement of the signal chain of a detector is available. For that the antenna was removed. Thus only thermal noise was measured by the digitizer and amplified. By comparing that laboratory measurement with in-situ measurements one can detect radio signals.

3.2 Converting measured voltage to spectrum

From the detectors one obtain traces. A trace consists of $N = 2048$ voltage measurements taken at a rate of $f_S = 3.2\,\text{GHz}$. So it covers a time span of 640 ns. By using the discrete Fourier transformation, the voltage values can be transformed into frequency domain which is called spectrum.

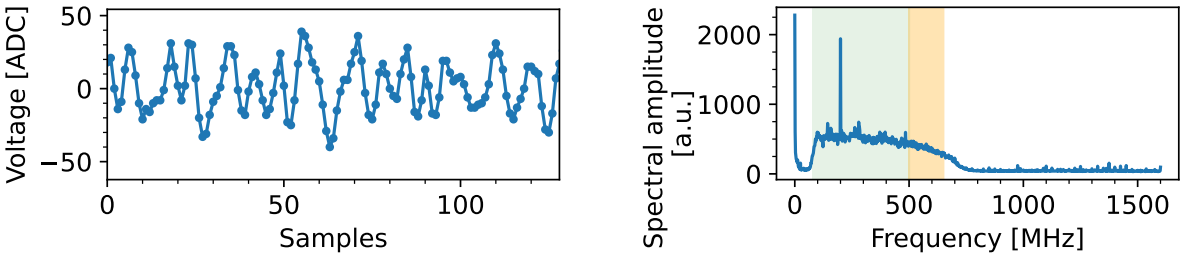

Figure 3: The left plot shows the first 128 samples of a trace. The mean spectrum over one run is visible in the right plot. In the green and orange area (80 - 650 MHz) the amplification works reliable. 500 - 650 MHz (orange area) is used for normalisation.

Due to the characteristics of the Fourier transformation the spectrum shows which frequencies are present in the voltage data. Since the trace is not continuous but discrete, the frequency domain is as well as discrete. The frequencies f_n can be calculated with the sample spacing $d := 1/f_S$ and the count of data points N:

$$f_n = \frac{n}{N \cdot d} = 1.5625\,\text{MHz} \cdot n \qquad n \in [0,\ 1,\ \ldots,\ \lfloor N/2 \rfloor] \tag{3}$$

Moreover as the Nyquist–Shannon sampling theorem states, the range of the discrete frequencies is limited to $f_S/2 = 1.6\,\text{GHz}$. This limitation is already included in equation 3. That equation shows that the obtained spectrum has a resolution of about 1.5 MHz.
If U_n are the voltage values at time point $t_n = n \cdot d$, the spectrum $S(f_k)$ at frequency f_k is calculated with the Fast Fourier transformation (FFT):

$$S(f_k) = \alpha \cdot \left| \sum_{n=0}^{N-1} U_n \cdot \exp(-i2\pi kn/N) \right| \tag{4}$$

The factor $\alpha \in \mathbb{R}$ is constant and is used for correct normalisation and conservation of power. In the following analysis only ratios of spectrum values matter which eliminate this factor. So the factor α is irrelevant.
The return values of the FFT are complex values containing the phase and the signal amplitude of a discretised frequency. For the following analysis, only the amplitude is important. Thus the absolute value of the complex return value is calculated which gives the amplitude. An example trace and spectrum is shown in figure 3.

There is a Python project called *NuRadioReco* which is a "reconstruction framework for radio detectors of high-energy neutrinos and cosmic-rays". It is available on GitHub (see paper [10]). It contains the function *NuRadioReco.utilities.fft.time2freq* which converts traces into a spectrum. This function is used to calculate the FFT of a trace.

When you calculate the spectrum for each detector with equation 4, you see that the results of different downward and upward facing LPDAs do not match even in frequency domains where no signals are expected because the gain values of the detectors are not calibrated. For that reason the spectrum must be normalised.

In general we expect that every LPDA measures at high frequencies (> 500 MHz) only thermal

noise. Thus this frequency region is ideal for normalisation. It is wise to limit the frequencies used for normalisation to about 650 MHz because the amplification vanishes in higher frequencies.
To gain a more stable normalisation, the following normalisation procedure uses the mean spectrum over one run which is denoted by $\overline{S}(f)$. Moreover the normalisation is applied to each channel group, which consists of one upward facing channel and its neighbouring two downward facing channels. $S_{\mathrm{U}}(f)$ is used for the spectrum of the upward facing channel and $S_{\mathrm{D},\,i}(f)$ is the spectrum of the i-th downward facing channel.
The normalisation procedure consists of two successive steps.

Firstly, the spectrum of the upward and downward facing channel must be approximately the same at high frequencies because the detectors measure only thermal noise. Formulated mathematically, this means that m_i should be one in the following expression:

$$m_i = \langle \frac{\overline{S}_{\mathrm{U}}(f)}{\overline{S}_{\mathrm{D},\,i}(f)} \rangle_{500\,\mathrm{MHz}<f<650\,\mathrm{MHz}} \stackrel{!}{=} 1 \tag{5}$$

The term $\langle\cdot\rangle_{f_1<f<f_2}$ stands for taking the average of $\cdot$ for all $f \in (f_1,\, f_2)$. Multiplying the spectrum of the downward facing channel $S_{\mathrm{D},\,i}(f)$ by m_i completes the first normalisation step.

Secondly, it is advisable to additionally normalise the downward facing channels to the laboratory measurement $S_{\mathrm{L}}(f)$. This ensures that the spectrum is always in a similar magnitude and can be easily compared with the laboratory measurement. For that, one wants $\tilde{m}$ to be one in the next equation:

$$\tilde{m} = \langle \frac{(m_1 \cdot \overline{S}_{\mathrm{D},\,1}(f) + m_2 \cdot \overline{S}_{\mathrm{D},\,2}(f))/2}{S_{\mathrm{L}}(f)} \rangle_{500\,\mathrm{MHz}<f<650\,\mathrm{MHz}} \stackrel{!}{=} 1 \tag{6}$$

Multiplying $S_{\mathrm{U}}(f)$, $S_{\mathrm{D},\,1}(f)$ and $S_{\mathrm{D},\,2}(f)$ additionally by the factor $1/\tilde{m}$ finishes the normalisation process.

The normalisation described above has the great advantage of removing the influence of variations of the amplifier temperature from the spectra. This can probably be proven mathematically.

With the normalised spectra one can calculate the lab-excess (short form for laboratory excess) $E_{\mathrm{L}}(f)$. The lab-excess indicates how much more a channel measures than it should be according to the laboratory measurement. Channel detect higher signals for example if a radio wave arrived at the detector.

$$E_{\mathrm{L}}(f) := \frac{S(f)}{S_{\mathrm{L}}(f)} \tag{7}$$

Furthermore one can calculate the excess $E(f)$. The excess indicates the strength of signals coming from the surface, atmosphere or space because the downward facing channels are hardly sensitive to these directions.

$$E(f) := \frac{S_{\mathrm{U}}(f)}{(S_{\mathrm{D},\,1}(f) + S_{\mathrm{D},\,2}(f))/2} \tag{8}$$

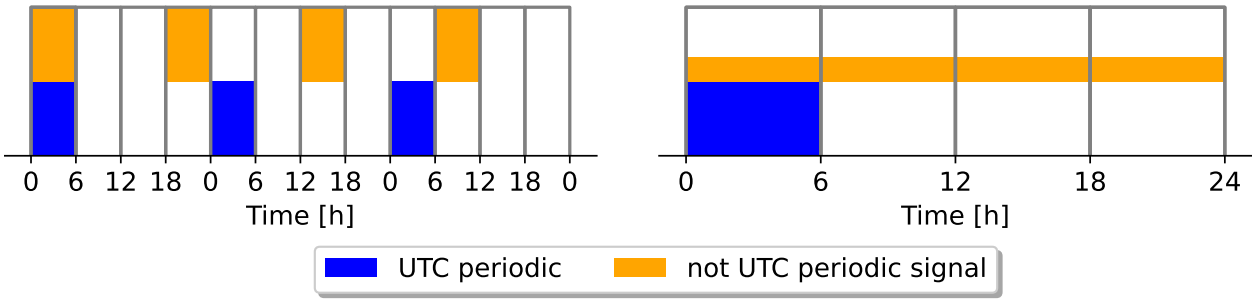

Figure 4: Illustration of the time binning process: three days are shown on the left with time bin sizes of 6 hours. The blue signal is measured every 24 hours. Therefore it is UTC periodic. The orange signal occurs every 18 hours. On the right the mean over all (three) bins at one time point is shown. Since the blue signal is in all three bins at 0 - 6 h, the mean is at full strength. Whereas the orange signal is only in one of three bins at a all times, thus the mean is at $^1/_3$-strength. The median, which is used in the time-binning procedure instead of the mean, would be the same for the blue signal and always zero for the orange signal.

The (lab-) excess is often given in decibels calculated by $20 \cdot \log_{10} E_{(\mathrm{L})}(f)$ [dB].

3.3 Time zones

Two time zones are particularly important for this analysis. The first one is coordinated universal time (UTC). UTC is synchronised with the rising and setting of the Sun. One UTC day consists of 24 hours. It is therefore also synchronous with human activity and anthropogenic signals are generally periodic in UTC time.

The second important time type is local sidereal time (LST). LST is synchronised with the rising and setting of distant (not moving) stars. A sidereal day lasts for as long as it takes the Earth to complete one complete rotation on its own axis from the perspective of a distant star. In one Earth year the Earth rotates 365 times around its own axis. But from the perspective of a distant star, the Earth makes an extra rotation on its own axis because the Earth's orbit around the Sun takes one Earth year to complete. Therefore the Earth rotates 366 times in one year and a sidereal day takes $^{365}/_{366} \cdot 24\,\mathrm{h} \approx 23\,\mathrm{h}\ 56\,\mathrm{m}\ 4\,\mathrm{s}$.
Since LST is synchronised with the rising and setting of stars, radio signals from stars or galaxies are periodic in that time.

Knowing whether a radio signal is man-made or produced by an astronomical object is a good first step in finding its source. Therefore it is helpful to know its period.

3.4 Time binning

For reasons of limited hardware resources of the private laptop used for this analysis, the normalised spectrum and excess are binned in time. The aim of this analysis is to find signals which are periodic in sidereal or UTC time. To find UTC periodic signals a time bin size is chosen such that it is a divisor of the UTC day. The advantage of this is that the times for the bins are repeated. Thus you can calculate the median over all bins with the same time. The result shows UTC periodic

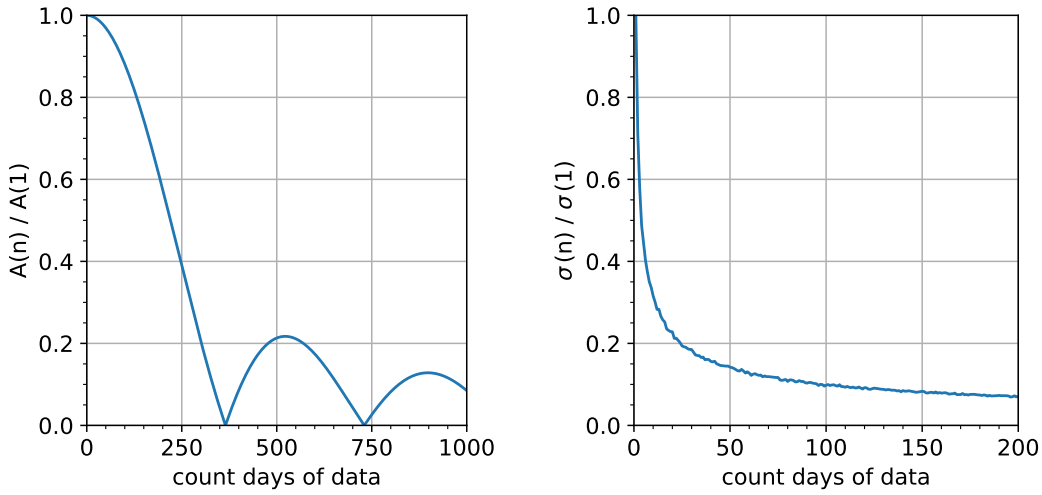

Figure 5: In the left picture an UTC time binning is applied to a sidereal periodic sinus function. The amplitude of the resulting sinus function is shown. In the right picture the UTC time binning is applied to time bins which have Gaussian distributed values. The standard deviation of the result values is plotted. The curve is, as expected, approximately the same as $1/\sqrt{n}$ where n is the count of days.

signals whereas other periods are smeared out or completely removed. This process is illustrated in figure 4. The same procedure is used for sidereal time.

The left plot of figure 5 shows how slow a sidereal periodic sinus function smears out in an UTC time binning. The result of the time binning is again a sinus function but with decreased amplitude. Binning 14 days of data results in an amplitude decrease of 0.3 % which is in a real measurement probably dominated by noise. To decrease the amplitude by 10 % 90 continuous measurement days are required. The amplitude falls to zero if there is $k \in \mathbb{N}$ with $n \cdot 24\,\mathrm{h} = k \cdot (23\,\mathrm{h}\ 56\,\mathrm{m}\ 4\,\mathrm{s})$ where n is the count of measurement days. Examples are $n = 365 \cdot \mathbb{N}$. Due to measurement pauses, it is unrealistic to achieve this condition in reality. Therefore time binning is normally useless to separate sidereal and UTC periodic signals. Instead, one can take two measurement time scopes, do the time binning for each and look at the time shift between the two result signals. The time shift is zero, if the period of the signal is the same as of the time binning.

The right plot of figure 5 shows that time binning reduces contributions of (high frequency) noises. For that the time binning is applied to time bins which have Gaussian distributed values. Already with 14 measurement days the standard deviation of the noise decreased by 74 %.

To increase time resolution the time binning process is enhanced by a simple moving average. For that time bins are overlapping. The time difference between the start of two consecutive bins is called time step size and is 20 % of the bin size. The concrete values for the time binning process can be seen in table 2.

Time	Day length	Bin size	Time step size	Count bins in one day
UTC	24 h	1600 s	320 s	270
sidereal	23 h 56 m 4 s	1657 s	331.4 s	260

Table 2: The values used for the time binning process.

