PROCEEDINGS OF SCIENCE

Likelihood-reconstruction for radio detectors of cosmic rays and neutrinos

Martin Ravn,[a] **Christian Glaser,**[a,*] **Thorsten Glüsenkamp,**[a,b] **Ayca Özcelikkale**[c] **and Alan Coleman**[a]

[a] *Dept. of Physics and Astronomy, Uppsala University, Box 516, S-75120 Uppsala, Sweden*

[b] *Oskar Klein Centre and Dept. of Physics, Stockholm University, SE-10691 Stockholm, Sweden*

[c] *Dept. of Electrical Engineering, Uppsala University, Box 65, 75103 Uppsala, Sweden*

E-mail: martin.ravn@physics.uu.se

Radio detection of cosmic rays and neutrinos is an established technique pursued by many experiments. However, current reconstruction methods ignore bin-to-bin correlations of the measured waveforms, which limits reconstruction resolution and, so far, has prevented calculations of event-by-event uncertainties. In this work, we solve this shortcoming and present a likelihood description of neutrino and cosmic-ray signals in radio detectors. Using simulated data, we demonstrate that with this method, parameters such as the neutrino energy and direction, or lower-level parameters such as the energy fluence, can be obtained more precisely and with correct event-by-event uncertainties. The reconstruction code will be available through the open-source software NuRadioReco.

39th International Cosmic Ray Conference (ICRC2025)
15–24 July 2025
Geneva, Switzerland

*Speaker

PoS(ICRC2025)271

1. Introduction

Radio detectors are a powerful tool for ultra-high-energy (UHE) cosmic ray and neutrino astronomy. In cosmic-ray physics, radio arrays have demonstrated high-precision measurements of both primary energy and shower geometry [1]. For instance, the Pierre Auger Observatory has used radio antennas to complement surface and fluorescence detectors, enabling energy estimates with small systematic uncertainties and improved reconstruction of the depth of shower maximum, X_{max}, a key observable for determining the mass composition of cosmic rays. The recently finished AugerPrime upgrade [2], future radio enhancement to IceTop [3], and extension of cosmic-ray capabilities to SKA [4] further underline the growing role of radio detection in cosmic-ray science.

In the field of UHE neutrino detection, experiments like ARA [5] and ARIANNA [6] have demonstrated the viability of the experimental concepts of in-ice radio neutrino detectors. With the ongoing construction of the Radio Neutrino Observatory in Greenland (RNO-G) [7] and the planned radio extension for the IceCube-Gen2 array [3], radio detection will likely play a crucial role in the characterization of the UHE neutrino flux.

Precise reconstruction of the signals in radio detectors is critical for any science analysis using cosmic-ray or neutrino radio observatory data. The critical step is reconstructing the electric field that arrived at each antenna. Current reconstruction techniques can be divided into two general categories: *unfolding* and *forward folding* methods. In unfolding-based reconstruction, the measured traces in two or more antennas with orthogonal polarization are unfolded with the antenna response to estimate the electric field from which observables like energy fluence and polarization can be estimated. This is widely used for cosmic-ray radio detectors and works well for data with high signal-to-noise ratios (SNRs).

In forward-folding-based reconstructions, a deterministic parameterized electric field signal model is used to calculate the electric field generated at the antenna, which is then folded through the antenna and detector response. The resulting signal is then compared to the measured data, and an objective function is calculated and minimized, e.g.,

$$\chi^2 = \sum_{\mathrm{n=0}}^{n_t-1} \frac{[\boldsymbol{x}_n - \boldsymbol{\mu}_n(\boldsymbol{\theta})]^2}{\sigma^2}, \tag{1}$$

where $\boldsymbol{x}$ is the measured trace represented as a vector, $\boldsymbol{\mu}$ is a predicted signal that depends on parameters $\boldsymbol{\theta}$, n_t is the number of samples in the trace, and σ is the standard deviation of the noise in the antenna. The forward folding method was proposed in Ref. [8], and was shown to improve reconstruction resolution, especially at low SNR. It was also successfully employed to reconstruct the polarization of electric field from cosmic-ray air showers with the ARIANNA detector [9]. The gain in sensitivity of this method comes from the fact that it fully exploits the time-domain structure of the waveform and that it constrains the estimated electric field to a physically meaningful signal model.

