PROCEEDINGS OF SCIENCE

Searching for comic-ray air-showers with RNO-G

J. Henrichs[a,b] **and A. Nelles**[a,b] **for the RNO-G Collaboration**
(a complete list of authors can be found at the end of the proceedings)

[a] *Deutsches Elektronen-Synchrotron DESY,*
Platanenallee 6, 15738, Zeuthen, Germany
[b] *Erlangen Center for Astroparticle Physics (ECAP), FAU Erlangen-Nuernberg*
Nikolaus-Fiebiger-Straße 2, 91058 Erlangen, Germany

E-mail: jakob.henrichs@desy.de

The Radio Neutrino Observatory – Greenland (RNO-G) is an in-ice neutrino detector currently under construction. The detector is designed to make the first measurement of neutrinos beyond energies of ~10 PeV. Each of the planned 35 stations of the detector includes three log-periodic dipole array antennas (LPDA) pointing towards the sky. The stations cover an area of ~ 50 km^2 and enable RNO-G to measure the radio emission of cosmic-ray induced air-showers, thus making it a cosmic-ray detector as well. As other experiments have shown, such radio emission can be used to make precision cosmic-ray measurements. Additionally, the location of the experiment at Summit Station, at a height of ~3000 m, enables RNO-G to study the phenomena of shower cores hitting the air/ice boundary and further developing in the ice itself. Moreover, RNO- G is also able to study high energetic muons, created in cosmic-ray induced air-showers, which penetrate into the ice from above.

In this contribution, we will give an overview of the cosmic-ray analysis of RNO-G and report the current status. This includes outlining the method used for identifying the air-shower signals using signal templates, showing the first cosmic-ray candidate events and discussing systematic uncertainties.

The 38th International Cosmic Ray Conference (ICRC2023)
26 July – 3 August, 2023
Nagoya, Japan

PoS(ICRC2023)259

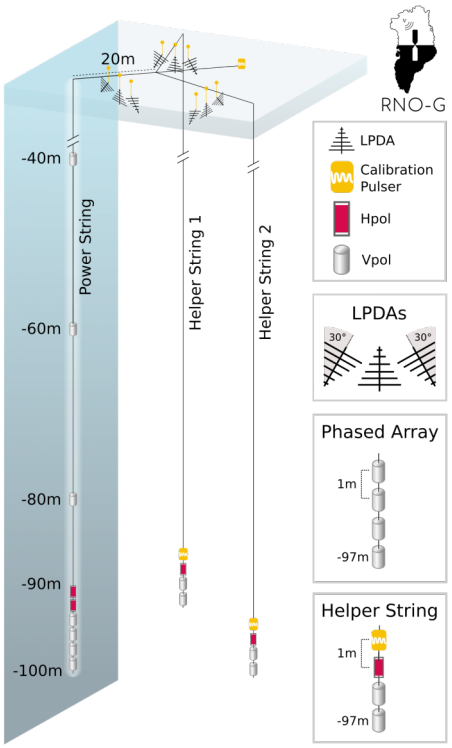

Figure 1: Schematic showing the structure of an RNO-G station. The station includes a deep and a shallow component. The deep component ($\sim$ 100 m) consist of three strings with antennas sensitive to the vertical (Vpol) and horizontal (Hpol) signal polarization, which are used to detect radio emission stemming from neutrino interactions. The shallow component ($\sim$ 3 m in vertical extent) consist of log-periodic-dipole-array antennas (LPDA), of which six point into the ice and can detect radio signals from neutrinos. Three additional LPDAs point towards the sky and are used to detect cosmic-ray air-showers.