4 Overview of background signals in spectrum

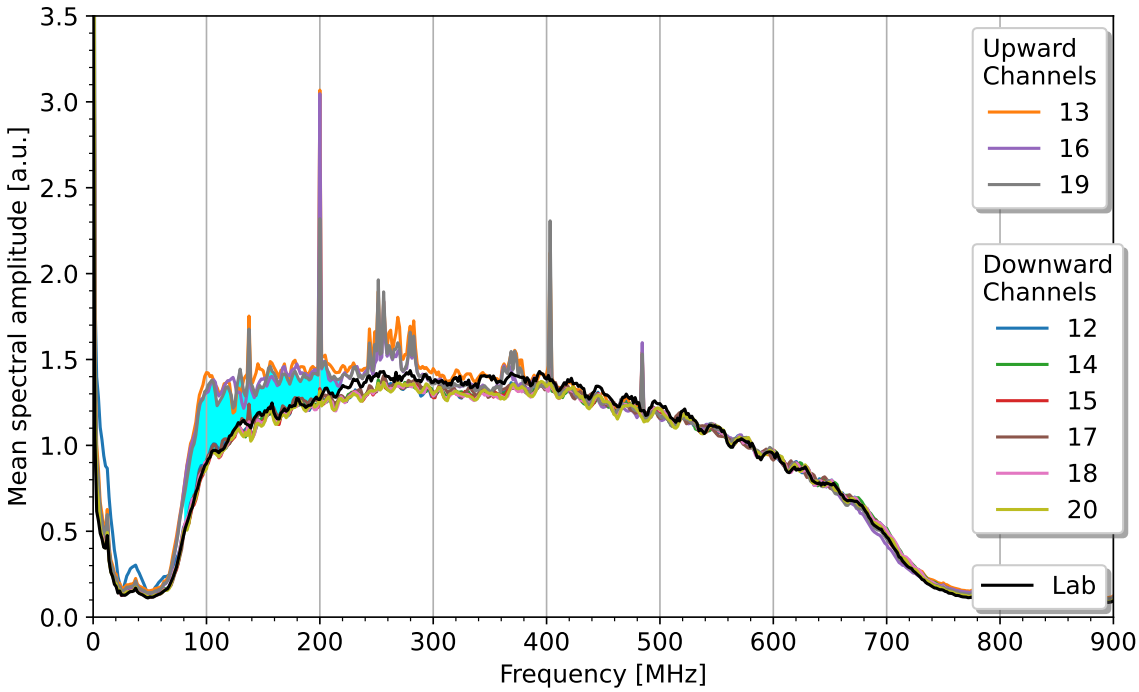

Figure 6: The mean normalised frequency spectra of station 23 measurements from 09.08.2022 1 am - 10.08 5 pm are shown. Channel 13, 16 and 19 are upward facing the rest are downward facing ones. The "Lab" curve is the laboratory measurement.

To improve the indirect measurement of neutrinos using the Askaryan effect via radio waves, it is necessary to suppress noise signals. Therefore, finding and eliminating noise sources is an important task. To assist this process, this chapter analyses periodically occurring noise signals which have been measured by detectors.

Calculating the spectrum of measured data and normalising it as described in section 3.2 results in diagrams like in figure 6. To this end, plot data from station 23 from 09.08.2022 1 am - 10.08 5 pm has been taken and averaged.

Is a signal stronger in upward than in downward facing channels, as explained in section 3.2, implies that it is a radio wave coming from the surface, atmosphere or space. Regarding that fact, one sees already some interesting features in the spectrum shown in figure 6:

1. Only upward facing channels see more signals than the laboratory measurement. Thus all visible signals come from surface, atmosphere or space.
2. Four sharp peaks are visible at 136, 200, 400 and 484 MHz.
3. Two accumulations of signals: 240 - 290 MHz & 360 - 380 MHz
4. Clear excess in received power in the 80 - 200 MHz frequency range (cyan coloured area)

In the following sections background signals like sharp peaks or signal accumulations will be discussed. The 200 MHz signal is covered in section 4.2. The signal excess in cyan coloured area is thought to be created by the Milky Way. Chapter 5 deals with that signal.

4.1 Background signals

In the following, regular occurring background signals are presented. Among other things the periodicity is determined. For that plots like in figure 8 are created. The plots use the time binning method to show the median development of a signal over a day. If the curves in a channel from different measurement time scopes are congruent, it means that the displayed signal has the same period as the time binning. For example, this is the case for all channels in the figure, therefore the shown signals are periodic in LST. This plot can also be used to determine which channel is detecting the signal.

4.1.1 Air traffic, Walkie-Talkie frequencies : ~125, 136, ~150 MHz

Next to Summit Station is a small runway for aircraft. The air planes use frequencies at ~125 MHz for communication. Moreover, there are several walkie-talkie frequencies with irregular signal transmission. The frequencies are in the range 151 - 157 MHz and at 136 MHz. The Walkie-Talkie frequencies are split into eight channels each with a bandwidth of 12.5 kHz which can be used for communication. A detailed list can be found in the following table or at [11].

Channel	Frequency [MHz]	Receive / Transmit
1	156.800	✓/ ✓
2	151.000	✓/ ✓
3	151.625	✓/ ✓
4	152.000	✓/ ✓
5	153.000	✓/ ✓
6	136.000	✓/ ✗
7	151.000	✓/ ✓
8	151.625	✓/ ✓

Table 3: The Walkie-Talkie frequencies used near the Summit Station [11].

4.1.2 Semi-circle signal: 243.7 & 248.4 MHz

In every station the upward facing channels detect two signals at discrete frequencies 243.7 and 248.4 MHz. Those two signals are periodic in sidereal time and look pretty much identical as one can see in figure 7. Therefore, one can conclude that both signals are originating from identical sources. On top of that in station 22 and 21 the downward facing channel 20 can see both signals. The start and end time of both signals vary between channels probably because of differences in spatial sensitivity. In general both signals start at 23 - 02 h and end at 13 - 16 h (LST). In this time range the signal looks like a semi-circle. Since the connection between signal start/end time and channel might be interesting, figure 8 has been created. It shows the LST time binned median

lab-excess (the red curve in figure 7) for all channels in which these signals are visible. For better visibility the lab-excess E_L is not directly shown but $(E_\mathrm{L} - \min E_\mathrm{L})/(\max E_\mathrm{L} - \min E_\mathrm{L}) \in [0, 1]$ is plotted. The data from 2023 is in general a lot noisier thus data lines of 2023 must be treated carefully. Since in nearly all channels the transformed median lab-excess curves of different measurement time scopes are congruent, one knows that the signal must be periodic in LST. The source of this signal is still unknown.

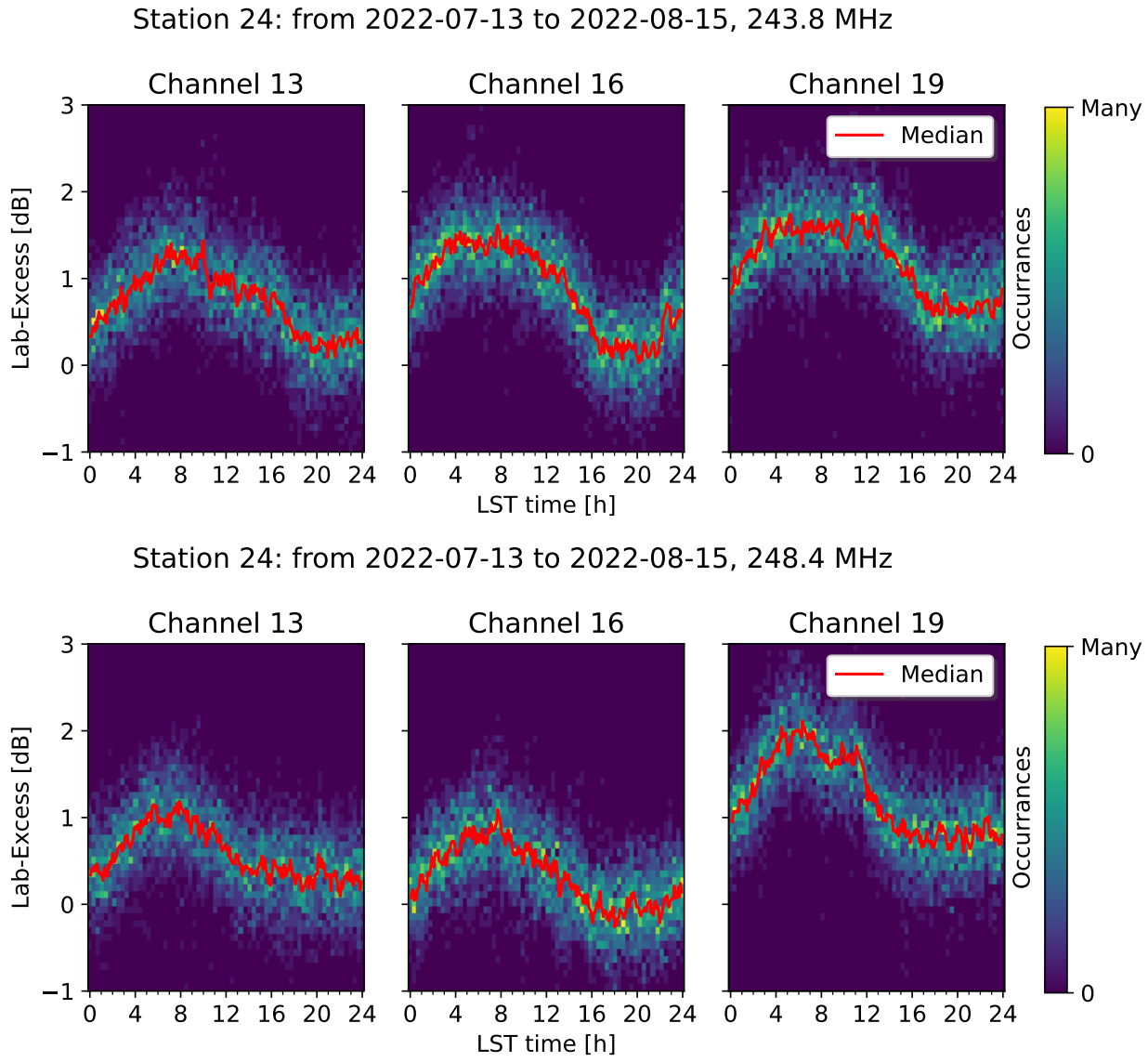

Figure 7: The signal at 243.7 MHz (top) and 248.4 MHz (bottom) are shown as a function of the local sidereal time binned lab-excess. Data is taken from station 24 from 2022-07-13 to 2022-08-15. The bluish quads show the underlying distribution of the time binning result. The red curve is the median of that distribution. All shown channels are upward facing.

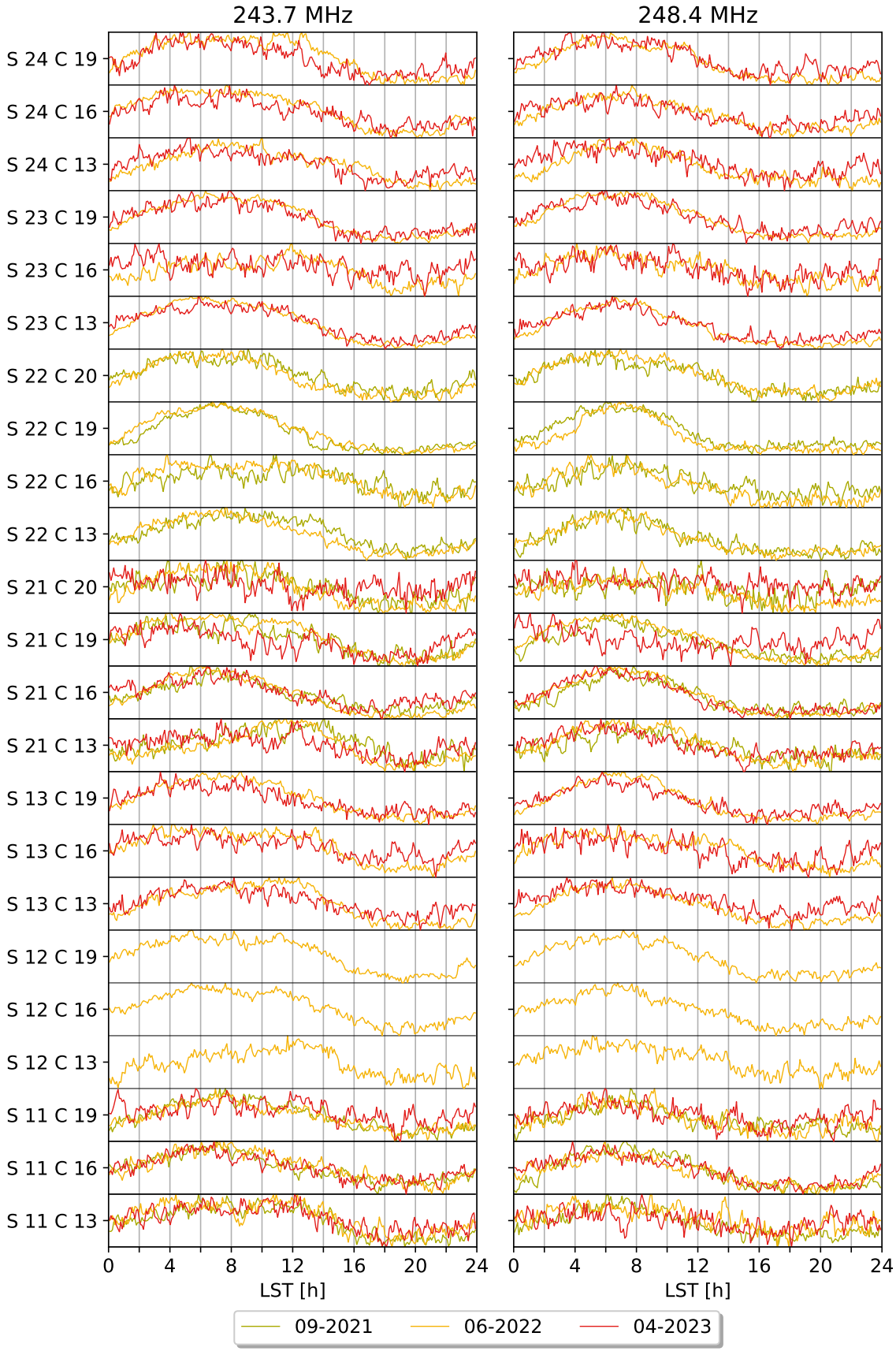

Figure 8: The signal at 243.7 MHz (left) and 248.4 MHz (right) are shown as a function of the median local sidereal time binned lab-excess. For better visibility the median lab-excess E_{L} is transformed through $(E_{\mathrm{L}} - \min E_{\mathrm{L}})/(\max E_{\mathrm{L}} - \min E_{\mathrm{L}}) \in [0, 1]$ which is shown. For each channel the bottom dark line is 0 and the top line 1.

4.1.3 Short range signals: 251.6 & 256.3 MHz

All upward facing channels and the downward facing channel 17 of station 12 detect a signal at 251.6 and 256.3 MHz. The signals are probably periodic in UTC time. Both signals are shown in figure 9 (station 21) and 10 (station 22). Due to the fact that the signals of both frequencies look similar in the same channel (see figures), one can assume that they have the same/a similar signal source. One sees in many channels a peak in the median lab-excess at around 9 h. Interesting is that in the most channels the lab-excess distribution shows that the peak is often moving in time, for example in figure 10 one sees a broader distribution of high lab-excess values in 8 - 11 h UTC time range. In some channels the signal looks like that of station 21 channel 13 at 256.3 MHz (see figure 9): an approximately rectangular signal shape with a peak.
The high variability of signal shapes between channels and stations, suggests that both signals are only detectable in a short range. Thus it is probable that the signals are emitted near or by channels.

4.1.4 Sharp peaks: 278 - 285 & 482 - 485 MHz

At 278 - 285 MHz and 482 - 485 MHz there is a sharp peak which is roughly periodic in sidereal time and can be detected about every six hours. The peak lasts less than two hours. Both signals are detected by all upward facing channels. The figures 11 and 12 show that both signals look identical, thus the signal source is similar or even the same. In the 278 - 285 MHz range the signal consists probably of two discrete signals but the low frequency resolution of the detectors cannot dissolve both perfectly.

Figure 11 shows that the peaks are moving in time between different measurement time scopes in a sidereal time binning. Therefore the peaks are not exactly periodic in sidereal time. To specify the signal period, one starts by estimating the time shift between different measurement time scopes. With the signal shape shown in channel 19 the time shift can be approximated to -3 ± 1 h (from 2021 to 2022) and -2 ± 1 h (from 2022 to 2023) (higher precision is possible but not necessary). The time gap between measurement time scopes is the difference between the mean dates of the time scopes. Between the measurement of 2022 and 2021 are 324 days and between 2022 and 2023 are 283 days. With this information, one can exclude UTC periodicity, because after 283 days the signal offset would be $283 \cdot 3\,\text{m}\ 56\,\text{s} = 18.6\,\text{h} = -5.4\,\text{h}$ which does not fit to -2 ± 1 h.
However, it is possible to calculate the mean time movement of the signal in one sidereal day by dividing the signal shift by the measurement time gap. The result is -33 ± 11 s for the signal shift from 2021 to 2022 and -25 ± 12 s for the shift from 2022 to 2023. Since both mean time shifts have nearly the same value, one can assume that the signal period is by 29 ± 12 s lower than the sidereal time. The signal period (all four peaks) is therefore 23 h 55 m 35 ± 12 s. Since the period is calculated for four peaks, one peak occurs every 5 h 58 m 53 ± 3 s.

Possible source: satellites
The earth orbit is populated with a lot of satellites which are transmitting data. Thus, one should consider satellites as a possible background signal source. An investigation was carried out to search for signals from the Starlink satellites, measured by the Low Frequency Array (LOFAR) ([12]). Several frequencies in the 110 - 188 MHz range have been found where Starlink satellites radiate although downlink communication is at 10.7 - 12.7 GHz. LOFAR has a small field of view, therefore, the signals were detected only for a few seconds.

A search for periodically signals in RNO-G in the 110 - 188 MHz frequency range is done with the time binning method. However, no satellite signals are found. It must be mentioned that the time binning averages data of about 25 minutes. Thus, signals, which last only for a few seconds, would vanish in the average. For the future, it might be helpful to increase time resolution.

If one assumes that a signal comes from a satellite, one must watch out, that the signal periodicity is not the same as the orbital period of a satellite since the Earth and thus the detector is rotating.

Therefore, a signal source creating the peaks with such a special periodicity is unknown.

4.1.5 Weather balloon: 403 MHz

A weather balloon is often started around ~11 h and ~23 h UTC for two hours. The balloon sends data at 403 MHz while it flies. That signal is visible in all stations and all channels including downward facing ones.

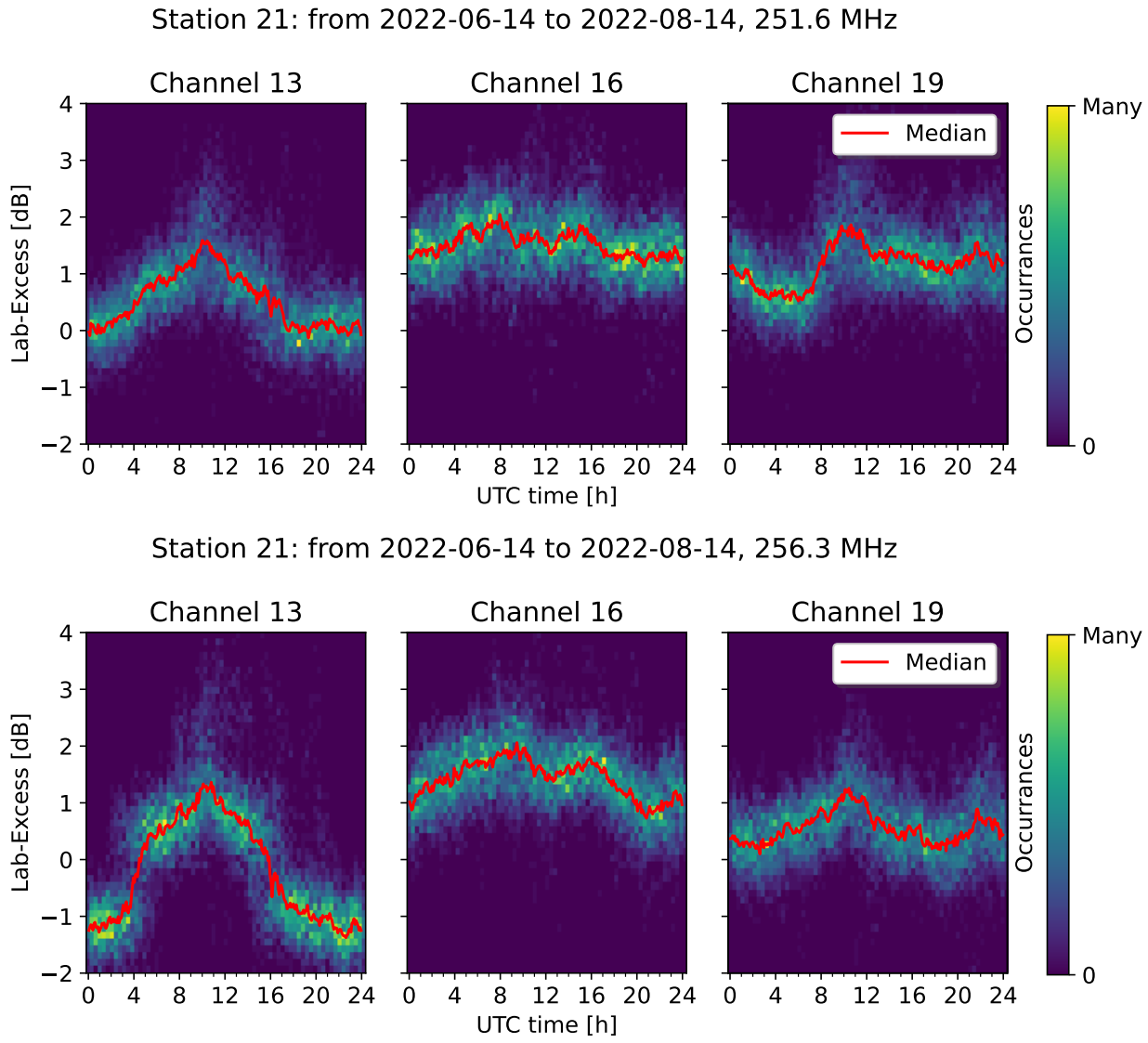

Figure 9: The signal at 251.6 MHz (top) and 256.3 MHz (bottom) are shown as a function of the UTC time binned lab-excess. Data is taken from station 21 from 2022-06-14 to 2022-08-14. The bluish quads show the underlying distribution of the time binning result. The red curve is the median of that distribution. All shown channels are upward facing. All shown channels are upward facing.

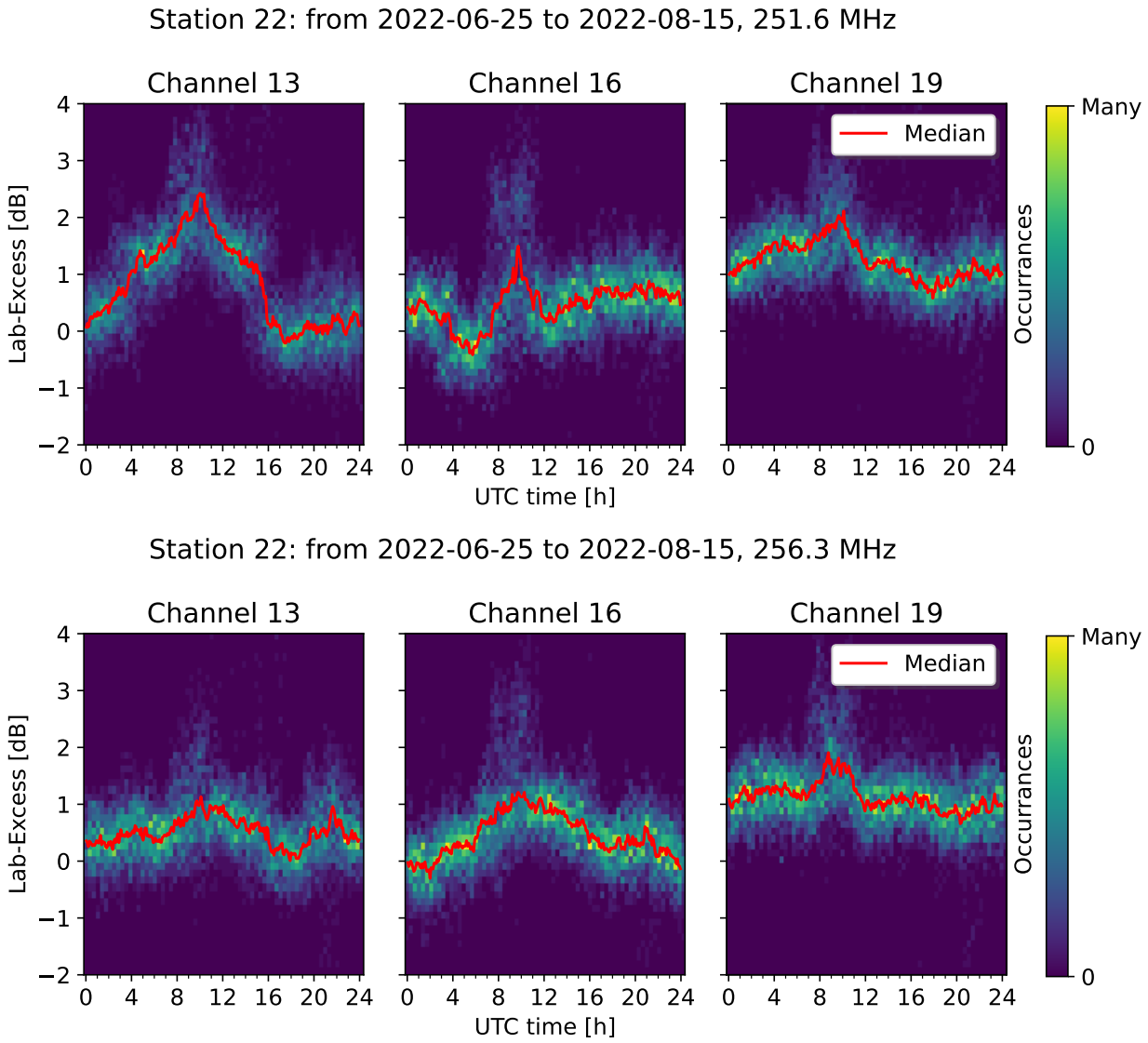

Figure 10: The signal at 251.6 MHz (top) and 256.3 MHz (bottom) are shown as a function of the UTC time binned lab-excess. Data is taken from station 22 from 2022-06-25 to 2022-08-15. The bluish quads show the underlying distribution of the time binning result. The red curve is the median of that distribution. All shown channels are upward facing. All shown channels are upward facing.

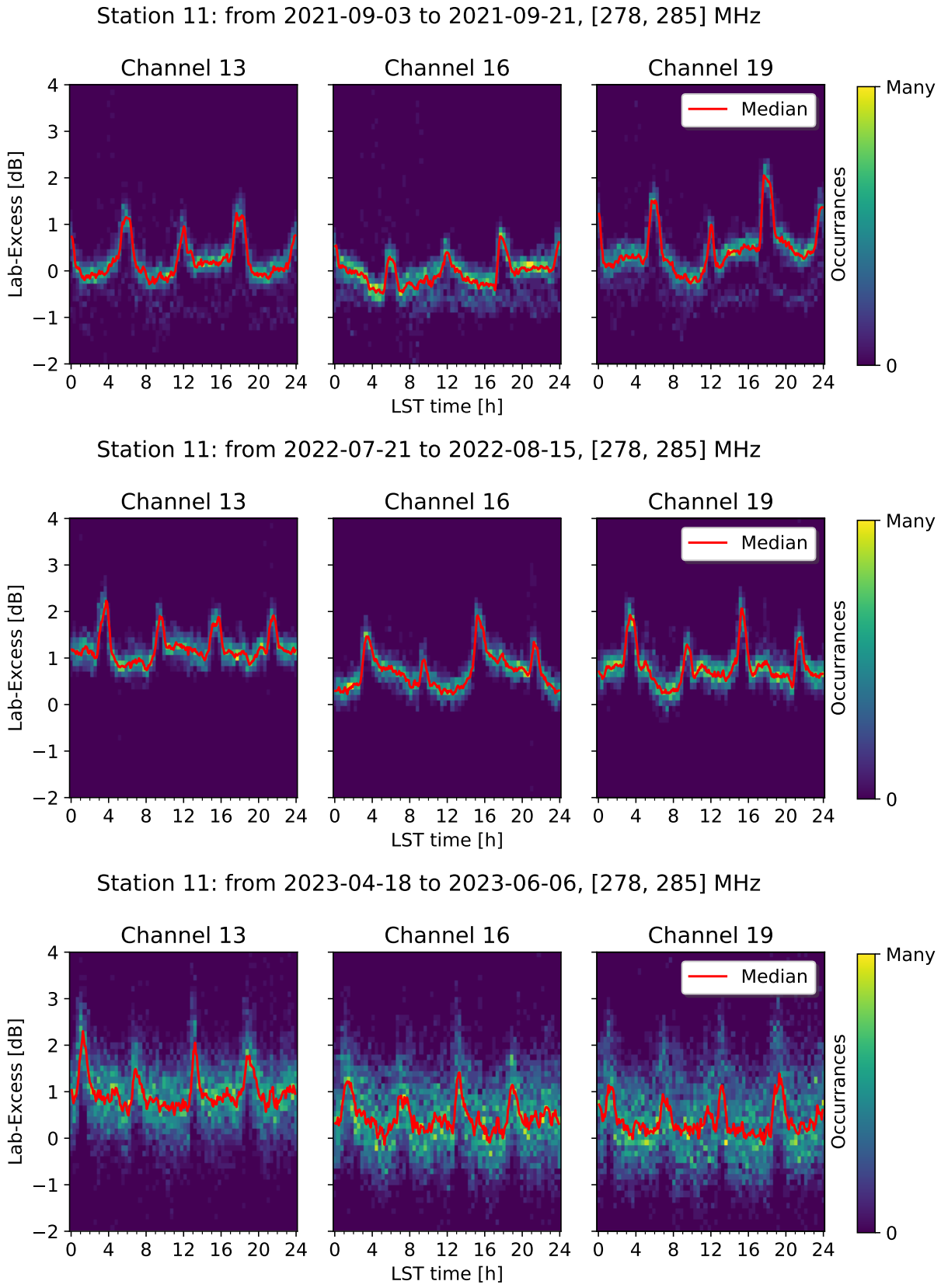

Figure 11: The average of the frequency range 278 - 285 MHz is shown as a function of the local sidereal time binned lab-excess. Data is taken from station 11 from 2021 (top), 2022 (middle) and 2023 (bottom). The bluish quads show the underlying distribution of the time binning result. The red curve is the median of that distribution. All shown channels are upward facing.