1. Introduction

The Radio Neutrino Observatory Greenland (RNO-G) [1] is an in-ice neutrino detector currently being built at Summit Station in Greenland [2]. The layout of a single station is presented in Fig. 1. Per station, three log-periodic-dipole-array antennas (LPDA) point upwards, making it possible to detect the radio emission of air showers [3]. The main goal of the upward pointing antennas is to veto down-coming backgrounds to a neutrino search. However, the large spacing between stations (1.25 km) implies a large aggregate area of $\sim$ 50km^2 and thus renders RNO-G an instrument with good cosmic-ray sensitivity. The station design allows RNO-G to study the distribution of cosmic-ray arrival direction and polarization and investigate high energetic muons that are created by air showers and interact in the ice [4]. Additionally, due to the altitude of RNO-G ($\sim$ 3000m), it will be possible to study air showers that hit the ice before they are fully developed. Such air showers will further propagate into the ice, creating additional signals via the Askaryan effect [5, 6]. Finally, measuring cosmic rays will help to calibrate the detector and refine our understanding of the instrument.

2. Analysis strategy

The main strategy for the cosmic-ray analysis is to use template matching, meaning template waveforms are correlated with data waveforms and the correlation value is used for signal identification. This approach is based [7]. However, in contrast to the ARIANNA analysis, which used numerous templates ($\sim$ 200,000), the RNO-G collaboration plans to exploit the fact that

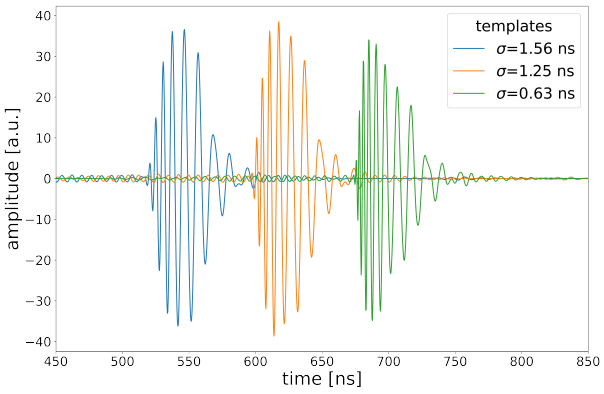

Figure 2: The three templates used in the cosmic-ray analysis. Their differences are entirely due to varying the width of the Gaussian function σ used as the input electric field, when creating them. Their similarity is caused by the identical system response.

the measured waveforms are mainly determined by the hardware response of the signal chain, to drastically reduce the number of considered templates. The electric field arrives at the antenna as a nanosecond long pulse, which is spread over a larger time and transformed into an oscillating signal by the group delay and limited bandwidth of the antenna and amplifiers. To mimic this behavior, we use Gaussian pulses with different widths as input signal electric fields and propagate these through the signal chain by convolving it with the simulated vector effective length of the antenna (for a single zenith angle) and the measured response of the amplifiers. With this technique, we are able to generate three templates (Fig. 2) that cover nearly the complete parameter space used for the cosmic ray search [8]. The similarity of the three templates illustrate the dominance of the system response in determining the expected air-shower waveforms. To quantify the match between template and data, the correlation is calculated. Since the expected pulse is still relatively short in time and can appear anywhere in the trace, a scan of the complete waveform is performed. The scan is calculated by shifting the data and template waveform relative to each other and retaining the maximal value of the resulting correlation values. The formula used is the following:

$$\rho = \max\left(\rho(\Delta n)\right) = \max\left(\frac{\sum_i^m (V_1)_i \cdot (V_2)_{i+\Delta n}}{\sqrt{\sum_i^m (V_1)_i^2} \cdot \sqrt{\sum_{j=\Delta n}^{m+\Delta n} (V_2)_j^2}}\right) \tag{1}$$

where V_1, V_2 are the data and template voltage traces and Δn is the number of samples by which the two traces are shifted relative to each other. In the analysis, we only use a 200 ns window around the pulse for the template waveform in order to mitigate the influence of noise in regions where no signal is expected. Zero padding is applied to the data waveform, to ensure that a 200 ns window can be used near the edges of the waveform. For a cosmic-ray event, we expect (for high signal-to-noise (SNR) values) a correlation score > 0.8, while the background should have smaller values. Detailed SNR-dependent cuts will be optimized at a later stage.

3. Search for cosmic rays in RNO-G data

For these proceedings, and consistent with the 'blinded analysis' strategy adopted by RNO-G, we have only analyzed the so-called 'burn' sample data of three RNO-G stations, which are the three

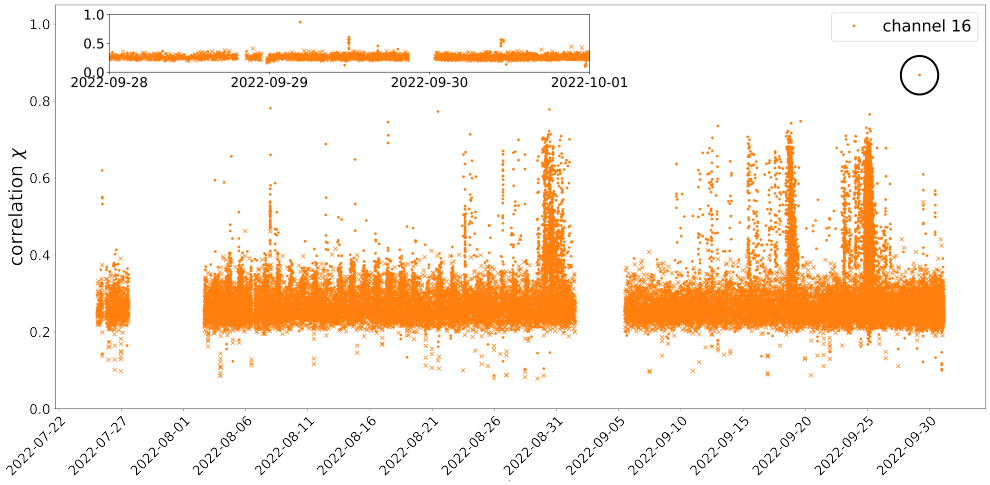

Figure 3: Correlation score of the burn sample data (station 23, channel 16) over the period 07/25/2022–10/01/2022 as a function of time. Forced trigger events are marked by crosses and the surface trigger events are marked by dots. The inset shows a zoomed-in version of the time period spanning observation of a cosmic-ray candidate event (Armstrong), which is circled in black.

stations with the optimal operation conditions for the upward facing antennas. The burn sample data is the data which is transferred by satellite and includes only ∼ 3% of the collected data. We select the data collected during good summer running (between June 25th and October 1st, 2022) for further analysis. In what follows, unless otherwise indicated, we will use the data from station 23, channel 16 (one of the three upward facing LPDAs) for illustration.

Correlation values as a function of time: The correlation values, as a function of time, provide a selection criterion (Fig. 3). The bulk of events are distributed around a correlation value of ∼ 0.3, which is the region where thermal noise triggers will cluster. Events with smaller correlation values can be attributed to continuous wave (CW) signals, e.g. the communications from a weather balloon, which flies twice a day at the RNO-G site, or from hand-held radios. These events are observable either as surface triggers (using two out of three upward facing LPDAs in coincidence) or in waveforms captured during a 'forced' trigger (which records an event every 10 s). In addition, there are time-clustered events with larger correlation values (spikes), evident as vertical bands in the plot. (see discussions about noise properties in [9, 10].) Since cosmic rays are expected isolated in time, we exclude events associated with these spikes from signal candidacy. As an alternative to excluding the correlation spikes, we could also use correlation-time cluster cuts (as done by ARIANNA [7]) or a high trigger rate cut. First tests on the latter are very promising and further work is ongoing. Some of the spikes are also coincide with high wind speeds. It was observed by RNO-G and other radio-ice experiments that the trigger rate increases during high wind periods, due to triboelectric discharges created by high winds [11]. We are also currently working on a wind event discriminator based on machine learning, which shows promising initial performance. If we zoom in to a smaller time period (07/27 - 09/30), we find one of the three cosmic-ray candidate events ("Armstrong") clearly isolated in time. In total, we find three well-isolated cosmic-ray candidate signal events, as detailed below and that there is no significant contribution of background events in the signal region.