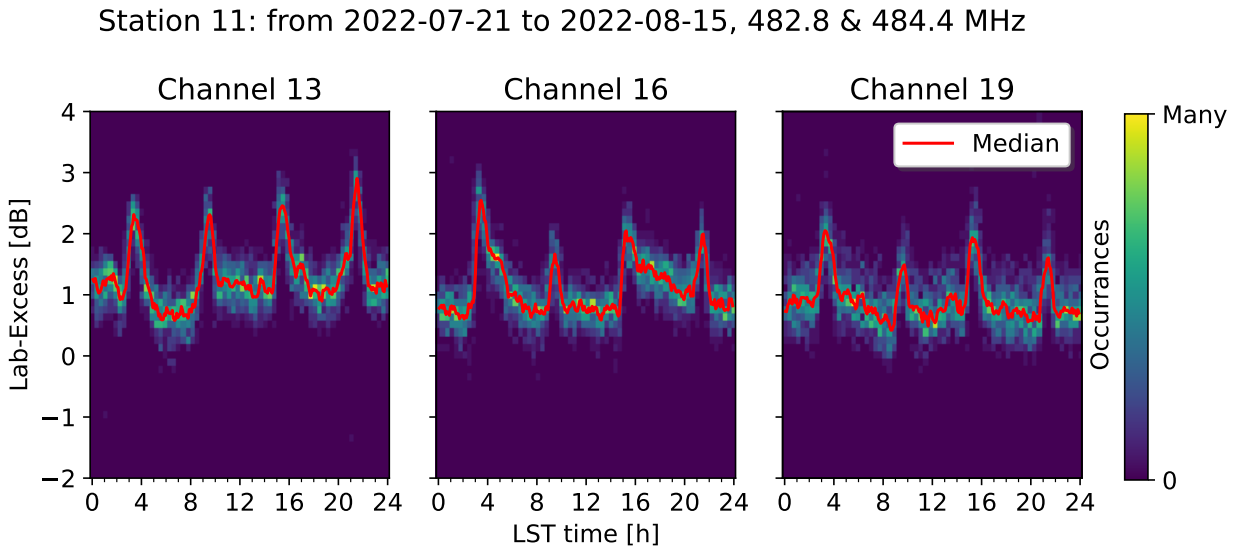

Figure 12: The average of the frequencies 482.8 & 484.4 MHz is shown as a function of the local sidereal time binned lab-excess. There are no other frequencies in the spectrum between the two. Data is taken from station 11 from 2022-07-21 to 2022-08-15. The bluish quads show the underlying distribution of the result excess values. The red curve is the median of that distribution. All shown channels are upward facing.

4.2 Constant signal: 200 MHz

The 200 MHz signal is special. It is seen in many different stations and channels: upward facing and even downward facing ones. Efforts to locate the source of this signal have been ongoing for a long time without any success.

The signal is narrowband and is constant in strength over several hours or even days. If the 200 MHz signal is visible, it is detectable in nearly every forced trigger event.

To analyse the occurrence of the 200 MHz signal, frequency spectra are averaged for each channel and run. After that, the spectral amplitude ratio of the 200 MHz line divided by the mean spectral amplitude of the neighbouring frequencies 180 - 220 MHz is calculated. Of course the 200 MHz is excluded from the neighbouring frequencies. The ratio is therefore a measure of how much stronger the 200 MHz signal is in comparison to the thermal noise in the neighbouring frequencies.
The ratio is independent of temperature effects since the effect would be multiplicative and for all frequencies identical. Furthermore, the ratio neutralises the measurement artefact "spectra jumps" in station 21 which is explained in section 4.4. The result is shown in figure 13 and 14. Moreover, calculating the average ratio over a time measurement scope can help determine which channel receives the strongest 200 MHz signal. The average value is shown after the colon in the y axis of the figures. A value of one means that the channel did not detect the 200 MHz signal.

Some interesting aspects can be extracted from the two figures 13 and 14:

1. Upward and downward channels see the signal. The signal strength is higher in upward facing channels than in downward facing ones.
2. In 2021: All available stations (11, 21, 22) have at least one channel which detects the signal.
3. In 2022: Only the stations 21, 22 and 23 see the signal. Station 11, 12, 13 and 24 do not detect the signal in any channel.
4. The signal strength can be higher/lower for a few days (see station 23 channel 13 from 15. - 25.07.2022 or station 21 channel 13, 16 and 19 from 13. - 17.07.2022)
5. In station 21 channel 13 and 19 hardly detected the signal in 2021 whereas in 2022 both see a decent amount.

Possible explanations are examined in the following.

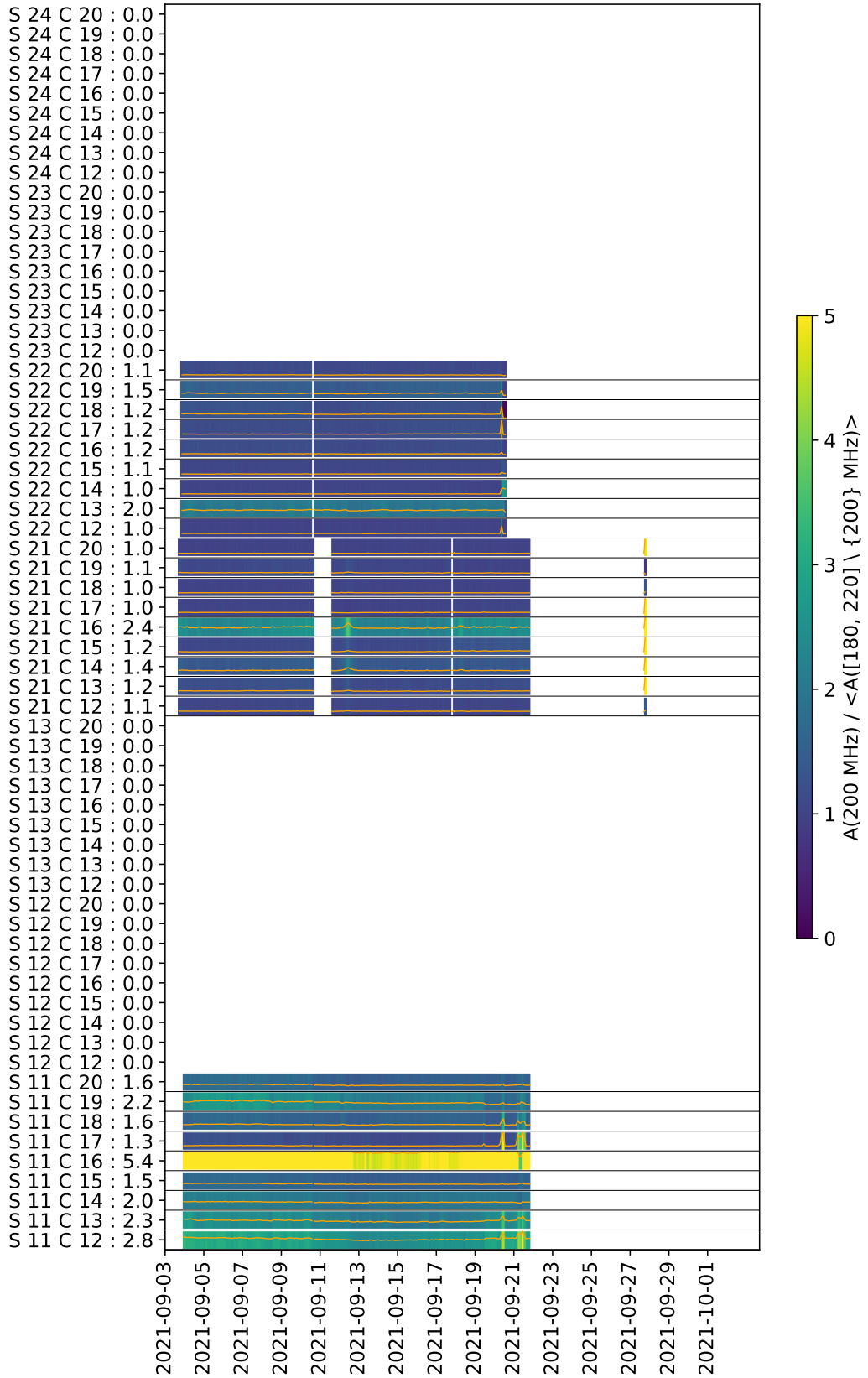

Figure 13: The normalised spectrum is averaged over each run. The orange lines and the background colours show the ratio of spectral amplitude at 200 MHz divided by the average of neighbouring frequencies 180 - 220 MHz except 200 MHz. In the y axis "S" stands for station, "C" for channel, the value after the colon is the mean of the shown ratio.

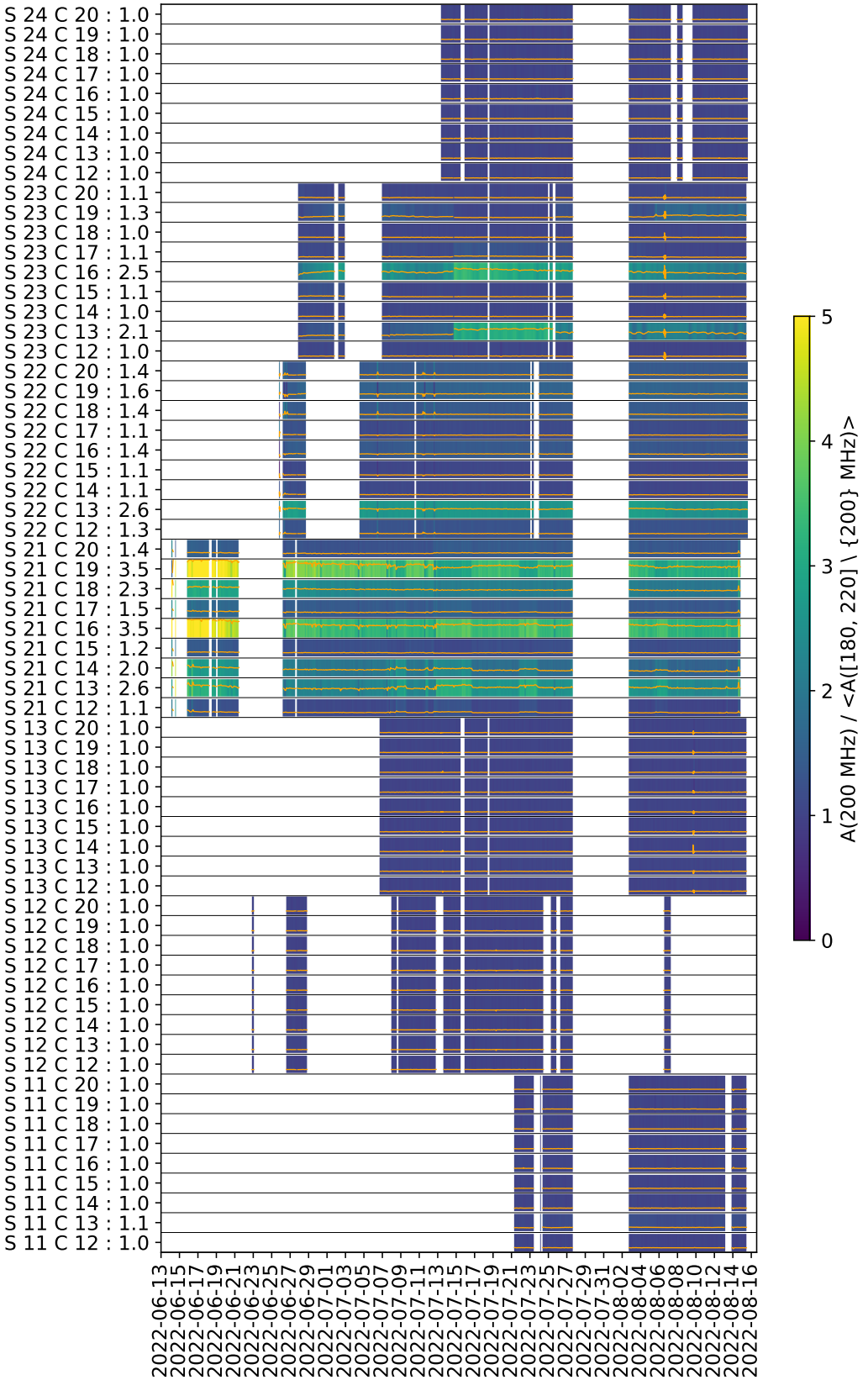

Figure 14: The normalised spectrum is averaged over each run. The orange lines and the background colours show the ratio of spectral amplitude at 200 MHz divided by the average of neighbouring frequencies 180 - 220 MHz except 200 MHz. In the y axis "S" stands for station, "C" for channel, the value after the colon is the mean of the shown ratio.

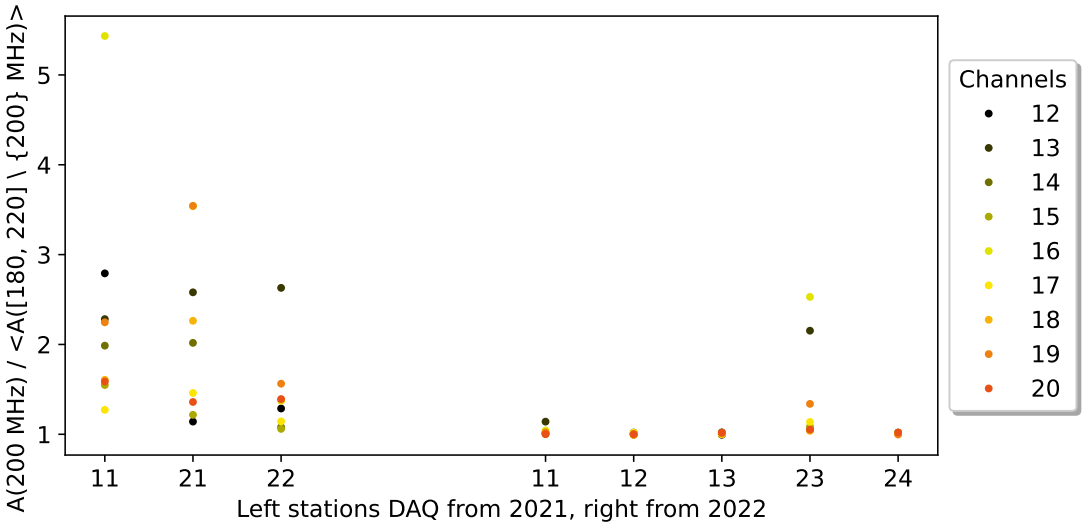

Figure 15: The plot shows the mean ratio of spectral amplitude at 200 MHz divided by the average of neighbouring frequencies 180 - 220 MHz except 200 MHz. The stations 11, 21 and 22 started with a DAQ box from 2021. The DAQ box of station 11 was changed to a new DAQ box model from 2022. Since the beginning of the measurement, stations 12, 13, 23 and 24 have installed a DAQ box from 2022.