Correlation values as a function of SNR: Using the same data as used for the time vs. correlation plot, the correlation value as the function of the SNR of each event (see Fig. 4) provides a second selection criterion. The signal-to-noise ratio (SNR) is calculated by dividing the peak amplitude of the trace (A) by the voltage RMS (V_{RMS}) of forced trigger data (SNR = A/V_{RMS}). In addition to the data, the correlation values of a cosmic-ray simulation are shown in Fig. 4. The air shower signals are simulated with CORSIKA and CoREAS [12, 13] and the surface trigger is simulated with a simple amplitude threshold of 10σ, which is purposely conservative, given the fact that precise modeling of the actual trigger is still on-going. For the simulation, the correlation values at high SNR are close to 1, with smaller values at lower SNR. The decrease in correlation values is expected due to the increasing influence of the noise for decreasing SNR; at sufficiently small SNR values, simulated correlation values approximate that of thermal noise. For forced trigger data (Fig. 4 left), we note that the majority of the events cluster around the putative thermal noise band, at small correlation values. The fact that most of the force triggers are thermal noise and have a very small background occupancy indicates the radio-quiet quality of the RNO-G site. A small tail of events towards higher SNR and smaller correlation values is also visible. This tail can be attributed to the CW events. The plot containing only surface trigger events (Fig. 4 right) also shows a large contribution of events clustered in the thermal noise region. In addition, we observe a cluster of events at rather small correlation values (~ 0.4), which are also clearly separated from the cosmic-ray region. The last large cluster is at higher correlation values (~ 0.6) and thus is relatively close to the signal region. This cluster is still well separated from the simulation, but has the highest potential for interfering with the signal region. An analysis of these events showed that they come from times with a higher than normal trigger rate, typically high wind periods. Therefore, the trigger rate cut or wind event discrimination discussed above could be very effective to suppress this cluster, if needed. The event with the highest correlation (at an SNR of ~ 1.5) is the candidate event Armstrong (same event as shown in the inset of Fig. 3). The other two candidate events at smaller SNR values are not clearly visible, since they are close to the boundary to the cluster with the highest correlation value. We observe that Armstrong does not attain the correlation values expected from the simulation, which indicates that the template is not describing the candidate event well enough. (More details about this are presented below.) Similar plots were created for stations 13 and 24. Both plots show a better separation of simulated signal and the bulk of our data. These two stations also have a higher trigger threshold compared to station 23 and thus have less noise contamination.

Candidate events: In total, three cosmic-ray candidate events are found in the burn sample data for stations 23, 13 and 24 (time period: 07/25/2022 - 10/01/2022). These three candidate events are all found in station 23, one with a high SNR (Armstrong) and two with lower SNR (see Fig. 6). We attribute the fact that candidate events are found exclusively in station 23 to the lower threshold compared to the other stations. In general, the number of candidate events found is within the expectations from simulation, given the currently large uncertainty on the trigger threshold. Armstrong's waveform and frequency spectrum is shown in Fig. 5. The frequency spectrum (especially channel 16) shows the expected behavior of a cosmic ray convoluted with the vector effective length and the amplifier response. As another check, the downward facing antennas were checked. All the downward facing antennas show a much smaller signal, suggesting that the signal came from above. Additionally, a preliminary direction and polarization reconstruction was