Possible source: anthropogenic (e. g. Summit station)
Since the signal is discrete and always visible, it is probably anthropogenic. However it is unlikely that the signal is emitted by one single source like the Summit Station. The reason for this is that the strength of the signal changes over time but those changes are limited to single stations, sometimes even to single channels. Examples for the former are changes in station 21 in 2022 which are not visible in station 22 and 23. Station 23 around 25.07.2022 is an example for the latter: channel 13 and 16 detect the signal; after the measurement pause (white space in the plot) channel 13 detects a lower strength whereas channel 16 remains at its before measured strength. That implies that the signal has a short transmission range. Therefore, there must be several signal sources at different locations to achieve this signal pattern.
The Summit station does not fulfil these requirements.

Possible source: station induced
The stations as the signal source can satisfy the above requested source characteristics. There is for example the possibility that solar panels or battery packs emit the signal.
In July 2023, a search was made for the 200 MHz signal at the Summit station using handheld antennas. However, they did not detect any 200 MHz signal which is possible if the signal is station induced.

Possible explanation: DAQ box
Especially the development of the signal in station 11 is interesting: in contrast to 2022, the signal is detected by station 11 in 2021. To get a hint for this change, one has to look at the data acquisition box (short: DAQ box) which is part of the detector setup. There are two DAQ box models: one from 2021 and one from 2022. When station 11, 21 and 22 were built, they have been

equipped with the old 2021 DAQ box model. Whereas the other stations (12, 13, 23, 24) have the new model from 2022. Station 11 is special regarding that the old DAQ box was replaced with the new model between the two measurement time scopes 2021 and 2022. All other stations kept their (old) DAQ box.

In general electronics emit radio waves because of alternating current. Thus, it is important to shield the electronics such that self emitted noise is not measured. It may be that the new DAQ box model is better protected against emitting electronic noise compared to the old model and therefore station 11 does not measure the 200 MHz signal with the new DAQ model. But this assumes that the signal is induced by a detector component.

Figure 15 examines the connection between DAQ box model and 200 MHz signal strength. It shows the average signal ratio of 200 MHz to neighbouring frequencies plotted against DAQ model. Since station 11 had both DAQ models, it is shown twice. It is visible that all old DAQ models see the signal whereas only one new DAQ model (station 23) detects the signal. Station 23 does not falsify the above named guess because the shielding of the DAQ box of station 23 might be damaged.
If the DAQ box leaked the 200 MHz signal by improper shielding, the signal would mainly reach the upward facing channels, because they are closer to the DAQ box than the downward facing channels.

But even if the signal is station induced and measured because of not sufficient shielding, there is still one question:

1. Why does the signal strength change for a few days and why does channel 13 and 19 of station 21 detect in 2021 nothing but in 2022 a decent amount?
   Self induced signals should be, besides temperature influence, more or less constant over time.

Therefore, the investigation of the 200 MHz signal is still ongoing.

In the next two sections two types of measurement failures are shown.

4.3 Side maximas

Sometimes a real existing discrete signal peak will cause several discrete side peaks in the measured spectrum. The side peaks often consist of a single excited frequency, sometimes there are another two smaller peaks to a side peak in the high frequency area (> 700 MHz) with an offset of 3 MHz or 6 MHz (see right plot of figure 16).

These side peaks are not caused by radio waves since they are also visible in frequency regions where the amplifier does not work. If the amplifier does not work, it is impossible that a radio wave creates a signal.
Therefore, this is a measurement artefact. It is observed in upward and downward facing channels and has already been seen in station 11, 12, 13, 21 and 23. It is not proven that station 22 and 24 are not affected by this artefact.

The artefact has different appearances, but all have in common that the frequency difference between the real existing peak and the side peaks is in general a multiple of *exactly* 25.0 MHz.
As the distance from the existing peak increases, the spectral amplitude of the side peaks decreases. However, in regions where the amplifier does not operate, the amplitude hardly decreases but it may fluctuate. The spectral amplitude of the side peaks is probably always smaller than that of the existing peak.

Side peaks are often visible for more than one forced trigger event but not longer than the real existing peak. For example, the radio signal from the weather balloon causes side maximas which are visible for the entire flight of about two hours. The left plot of figure 16 shows the temporal development of the spectral amplitude of (discrete) side peaks caused by the weather balloon. The spectral amplitude is scaled such that the maximal amplitude is one. Moreover the moving average of the side peaks is plotted. The 403 MHz amplitude curve shows that the balloon starts shortly before 0 h. Followed by high signal emission which rapidly decreases after a short time. Around 1.5 h another increase in signal emission is detected after which the data transmission stops. It is visible that the amplitude of side peaks is besides statistical fluctuations direct proportional to that of the real peak.

It is interesting that even in the laboratory measurement are small peaks visible at frequencies $800.0 + 25.0 \cdot k\,\text{MHz}$ with $k \in \mathbb{N}_0$ up to 1.6 GHz. The peaks have an height of about 5 - 10 % of the surrounding frequencies amplitude which is more or less constant. Maybe they have a similar origin as the side peaks.

Before one can analyse the side peaks, one must find them. A small simple algorithm has been created to search for the side peaks. It loads a forced trigger event and the python module *scipy* searches for peaks in the spectrum. To reduce considerably the false-positive rate, the found peaks must fulfil strict conditions since most of the peaks are created by noise or statistical fluctuations. Of course some or maybe a lot of side peak occurrences will not be found due to the strict conditions but they are necessary to get a manageable amount of results.

In figure 17 are four examples of side peaks shown. The spectrum of one single forced trigger

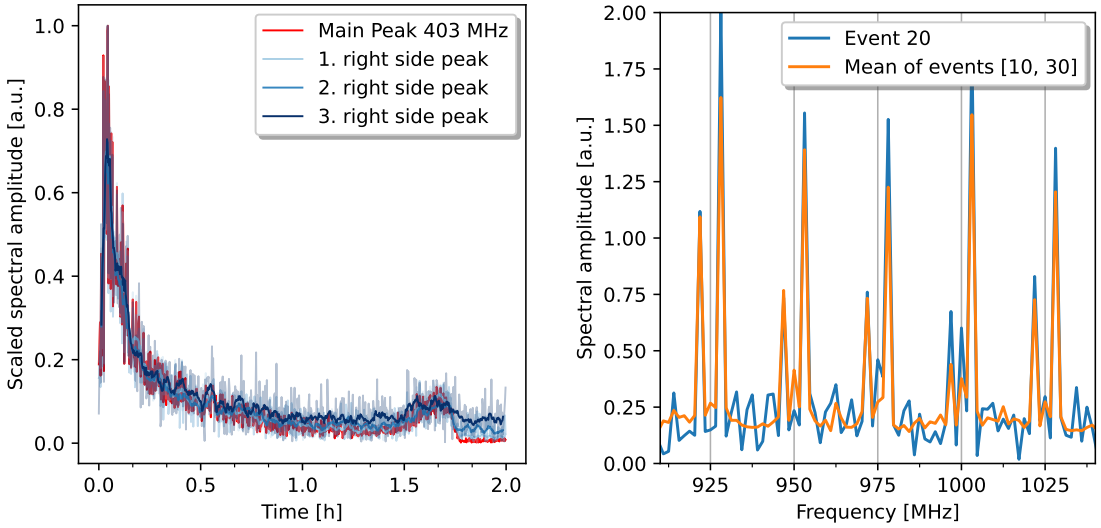

Figure 16: An example of side peaks is shown. The spectrum of channel 13 of station 21 (run 1441) is displayed while the weather balloon emits the 403 MHz signal. In the left plot the temporal development of side maximas at frequencies 428, 453, 478 MHz can be seen. The opaque lines are moving averages of ten forced trigger events. For better comparison, each spectral amplitude line is scaled such that the maximal value is one. The right plot shows the spectral amplitude of channel 13 of one event (blue line) and the average over 20 events (orange).

event is used to gain these plots. There are several interesting aspects in the diagrams:

1. The frequencies with maximal spectral amplitude are well known:
   123 MHz (air traffic), 144 MHz (probably Walkie-Talkie)[1], 152 MHz (Walkie-Talkie) and 403 MHz (weather balloon)

2. Two different real existing peaks may create at the same time side maximas (lower left plot)

3. There are several different appearances of this artefact:
   The side peaks may be everywhere (top left plot), only located near the real peak (top right plot), around real peak and above ~700 MHz (lower left plot) or around real peak with two single side peaks in the high frequency range (lower right plot)

In the lower right plot you can see two single side maximas in the high frequency range. They are at 806 MHz and at 1197 MHz. The first has twice the frequency of the real peak (403 MHz), the second is off by 6 MHz such that the difference would be a multiple of 25 MHz.

One can conclude that there are many different appearances of the side maximas. Regarding figure 17 one might think that side maximas occur when the spectral amplitude exceeds a (high) threshold. However, the figure is not representative as it is much easier to find high amplitude peaks than low amplitude peaks especially in an automated process.

[1] According to [11] the Walkie-Talkie frequencies are at 136 MHz and in 151 - 157 MHz. Specific upward and downward facing channels of station 21 detect a signal at 144 MHz which changes between measurement time scopes. That signal has not a specific shape. Therefore, periodicity is not defined and it is not covered in this thesis. Regarding, that it is in between Walkie-Talkie frequencies, it is assumed that it has a similar origin.

Possible sources
Since the side maximas are visible as well as in regions where the amplifier does not work, it is clear that this artefact is caused by detector components. There are two ideas how the side peaks may be created:

Intermodulation is a process which creates new signals, if at least two different real signals are measured in a non-linear measurement setup. The newly created frequency f is a linear combination of the measured frequencies f_i: $f = k_1 \cdot f_1 + k_2 \cdot f_2$, $k_i \in \mathbb{N}$. There can be a wide range of passive and active detector component types creating non-linearities in the measurement setup (e. g. [13]). It is difficult to estimate the probability that intermodulation is the source of the side maxima.

The frequency difference of 25.0 MHz of the side maximas is interesting regarding that the sampling rate of voltages is 3.2 GHz. If the complete spectrum is populated by side peaks, one would have 128 ($= 3.2\,\text{GHz}/25\,\text{MHz}$) side peaks (ignoring Nyquist-Shannon sampling theorem increases spectrum to 3.2 GHz). 128 is as well as the count of samples, which save one voltage value, in a digitizer block. This might be a hint for the source but it is unknown how digitizer blocks can create this artefact. There is a known issue where digitizer blocks have a constant voltage offset. This offset has not been removed by default from the 2021 data set since it is normally unproblematic. Therefore, it is tried to remove the offset but it does not solve the side maxima artefact.

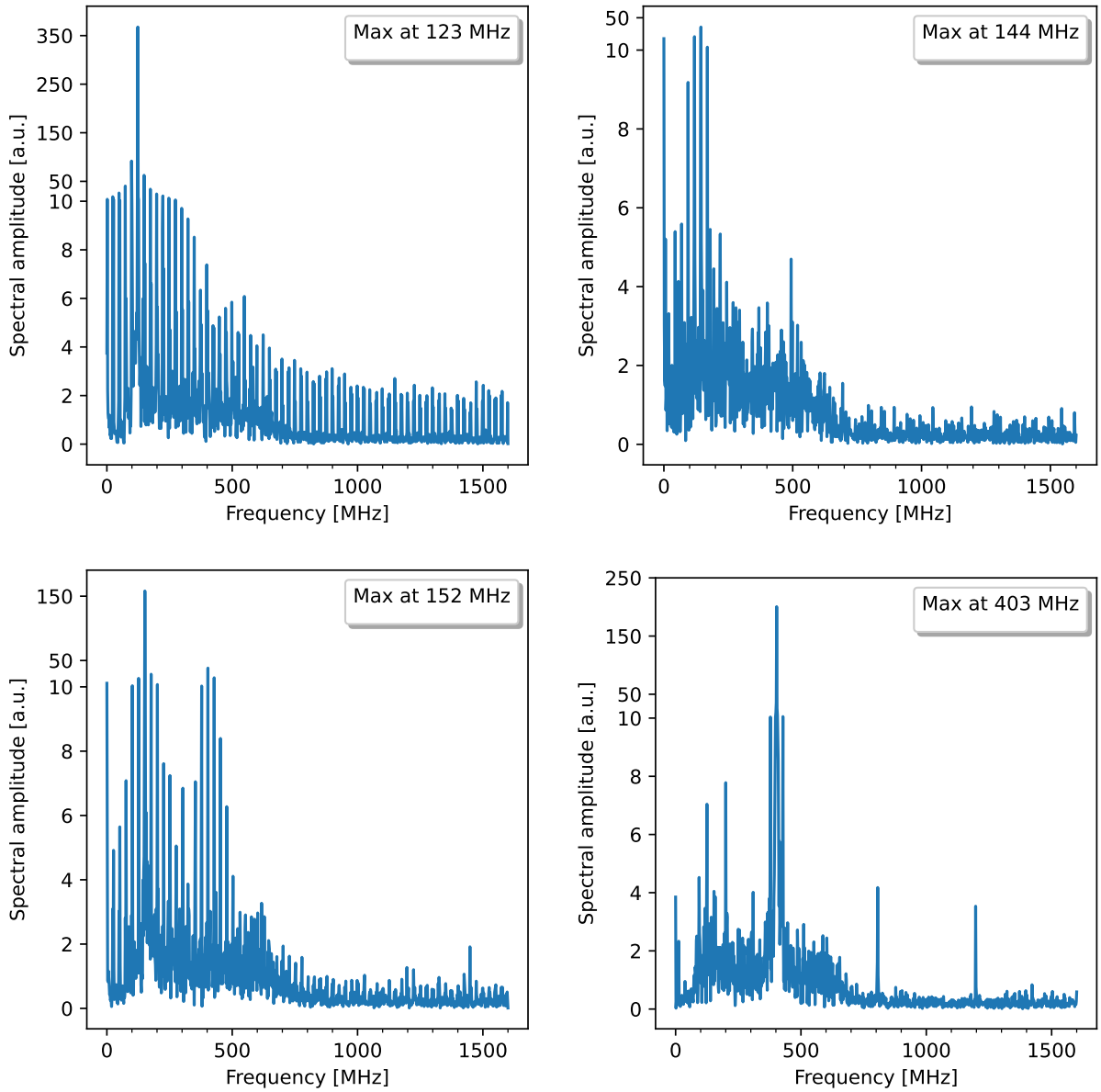

Figure 17: Examples of side peaks are shown. The diagrams display the spectrum of specific channels of one forced trigger event from September 2021. The left two plots show channel 19 and 13 of station 21. The right plots show channel 17 of station 21. Watch out for the y axis values in the plots since they might jump to high regions.
Used runs/events from left to right, up to down: 864/344, 765/669, 1536/106, 1081/9

4.4 Spectra jumps

In September 2021, something special happened in station 21:
the high frequency part of the spectrum is shifted.

In the available data set of station 21 in September 2021 are two measurement pauses (see table below) which divide the data set in thirds (before break one, between break one and two, after break two). These breaks are usually nothing special, they happen regularly and in other stations too.

Pause index	Start (UTC)	End (UTC)	Duration
1	10.09, 14:40	11.09, 12:50	22.3 h
2	17.09, 17:20	17.09, 19:10	1.8 h

Table 4: The data of the two measurement pauses of station 21 in 03.09 - 21.09.2021 is shown. The accuracy of the dates and times given is 10 minutes.

But if the mean spectrum of each data set third is calculated, in some, a higher spectral amplitude is observed in channel 13, 17 and 19 almost across all relevant frequencies. This is not a statistical fluctuation because each data third lasts at least 4 days. The averaged spectrums can be seen in the left part of figure 18.

It is not just the high frequency range (550 - 800 MHz) that looks strange. On closer inspection, channel 13 and 19 see a new peak at 459.375 MHz after measurement pauses, and the peak at 484.375 MHz has disappeared. The channel 13 and 19 are the only channels which detect this new peak every other upward facing channel sees the old peak at 484 MHz. By analysing the new peak at 459 MHz with the time binning method shows that the signal shape of the 459 MHz line is exactly the same as of the original 484 MHz signal. It looks like the peak has shifted by 25.000 MHz to lower frequencies. However, not all peaks are shifted. For example the 200 MHz and the ~280 MHz signals are correct.
Interesting is that in channel 13 and 17 the frequency shift starts after the first break. In channel 19 the shift is not until the second break.

Because of the above observation, it is tried to shift the average spectrum of channel 13 and 19 above 320 MHz by 25 MHz to higher frequencies. This shift changes the high frequency range, therefore it is necessary to normalise the spectrum again. The result is shown in the right plots of figure 18. It is clearly visible that the result looks a lot more like it is expected. Thus, it is as well as done at channel 17 but there it turns out that two shifts are required to get a spectrum as expected. The first one shifts all frequencies higher than 350 MHz by 50 MHz to higher frequencies. The second shifts all frequencies in between 210 MHz and 350 MHz by 25 MHz.

It must be mentioned that the shift starting frequencies 210 & 320 MHz are estimated with the eye to gain a good congruency; they and the shift itself have no physical origin. However, it has decent success.

It is visible that the shift does not solve all problems: the spectrum in the 300 - 400 MHz range is still not as it should be. In channel 13 and 17 the green and orange curve do not fit on the blue curve in this frequency region. Therefore, it can be assumed that the frequency shift artefact is more complicated than only shift by 25 MHz.

The reason for this issue is unknown.

Since this artefact has an huge effect on the measured spectrum, it is important to detect this artefact. As long as the signal source of the 484 MHz line has not been eliminated, one can use this signal to search for the frequency shift artefact in upward facing channels by controlling that the signal is visible at the correct frequency. Regarding channel 17, it is probably best to look at the average spectral amplitude of frequencies ~500 MHz to search for the artefact in downward facing channels. If the average over several runs/days changes significantly, one might have this artefact. Besides the named occurrences in the channels 13, 17 and 19 in station 21 in 09.2021, no further measurements have been found with this type of artefact (in the available data set).

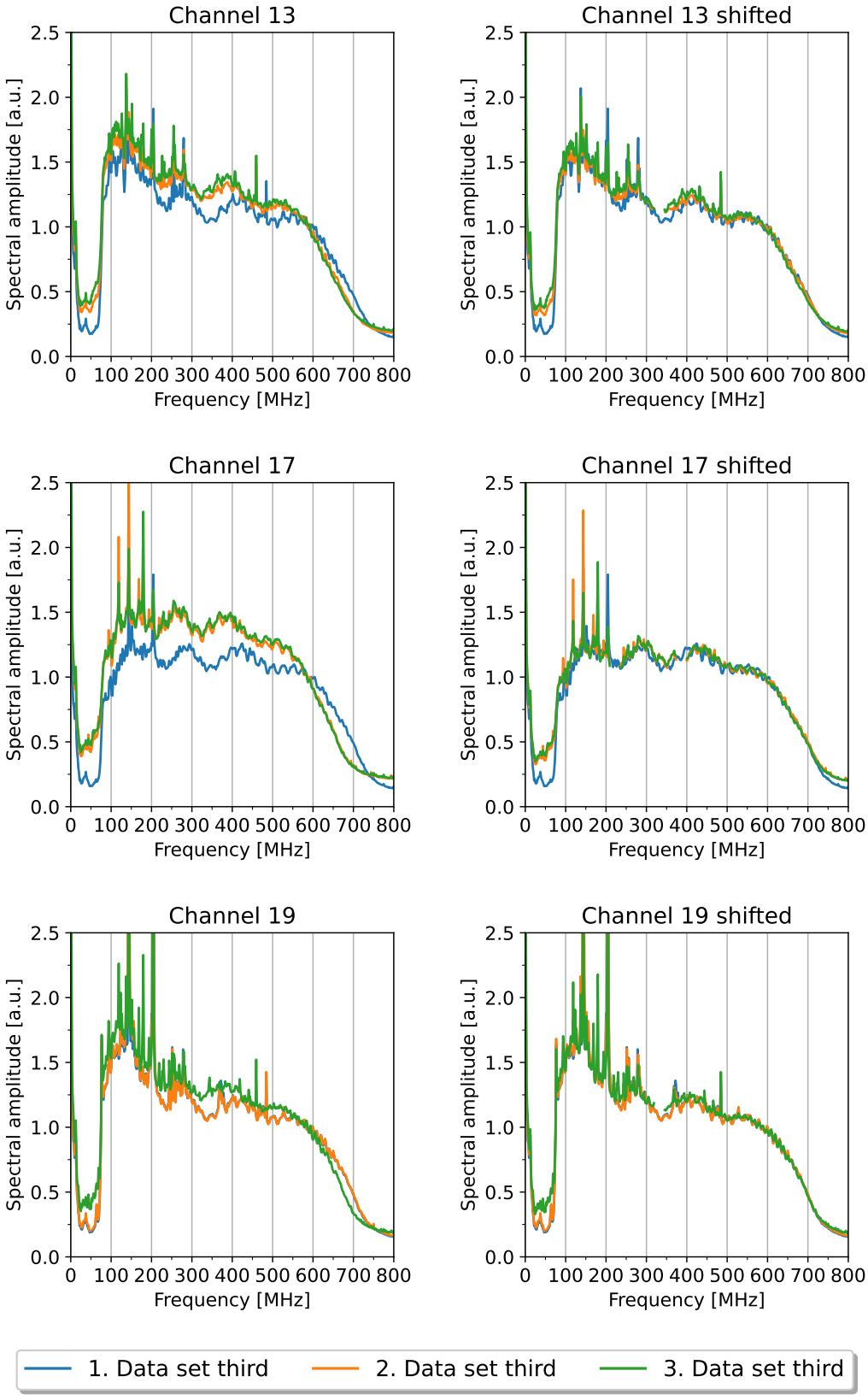

Figure 18: The spectral shift is illustrated. Data has been taken from station 21 from 2021. The data set is divided into thirds by two measurement pauses. The spectrum of each data third is averaged and plotted on the left. The blue spectrum is before the first break, orange curve is between first and second pause, green curve is after the second break. The plots on the right show the shifted averaged spectrum.

5 Galactic signal noise

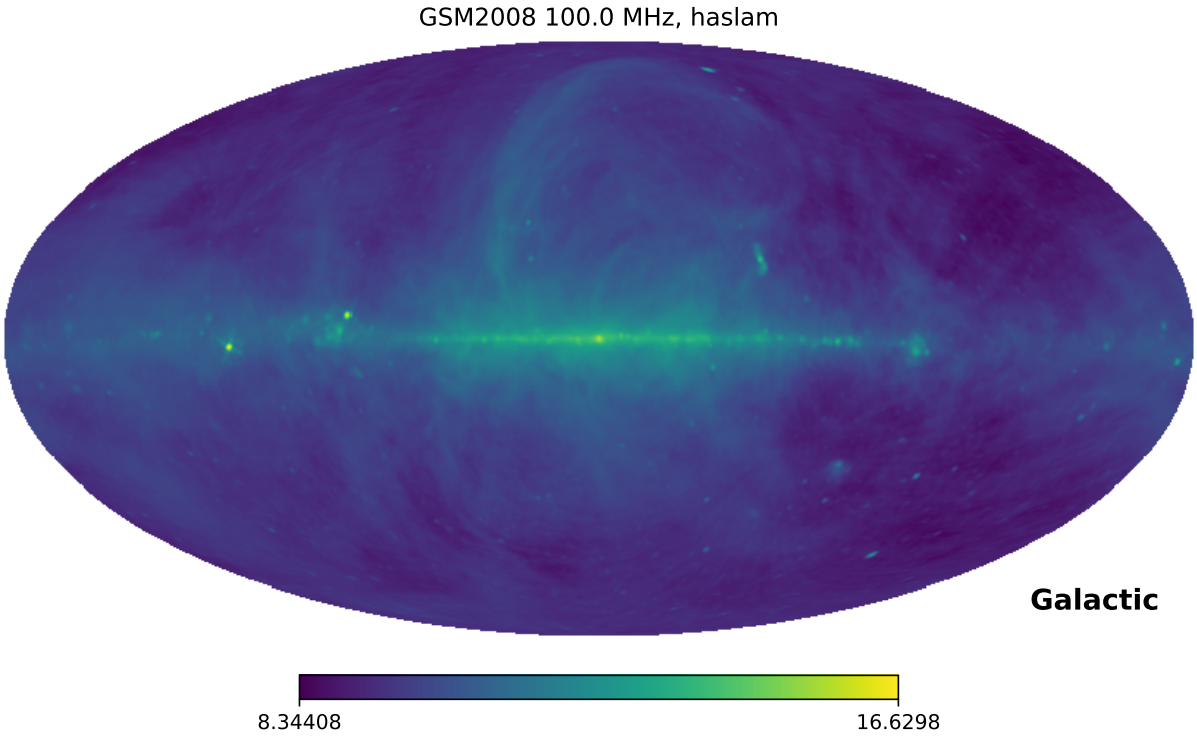

Figure 19: The search for the galactic signal uses this global sky model [14]. It is provided by the python library PyGDSM. The Galactic center and disk are visible in the middle of the image.

It has already been mentioned that in figure 6 we see an excess of signals in the 80 - 200 MHz range in the upward facing channels. It is thought that these signals come from the Milky Way. This chapter analyses this signal to provide evidence for this assumption.

5.1 Astronomical sources of radio signals

There is a wide range of processes creating cosmic radio signals. Many are mentioned in ref. [15].

1. Black body radiation:
   Every object with a positive temperature emits electromagnetic radiation according to the Plancks-Law.

2. Synchrotron radiation:
   If electrical charges are moving in a magnetic field, they are redirected. Conservation of momentum requires that photons must be radiated. The emitted spectrum is continously. It is the main radiation source of the Galactic disk. Galaxies with extraordinary high synchrotron radiation are called radio-galaxy. For example Cygnus A is a radio galaxy which can be detected by RNO-G.

3. Supernova remnants and pulsars

5.2 Expected frequency range of galactic signals

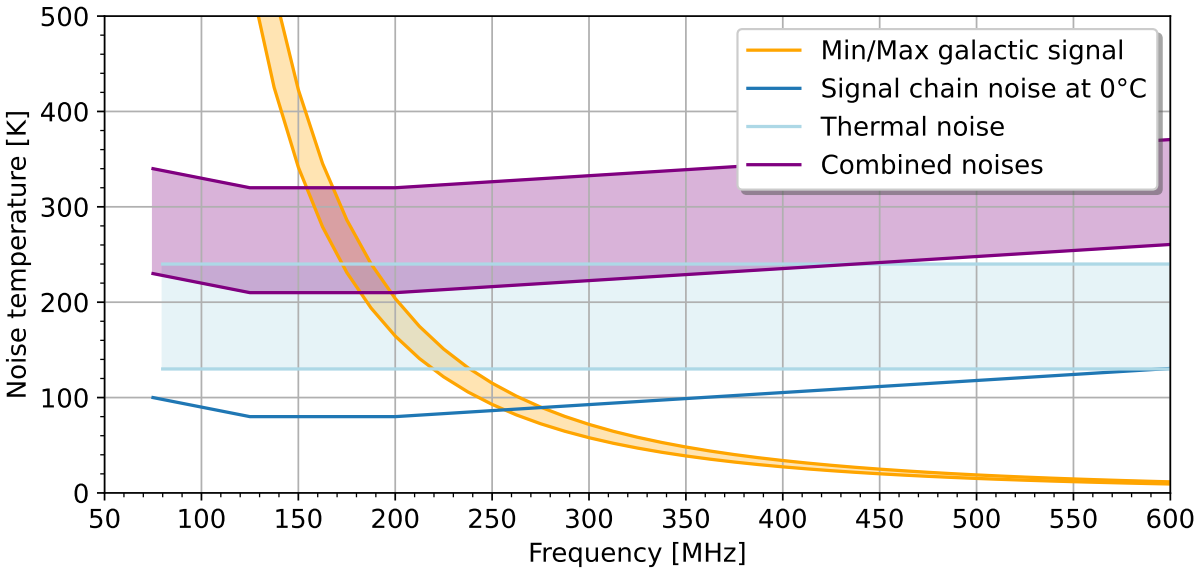

Figure 20: The contributions to an idealized frequency spectrum are shown. The result data from [16] is approximated to get the signal chain noise curve. Thermal noise lies in 120 K - 240 K ([16]). Galactic signal is the average temperature of the sky seen by station 23. The sky temperature is taken from the global sky model ([14]).

Before searching for the Milky Way signal, it is necessary to determine the frequency ranges in which the signal is expected. There are mainly three types of signals which must be considered. The first one is the Milky Way signal itself. Secondly, there are background signals like Walkie-Talkie and air traffic frequencies which would make it harder to detect the Milky-Way signal in that frequency region. And lastly, there is thermal and electronic signal chain noise. To detect the Milky Way signal, a frequency region must be found which hardly has background signals and the sum of noise signals must be smaller than the galactic signal.

From section 2.2 is known that the galactic signal can be specified by the sky temperature. Furthermore, it is possible to describe thermal and chain noise with a temperature. Thus, one can use the temperatures to compare signals against each other instead of signal strengths which are a lot harder to calculate.

The global sky model taken from ref. [14] provides the sky temperature. The data of that model can be accessed by the python library PyGDSM. In figure 19 one can see that model. This sky model contains a few extra-galactic radio sources, but mainly the diffuse radio emission from the Milky Way in the 10 MHz - 100 GHz frequency range.
The galactic signal can be calculated by averaging the sky temperature seen by the detector (here: station 23). Because the Earth rotates, the Milky Way rotates in the sky. Thus, the average temperature oscillates in LST. The galactic signal is limited by the minimum (~6 h LST) and maximum (~18 h LST) of that oscillation. These bounds are used for this analysis.

Moreover, a research of the strength of signal chain noise has been done by [16]. They showed that it has a temperature of about 90 K in the 100 MHz region. An approximation of their result data is used for this investigation.

To estimate the thermal noise, one must know that it is created by black body radiation. Therefore, thermal noise depends mainly on the temperature of the surrounding and its emissivity. For downward facing antennas, it is a good approximation to assume that the antenna is completely surrounded by an infinite amount of ice which radiates. The ice around downward facing antennas has a temperature of about $-30\,^{\circ}\mathrm{C}$ (~240 K) and an emissivity of $0.97 - 0.99 \approx 1$ ([17]). Thus thermal noise is at ~240 K. Whereas upward facing antennas have only two to three meters ice above them. Therefore, they receive only 50 % of the thermal noise of the ice seen by downward facing antennas which is 120 K. That value agrees with measurements ([16]). Thus, the thermal noise ranges from 120 K to 240 K.