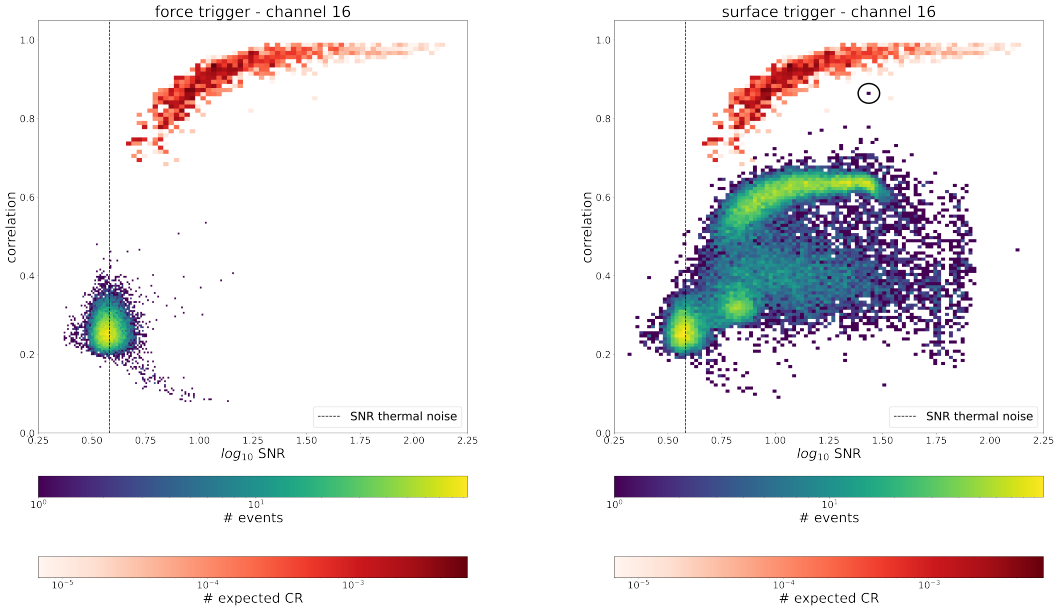

Figure 4: Correlation score as a function of SNR, with the dotted line showing the value for thermal noise. The red histogram shows the distribution of expected cosmic-ray events simulated for a surface amplitude trigger with a threshold of 10 σ. The blue-green-yellow histogram shows the distribution of the burn sample data (station 23, channel 16) between 07/25/2022 and 10/01/2022. The candidate event is circled in black.

performed. Both reconstructions are compatible with expectation for a cosmic ray and inconsistent, e.g., with the signal coming Summit Station.

Influence of incorrect modeling of the hardware response: As mentioned above, the correlation value of the candidate event differs from expectations based on our simulation. To investigate whether this is a consequence of incorrect modeling of the hardware response, we conducted measurements in the lab. With a signal generator, we injected a simulated template (without amplifier response) into our signal chain downstream of the antenna. These lab measurements can be compared to the fully simulated template,including amplifier response. We found differences between the measured and simulated template group delays, and that there is a deficit of high-frequency content in the measured template, however the latter is likely arising from deficiencies in the arbitrary waveform generator, which will be corrected in future studies. In order to assess how these differences impact the correlation, the correlation of the lab-based template with the complete data set is calculated Fig. 7). With the lab-based template, we found an increase of the correlation value for all three candidate events, while the background distribution roughly stays the same. In addition, the Armstrong event is compatible with the simulation using this lab-based template. Therefore, it seems that the mismatch in correlation between candidate events and simulation arises from an incorrect modeling of the group delays through the signal chain. Investigating this additional group delay is currently ongoing, it may stem from currently not-measured hardware-components such as connectors.