The galactic signal and noises are shown in figure 20. The figure shows that the galactic signal is visible up to ~150 MHz. Regarding that the amplifier starts to amplify at 80 MHz and above 120 MHz are many noise signals from air traffic and Walkie-Talkies, the 80 - 120 MHz frequency range is chosen for the search.

5.3 Experimental data

We want to apply the time binning to excess data in the 80 - 120 MHz frequency range, in order to detect the galactic signal if there is one. Therefore the mean (lab-)excess in that frequency region is calculated. After that the time binning as described in section 3.4 is performed. This has been done for all stations and measurement time scopes. All upward facing channels of every station see a sinusoidal function. The visibility of that signal varies strongly between channels and stations. Station 23 sees that function most clearly. It is shown in figure 21.

Comparing the minimum, maximum and shape of the sinusoidal function of different measurement time scopes, shows that the function rises and falls with sidereal time and the shape does not change between different time scopes. Furthermore the lower diagram of figure 21 shows that mainly the upward facing channels see the signal. Whereas the downward facing channels are more or less constant.

For the reasons given above, it can be concluded that the source of the sinusoidal signal is located outside the solar system. Therefore, the source is probably an astronomical object. To find evidence that the sinusoidal signal comes from the Milky Way, the next section simulates the expected signal from the Milky Way.

Moreover the sinusoidal signal is visible in all frequencies in 80 - 120 MHz. The signal strength decreases with increasing frequency.

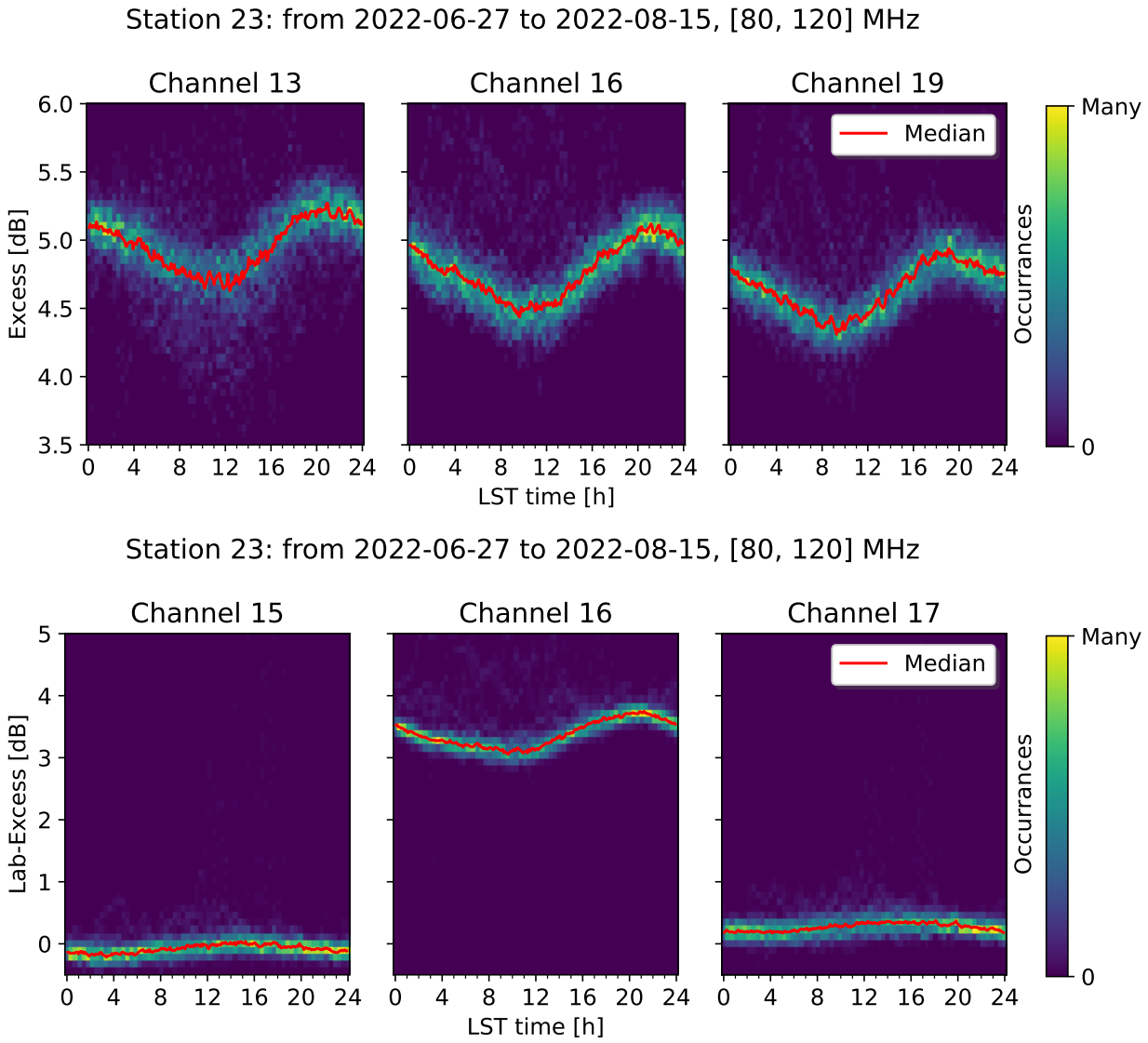

Figure 21: The average of the frequency range 80 - 120 MHz is shown as a function of the local sidereal time binned excess (top) and lab-excess (bottom). Data is taken from station 23 from 2022-06-27 to 2022-08-15. The bluish quads show the underlying distribution of the result excess values. The red curve is the median of that distribution. Channel 13, 16 and 19 are upward facing whereas 15 and 17 are downward facing ones.

5.4 Simple simulation model of galactic signals

In this section, a simple simulation model for the galactic signal strength is studied. From section 2.2 is known that the received power in a channel can be calculated from the sky temperature via spatial integral. Aim is to implement that formula (see equation 9) in the simulation, thus it is discretised with the simple approximation $\int_a^b f(x)\,\mathrm{d}x \approx (b-a)\cdot f(a)$.

$$P_{\mathrm{Sky}}(t,f) = \frac{Z_0}{Z_L}\frac{k_B}{c^2}\int_\Omega f^2\, T_{\mathrm{Sky}}(t,f,\alpha,\delta)\cdot|H_i(f,\alpha,\delta)|^2\,\mathrm{d}\Omega \approx \tag{9}$$

$$\approx \frac{Z_0}{Z_L}\frac{k_B}{c^2}\cdot f^2\cdot\sum_j T_{\mathrm{Sky}}(t,f,\alpha_j,\delta_j)\cdot\left|H_i(f,\alpha_j,\delta_j)\right|^2\cdot|\mathrm{d}A| \tag{10}$$

Since we are integrating over two angles α and δ, $\mathrm{d}A$ is given by $|\mathrm{d}A| = |\sin(\delta)\,\mathrm{d}\alpha\cdot\mathrm{d}\delta| = |\mathrm{d}\alpha|\,|\mathrm{d}(\cos\delta)|$. $\mathrm{d}A$ can be made constant by choosing $\mathrm{d}\alpha$ and $\mathrm{d}(\cos\delta)$ constant. Moreover, the spectral amplitude S is proportional to the measured voltage U. Therefore, one can write with $P = U\cdot I = {}^{U^2}/_{R}$:

$$S \propto U = \sqrt{P\cdot R} = \underbrace{\sqrt{\frac{Z_0}{Z_L}\frac{k_B}{c^2}|\mathrm{d}A|\cdot R}}_{=:\tilde{c}=\text{const.}}\cdot f\cdot\sqrt{\sum_j T_{\mathrm{Sky}}(t,f,\alpha_j,\delta_j)\cdot\left|H_i(f,\alpha_j,\delta_j)\right|^2} \tag{11}$$

$$\propto f\cdot\sqrt{\sum_j T_{\mathrm{Sky}}(t,f,\alpha_j,\delta_j)\cdot\left|H_i(f,\alpha_j,\delta_j)\right|^2} \tag{12}$$

The sky temperature is provided by the global sky model ([14]).

In the previous section it was stated that the sinusoidal signal is clearly visible in every frequency at the 80 - 120 MHz range. Thus, it is sufficient to simulate the galactic signal just for 100 MHz. By simulating only one frequency, the frequency-dependent voltage gain can be ignored as it is a proportional constant. As well as the frequency factor f can be ignored.

In contrast to reality this simulation model does not consider any thermal noise.

The aim of the simulation is to compare it's results with experimental data. Since section 5.3 obtained a likely galactic signal shape over a sidereal day, it makes sense to simulate the signal over the course of a day. On top of that the simulated signal must/will be periodic in sidereal time.

The simulation is structured as follows:

1. Select time point and frequency for which galactic signal strength is requested.
2. Determine Ω, which is the part of the sky seen by the detector.
   Significant signal strength from below the horizon is not expected, and simulating signal propagation in ice is complicated, so it is ignored.
3. Calculate for each sky position (α_j,δ_j) the product of sky temperature and channel sensitivity:

The global sky model is discretized. To increase precision the temperature values are interpolated with splines. An example is shown in figure 22.

4. Calculate the rest of equation 12:
   The result of this step can be seen in the two left plots of figure 23.

5. Compute the excess of the upward facing channels:
   The proportional constants are not easy to determine. By calculating the excess, these constant factors are eliminated. The result is presented in the right diagram of figure 23.

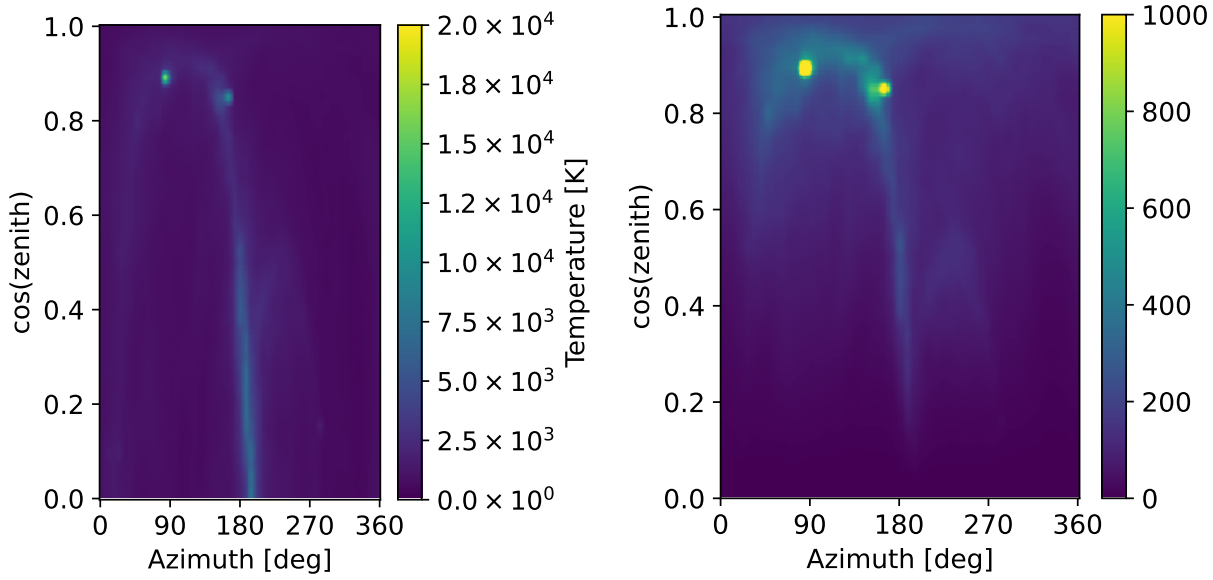

Figure 22: The left plot shows the sky section seen by the detector on 2022-07-04 1 h UTC. The sky temperature, which is indicated by the colour, is interpolated using splines. The right plot shows $T_{\text{Sky}} \cdot H^2$ where H is the sensitivity of channel 13 station 23 which can be seen in figure 24.

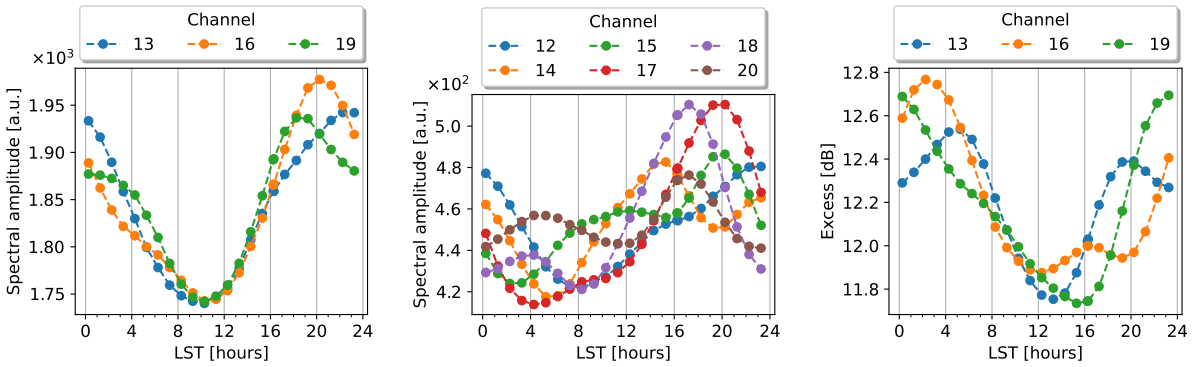

Figure 23: The diagrams show the simulation result at a frequency of 100 MHz. The simulated spectral amplitude is shown in the two left plots (upward & downward facing channels). With the simulated spectra the excess is calculated which is visible in the right diagram.

Several aspects emerge from the simulation results for 100 MHz in figure 23. Figure 21 is used for the following comparisons.

1. The simulated excess values are very different from the measured ones. For example, the time positions of the minimum and maximum are different and the simulation predicts more complicated shapes such as two maximas (instead of just one) in channel 13 and 16. On top of that the excess values are much too high.

2. Comparing the shape of the simulated spectral signal in upward facing channels with the lab-excess of real measurements (cf. lower figure 21, channel 16), one sees that the time positions of minimum and maximum lab-excess/spectra values are in the same time region. Only the maximum of channel 13 has a small deviation of about two hours (cf. table 5). The comparison makes sense since the lab-excess $E_{\mathrm{L}}(f,\ t) = 20\log_{10}(S(f,\ t)/S_{\mathrm{L}}(f))$ has the same extremum time positions as $S(f,\ t)$ which is proportional to the simulated spectral amplitude.
   The time points in the table are determined as follows:
   Due to limited time resolution in the simulation, the three nearest data points to the extremum are taken. The median lab-excess is calculated from the time-binned measurement data in the 80 - 120 MHz range. Since the median varies a little bit, the time range comprises all lab-excess data points which have an absolute excess difference of less than 0.05 dB to the exact extreme value.

Station 23 Channel	Minimum / Maximum position [h] Simulation	Measurement
13	[9.3, 11.3] / [21.2, 23.3]	[7.7, 11.8] / [17.5, 21.1]
16	[9.3, 11.3] / [19.2, 21.2]	[9.1, 11.5] / [19.6, 21.8]
19	[9.3, 11.3] / [17.2, 19.2]	[7.6, 10.1] / [16.5, 19.3]

Table 5: The LST time position of minimum and maximum of the simulation at 100 MHz and measured median lab-excess in 80 - 120 MHz range of each upward facing channel.

3. The simulated downward facing channels see as well as a sinusoidal function whereas they see no galactic signal in real measurements (see figure 21) since the lab-excess is constant. In comparison to the upward facing channels the downward facing ones detect in the simulation 4-times less signal strength, as one would expect.

Because of the second mentioned point and the congruence of simulated spectral amplitude and measured lab-excess it can be stated that the measured sinusoidal signal is most probably created by the Milky Way.

Furthermore thermal noise considerations are important, since thermal noise would dominate the signal in downward facing channels. Therefore, the signal of these channels would be approximately constant and the shape of the simulated excess would better match the spectral amplitude of the upward facing channels, and thus match as well as the measured excess.

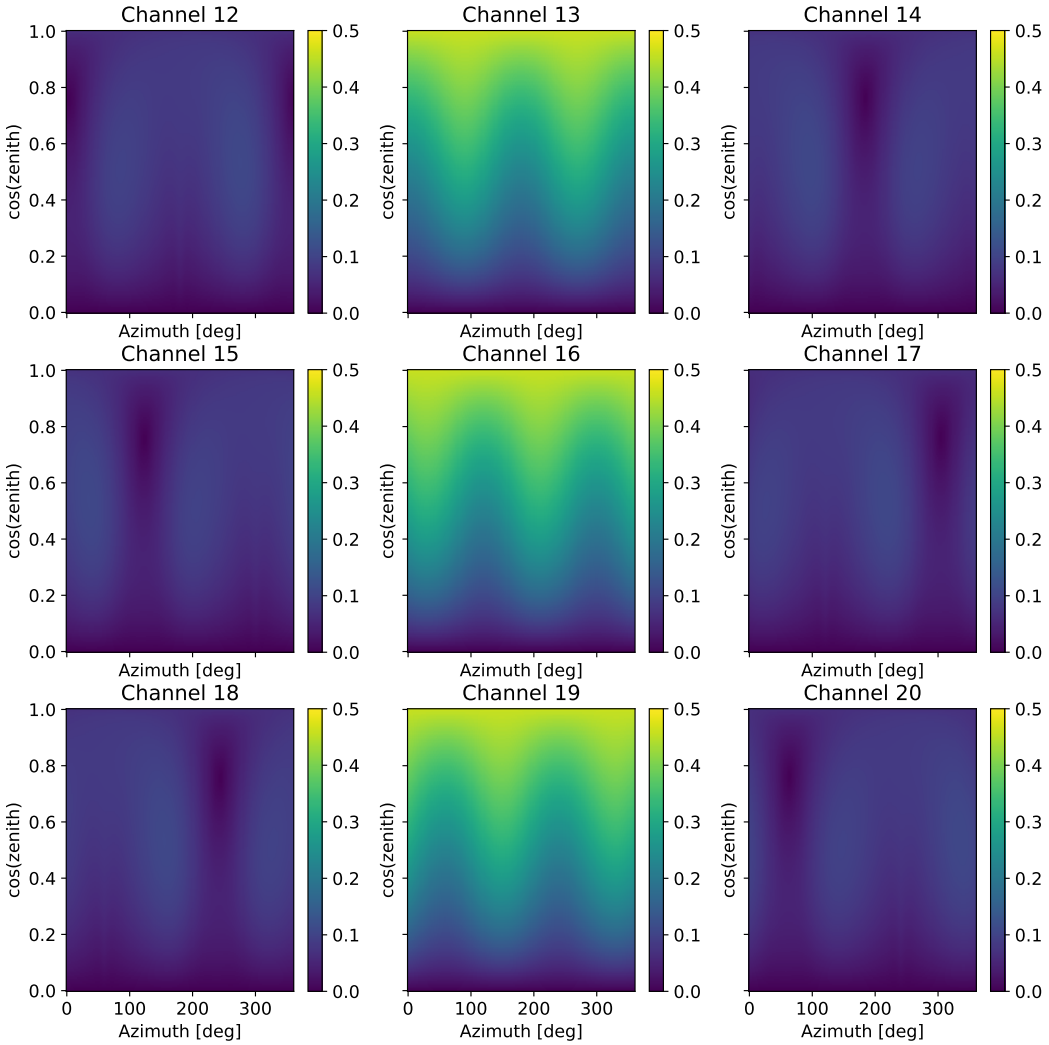

Figure 24: The spatial dependency of channel sensitivity at station 23 is shown.

6 Discussion and conclusion

In this thesis periodically occurring background signals in RNO-G at Summit Station in Greenland were investigated to support the search for noise signal sources. Furthermore, two new measurement artefacts were found and analysed.

RNO-G detects radio waves by measuring voltages applied to antennas at a sampling rate of 3.2 GHz. Because of limited hardware disk space, it is necessary to use triggers to decide whether a voltage trace should be saved or dismissed. For this thesis, only data saved by forced triggers is used. A forced trigger saves the current voltage trace every 10 s. The saved voltage values can be transformed into a spectrum by using the Fast-Fourier-Transform (FFT). The resulting spectrum shows the radio frequencies that are present. The detector setup is sensitive in the frequency region 80 - ~700 MHz.
As the voltage values are not gain-calibrated, a normalisation procedure is applied to the frequency spectra in order to be able to easily compare the measured spectra values with a laboratory measurement to gain further insight into the radio signals.

Knowing the signal period and the development of a signal over the course of a day is helpful in finding the source of background signals. To gain these information, a time binning procedure is applied to the measured spectra values. This procedure is capable of determining the period of a signal. Moreover, it shows the time development over the course of a day.

With this procedure it was possible to find identical narrow-band semi-circle signals at 243.7 and 248.4 MHz which are periodic in local sidereal time (LST). Their start and end time vary slightly between different detectors. They start at 23 - 02 h and end at 13 - 16 h (LST). Because of the similarities of both signals, it is assumed that the signal sources are similar or even the same. However, the source of this signal is still unknown.

Moreover, there are short range signals at 251.6 and 256.3 MHz which are narrow-band and probably periodic in UTC time. The signal shape varies greatly between detectors. Therefore, it was suggested that both signals are originating near or by the detectors.

Sharp signal peaks were found in the frequency ranges 278 - 285 MHz and 482 - 485 MHz. The peaks have an identical appearance. They last for less than two hours and have a period of $5\,\mathrm{h}\ 58\,\mathrm{m}\ 53 \pm 3\,\mathrm{s}$. Thus, the period of four peaks is by $29 \pm 12\,\mathrm{s}$ lower than a sidereal day. The signal source with this special periodicity is still unknown.

Furthermore, RNO-G detects air traffic communication at ~125 MHz and Walkie-Talkie frequencies at 136 MHz and 151 - 157 MHz. Data transmission from a weather balloon is also detected at 403 MHz.

Moreover, there is a special narrow-band signal at 200 MHz. It is seen by some but not all detectors. It is constant over time, except that its strength increases for a few days and then decreases to its original strength. Further investigations showed that it is probable that the signal is anthropogenic and maybe caused by the detectors themselves. A correlation between visibility of the signal in a detector and an old data-acquisition box model in the detector was investigated. There are reasons for causality. However, there is a big aspect against causality. Therefore, the search for the signal source is still ongoing.

After analysing these background signals, two measurement artefacts were studied.
The first one is called "side maximas". It describes the effect where a real existing signal peak in the spectrum causes side peaks that do not belong to real radio signals. This artefact has different appearances, but all have in common that the frequency difference between real existing peak and the side peaks is in general a multiple of *exactly* 25.0 MHz. The side peaks are observable in high frequency regions where the amplifier of the detector does not amplify anymore. The number of side peaks varies from a few to the entire spectrum is populated with these peaks. Two possible sources of the side peaks were investigated. Namely intermodulation and an issue with the digitizer. However, it was not possible to pinpoint the source of the problem.

The second measurement artefact is called "spectra jumps". In 2021, three detectors had detected incorrect spectra for the time between measurement pauses. After normalisation the incorrect spectras are higher in frequencies smaller ~600 MHz and lower in ~600 - 750 MHz than expected. Moreover, it has been found that the background signal at 484.4 MHz has been shifted by 25.0 MHz to 459.4 MHz. However, background signals at smaller frequencies (< ~300 MHz) are not shifted. Thus, it was tried to solve the issue by reversing the 25 MHz shift in high frequencies. This operation had decent success but could not solve all issues. Since this artefact has a huge impact on measurements, a method was described to find this issue with an automated process. The source of this problem is unknown.

The last section of this thesis analyses the Galactic signal which is visible in the detectors. It starts with a theoretical investigation where the signal is expected to be detectable. For this, the Galactic signal is simulated and compared against background, thermal and signal chain noises. It was concluded that the Galactic signal is clearly observable in the 80 - 120 MHz frequency range. The time binning method showed, that in this frequency region a sinusoidal signal is detected. To find evidence that this signal is originating from the Milky Way, a simple simulation has been created. The theoretically expected signal from the Milky Way was simulated by considering the sky temperature and detector sensitivity. The sky temperature is a measure for the signal strength receiving from a sky position. However, thermal and signal chain noises were ignored. It was shown that the simulated spectral amplitude of the Milky Way is in good coincidence with the experimental sinusoidal signal. Thus it was concluded that the sinusoidal signal in the 80 - 120 MHz frequency range is probably from the Milky Way.

In the future, the search for background signal sources will go on. To support this process, this thesis only used results from the time binning method which averages data over a time span of about 25 minutes. Maybe, one finds helpful or interesting information, if one looks at much smaller time scales.
Regarding the measurement artefacts, it is really important to investigate them further. The sources of the issues are unknown and thus it is currently hardly possible to estimate how often and in which appearances they occur. However, it is necessary to be sure that measured data is correct and not significantly affected by the artefacts.
Furthermore, it is planned for the future to use Galactic calibration at RNO-G. To support this, one would need to extend the Galactic signal simulation by adding additional signal chain characteristics like noises and the amplifier.

References

[1] Imre Bartos and Marek Kowalski. *Multimessenger astronomy*. 2399-2891. IOP Publishing, 2017. ISBN: 978-0-7503-1369-8. DOI: 10.1088/978-0-7503-1369-8. URL: https://dx.doi.org/10.1088/978-0-7503-1369-8.

[2] B. P. Abbott et al., LIGO Scientific, Virgo. "Observation of Gravitational Waves from a Binary Black Hole Merger". In: *Phys. Rev. Lett.* 116.6 (2016), p. 061102. DOI: 10.1103/PhysRevLett.116.061102. arXiv: 1602.03837 [gr-qc].

[3] J. A. Aguilar et al., RNO-G. "Design and Sensitivity of the Radio Neutrino Observatory in Greenland (RNO-G)". In: *JINST* 16.03 (2021). [Erratum: JINST 18, E03001 (2023)], P03025. DOI: 10.1088/1748-0221/16/03/P03025. arXiv: 2010.12279 [astro-ph.IM].

[4] J. A. Aguilar et al. "Triboelectric backgrounds to radio-based polar ultra-high energy neutrino (UHEN) experiments". In: *ASTROPART. PHYS.* 145 (2023), p. 102790. DOI: 10.1016/j.astropartphys.2022.102790.

[5] S. W. Barwick et al., ARIANNA. "Livetime and sensitivity of the ARIANNA Hexagonal Radio Array". In: *34th International Cosmic Ray Conference*. Aug. 2015. arXiv: 1509.00115 [astro-ph.IM].

[6] Max Büsken, Tomáš Fodran, and Tim Huege. "Study of the Uncertainties of the Galactic Radio Background as a Calibration Source for Radio Arrays". In: *PoS* ARENA2022 (2023), p. 031. DOI: 10.22323/1.424.0031. arXiv: 2301.11438 [astro-ph.IM].

[7] Steven Barwick and Christian Glaser. "Radio Detection of High Energy Neutrinos in Ice". In: (Aug. 2022). arXiv: 2208.04971 [astro-ph.IM].

[8] Rogerio de Almeida et al., Pierre Auger. "Absolute calibration and investigation of ageing of the AERA radio detectors". In: *PoS* ARENA2022 (2023), p. 039. DOI: 10.22323/1.424.0039.

[9] RNO-G-wiki. URL: https://radio.uchicago.edu/wiki/index.php/Infrastructure.

[10] Glaser, C., Nelles, A., Plaisier, I. et al. "NuRadioReco: a reconstruction framework for radio neutrino detectors". In: *THE EUROPEAN PHYSICAL JOURNAL C 79, 464* (June 2019). URL: https://doi.org/10.1140/epjc/s10052-019-6971-5.

[11] RNO-G-wiki. *Walkie-Talkie frequencies*. URL: https://radio.uchicago.edu/wiki/images/4/41/Handheld_freqs.pdf.

[12] Di Vruno, F. et al. "Unintended electromagnetic radiation from Starlink satellites detected with LOFAR between 110 and 188 MHz". In: *A&A* 676 (2023), A75. DOI: 10.1051/0004-6361/202346374. URL: https://doi.org/10.1051/0004-6361/202346374.

[13] Murat Eron. "Passive Intermodulation Characteristics". In: *MICROWAVE JOURNAL* (Mar. 2014). URL: https://www.microwavejournal.com/articles/21769-passive-intermodulation-characteristics.

[14] Angélica de Oliveira-Costa et al. "A model of diffuse Galactic radio emission from 10 MHz to 100 GHz". In: *MNRAS* 388.1 (July 2008). Provided by the SAO/NASA Astrophysics Data System, pp. 247–260. DOI: 10.1111/j.1365-2966.2008.13376.x. arXiv: 0802.1525 [astro-ph]. URL: https://ui.adsabs.harvard.edu/abs/2008MNRAS.388..247D.

[15] Clive Dickinson. “Large-Scale Features of the Radio Sky and a Model for Loop I”. In: *GALAXIES* 6.2 (2018). ISSN: 2075-4434. DOI: 10.3390/galaxies6020056. URL: https://www.mdpi.com/2075-4434/6/2/56.

[16] Eric Oberla et al., RNO-G. “Low-Power Radiofrequency Systems for the RNO-G Project”. In: *PoS* ICRC2023 (2023), p. 1171. DOI: 10.22323/1.444.1171.

[17] M. Quinn Brewster. “Thermal Radiative Transfer and Properties”. In: https://books.google.com/books?id=z_anVNTmQLUC&pg=PA56. John Wiley & Sons, Inc., 1992. Chap. Radiative properties of opaque solids, p. 56.

Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe, dass alle Stellen der Arbeit, die wörtlich oder sinngemäß aus anderen Quellen übernommen wurden, als solche kenntlich gemacht sind und dass die Arbeit in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegt wurde.

Erlangen, den 31.August, 2023

Unterschrift