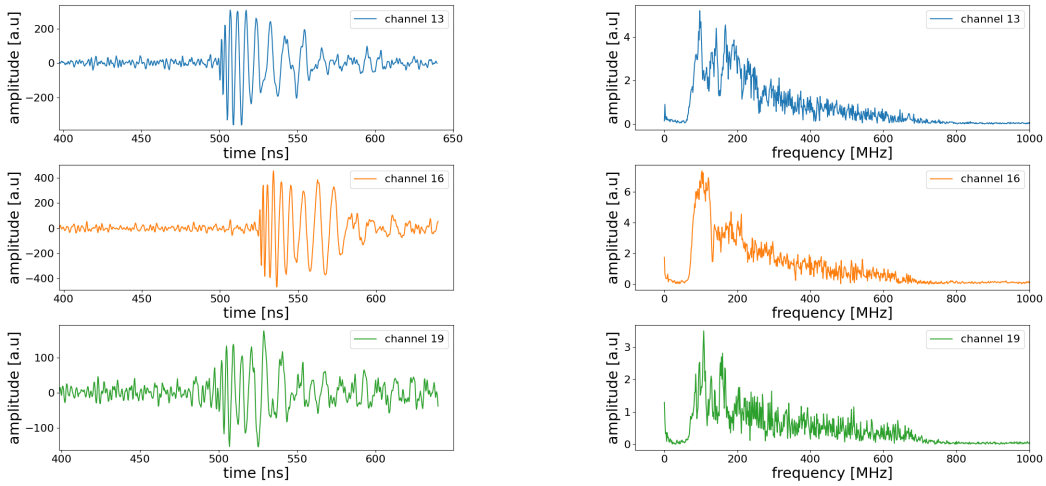

Figure 5: Voltage traces (left) and the frequency spectra (right) for the three upward facing channels (13, 16, 19) of the candidate event, Armstrong

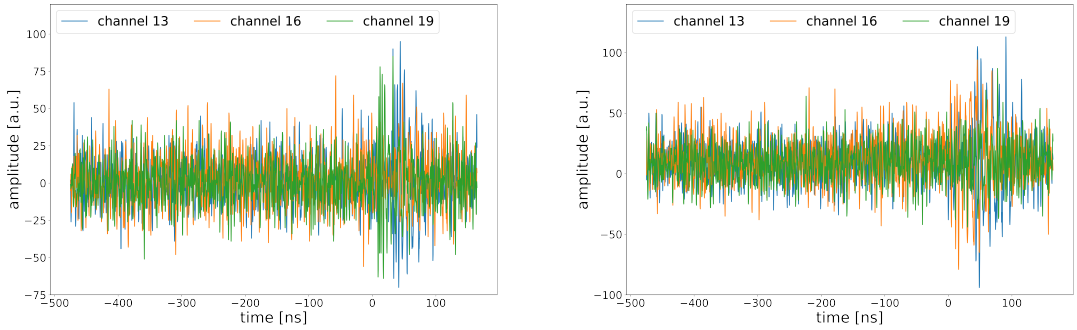

Figure 6: Voltage traces of the two candidate events with a low SNR for the three upward facing channels (13, 16, 19).

4. Conclusion

In this work, we have presented the current status of the search for cosmic-ray induced air-showers in the RNO-G data. We showed the results from 68 days of burn sample ($\sim$ 3%) data of three stations. From the data, it is evident that signal candidates can be well-separated from background. Additionally, we considered a high trigger rate cut and a wind event discrimination as promising means of reducing background events close to the signal region. In the burn sample data, we found three cosmic-ray candidate events, of which one has a clear signal in all three channels (high SNR). Following a preliminary direction and polarization reconstruction and the behavior of the frequency spectrum, we are confident that the high SNR event arises from a cosmic-ray. Moreover, a systematic offset of the candidate event's correlation value suggests that the signal chain used for the template creation may not be modeled correctly yet. Additional systematic studies are on-going. After, it is planned to extend the analysis to more stations and more data (e.g. the data from 2023) and un-blinding the full data-set. This will give more cosmic-ray events

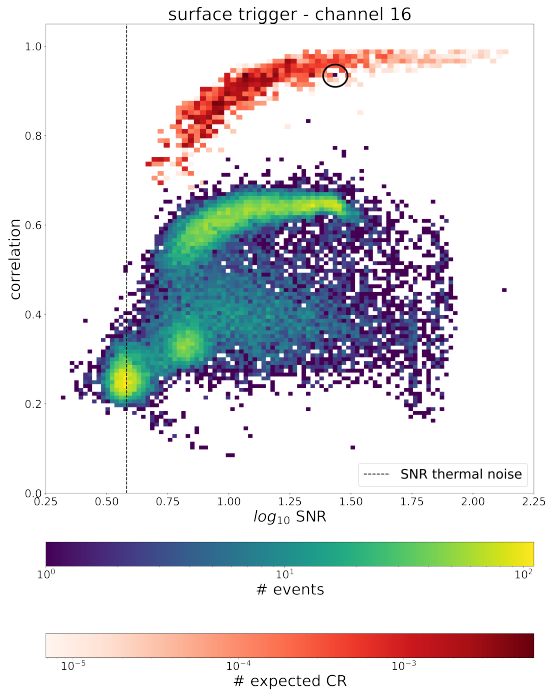

Figure 7: Correlation score of the same data as in Fig. 4, but calculated using the lab-based, rather than the simulated template. Only the surface trigger events are shown; the dotted line shows the SNR value of thermal noise. The blue-green-yellow histogram shows the data, while the red histogram shows the same simulation as in Fig. 4. The candidate event is circled in black.

and thus allow us to compare the distribution of arrival directions, polarization, and the number of expected events to simulations. In addition, it is planned to search the data for cosmic-ray induced air showers that impact the ice before they are fully developed and therefore continue to propagate in the ice.

References

[1] **RNO-G** Collaboration, J. A. Aguilar *et al.* *JINST* **16** no. 03, (2021) P03025. [Erratum: JINST 18, E03001 (2023)].
[2] S. Hallmann for the RNO-G collaboration *these proceedings, PoS(ICRC2023)1043* .
[3] F. D. Kahn and I. Lerche *Proc. R. Soc. Lond. A* **289** (1966) 206–213.
[4] L. Pyras and I. Plaisier for the RNO-G collaboration *PoS* **ECRS** (2022) 088.
[5] G. A. Askar'yan *Zh. Eksp. Teor. Fiz.* **41** (1961) 616–618.
[6] S. D. Kockere *et al.* *Physical Review D* **106** (2022) 043023.
[7] S. W. Barwick *et al.* *Astropart. Phys.* **90** (2017) 50–68.
[8] **RNO-G** Collaboration, J. Henrichs *et al.* *PoS* **ARENA2022** (2023) 007.
[9] Z. Meyers for the RNO-G collaboration *these proceedings, PoS(ICRC2023)1142* .
[10] T. Glüsenkamp for the RNO-G collaboration *these proceedings, PoS(ICRC2023)1056* .
[11] J. Aguilar *et al.* *Astroparticle Physics* **145** (2023) 102790.
[12] D. Heck *et al.* *Wissenschaftliche Berichte* **FZKA-6019** (1998) 1–90.
[13] T. Huege, M. Ludwig, and C. W. James *AIP Conference Proceedings* **1535** no. 1, (2013) 128–132.

Full Author List: RNO-G Collaboration

J. A. Aguilar[1], P. Allison[2], D. Besson[3], A. Bishop[10], O. Botner[4], S. Bouma[5], S. Buitink[6], W. Castiglioni[8], M. Cataldo[5], B. A. Clark[7], A. Coleman[4], K. Couberly[3], P. Dasgupta[1], S. de Kockere[9], K. D. de Vries[9], C. Deaconu[8], M. A. DuVernois[10], A. Eimer[5], C. Glaser[4], T. Glüsenkamp[4], A. Hallgren[4], S. Hallmann[11], J. C. Hanson[12], B. Hendricks[14], J. Henrichs[11,5], N. Heyer[4], C. Hornhuber[3], K. Hughes[8], T. Karg[11], A. Karle[10], J. L. Kelley[10], M. Korntheuer[1], M. Kowalski[11,15], I. Kravchenko[16], R. Krebs[14], R. Lahmann[5], P. Lehmann[5], U. Latif[9], P. Laub[5], C.-H. Liu[16], J. Mammo[16], M. J. Marsee[17], Z. S. Meyers[11,5], M. Mikhailova[3], K. Michaels[8], K. Mulrey[13], M. Muzio[14], A. Nelles[11,5], A. Novikov[19], A. Nozdrina[3], E. Oberla[8], B. Oeyen[18], I. Plaisier[5,11], N. Punsuebsay[19], L. Pyras[11,5], D. Ryckbosch[18], F. Schlüter[1], O. Scholten[9,20], D. Seckel[19], M. F. H. Seikh[3], D. Smith[8], J. Stoffels[9], D. Southall[8], K. Terveer[5], S. Toscano[1], D. Tosi[10], D. J. Van Den Broeck[9,6], N. van Eijndhoven[9], A. G. Vieregg[8], J. Z. Vischer[5], C. Welling[8], D. R. Williams[17], S. Wissel[14], R. Young[3], A. Zink[5]

[1] Université Libre de Bruxelles, Science Faculty CP230, B-1050 Brussels, Belgium
[2] Dept. of Physics, Center for Cosmology and AstroParticle Physics, Ohio State University, Columbus, OH 43210, USA
[3] University of Kansas, Dept. of Physics and Astronomy, Lawrence, KS 66045, USA
[4] Uppsala University, Dept. of Physics and Astronomy, Uppsala, SE-752 37, Sweden
[5] Erlangen Center for Astroparticle Physics (ECAP), Friedrich-Alexander-University Erlangen-Nürnberg, 91058 Erlangen, Germany
[6] Vrije Universiteit Brussel, Astrophysical Institute, Pleinlaan 2, 1050 Brussels, Belgium
[7] Department of Physics, University of Maryland, College Park, MD 20742, USA
[8] Dept. of Physics, Enrico Fermi Inst., Kavli Inst. for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA
[9] Vrije Universiteit Brussel, Dienst ELEM, B-1050 Brussels, Belgium
[10] Wisconsin IceCube Particle Astrophysics Center (WIPAC) and Dept. of Physics, University of Wisconsin-Madison, Madison, WI 53703, USA
[11] Deutsches Elektronen-Synchrotron DESY, Platanenallee 6, 15738 Zeuthen, Germany
[12] Whittier College, Whittier, CA 90602, USA
[13] Dept. of Astrophysics/IMAPP, Radboud University, PO Box 9010, 6500 GL, The Netherlands
[14] Dept. of Physics, Dept. of Astronomy & Astrophysics, Penn State University, University Park, PA 16801, USA
[15] Institut für Physik, Humboldt-Universität zu Berlin, 12489 Berlin, Germany
[16] Dept. of Physics and Astronomy, Univ. of Nebraska-Lincoln, NE, 68588, USA
[17] Dept. of Physics and Astronomy, University of Alabama, Tuscaloosa, AL 35487, USA
[18] Ghent University, Dept. of Physics and Astronomy, B-9000 Gent, Belgium
[19] Dept. of Physics and Astronomy, University of Delaware, Newark, DE 19716, USA
[20] Kapteyn Institute, University of Groningen, Groningen, The Netherlands

Acknowledgments

We are thankful to the staff at Summit Station for supporting our deployment work in every way possible. We also acknowledge our colleagues from the British Antarctic Survey for embarking on the journey of building and operating the BigRAID drill for our project. We would like to acknowledge our home institutions and funding agencies for supporting the RNO-G work; in particular the Belgian Funds for Scientific Research (FRS-FNRS and FWO) and the FWO programme for International Research Infrastructure (IRI), the National Science Foundation (NSF Award IDs 2118315, 2112352, 211232, 2111410) and the IceCube EPSCoR Initiative (Award ID 2019597), the German research foundation (DFG, Grant NE 2031/2-1), the Helmholtz Association (Initiative and Networking Fund, W2/W3 Program), the University of Chicago Research Computing Center, and the European Research Council under the European Unions Horizon 2020 research and innovation programme (grant agreement No 805486).

PoS(ICRC2023)259