The main limitation of the forward folding method is that the objective function is a simple χ^2, which ignores bin-to-bin correlations in the noise. However, the noise present in radio detectors is band-limited, which gives strong correlations between neighboring bins, and the χ^2 can therefore not be interpreted statistically. Hence, the method does not give correct event-by-event uncertainty estimates, and reconstruction resolution is limited. As previously discussed in Ref. [10], we

propose optimizing a likelihood in forward-folding-based methods, which is derived from the correct probabilistic description of a deterministic signal plus random band-limited noise in a radio detector. This gives the most precise possible reconstruction for known physical processes and detector responses, and enables correct estimations of event-by-event uncertainties. Since we have previously published on this topic with a focus on neutrino reconstruction, in this publication, we will briefly summarize the method and demonstrate its application to the reconstruction of polarization and energy fluence of electric fields from cosmic-ray air showers.

2. Likelihood description of radio signals and noise

The noise observed in a radio detector can be modelled as a multivariate normal distribution with dimensionality of the length of the measured trace, n_t, with a non-diagonal covariance matrix that parameterizes the correlations of the noise. The $n_t \times n_t$ covariance matrix, $\mathbf{\Sigma}$, can be derived from the spectrum of the noise or many datasets containing purely noise (see Ref. [10]).

The minus two log likelihood of a signal, $\boldsymbol{\mu}$, given a measured trace in n_{ant} antennas/channels is then

$$-2 \ln \mathcal{L}(\boldsymbol{\mu}(\theta); \boldsymbol{x}, \mathbf{\Sigma}) = \sum_{i=1}^{n_{\text{ant}}} \big(\boldsymbol{x}_i - \boldsymbol{\mu}_i(\theta)\big)^{\mathrm{T}} \mathbf{\Sigma}_i^{-1} \big(\boldsymbol{x}_i - \boldsymbol{\mu}_i(\theta)\big) + \text{constants}, \tag{2}$$

where the constants can be ignored in a minimization since only relative values are relevant. By minimizing this likelihood in a reconstruction where $\boldsymbol{\mu}$ represents the expected electric field – obtained either from physical simulations or parameterized models and subsequently folded with the known detector response – the most accurate reconstruction should be achieved.

In Ref. [10], it was shown with a simulation study that for radio detectors with band-limited Gaussian noise, the reconstruction resolution was improved using the likelihood compared to a χ^2. This was shown in the context of an in-ice radio neutrino detector, where the signal, μ, was simulated using NuRadioMC. Furthermore, event-by-event reconstruction uncertainties were estimated using Wilks' theorem and were shown to have perfect coverage. The likelihood-based forward folding method is thus a promising reconstruction method, and it should be explored to which other problems within the field of radio detection of UHE particles it can be applied.

3. Application to cosmic-ray electric field reconstruction

In this section, we apply the likelihood reconstruction to electric fields from cosmic-ray air showers to extract the electric field energy fluence and polarization at the location of a single dual-polarized antenna as used in most radio air-shower arrays [1]. This constitutes one component of a full cosmic-ray air shower reconstruction, where the fluences and polarizations at each antenna are combined to estimate the energy (e.g. [11]) and X_{max} (e.g. [12]) of the shower. The currently most widely used method for estimating the fluence and polarization of an electric field at a dual-polarized antenna relies on unfolding and calculating the fluence directly from the unfolded trace in a time window (see Section 3.2 for full description). Hence, it does not utilize the full time-domain information of the waveform, which limits reconstruction resolution, and it is known to be biased at low SNR. In the following, we show that this can be improved by using a likelihood-based forward folding method in a simulation study.

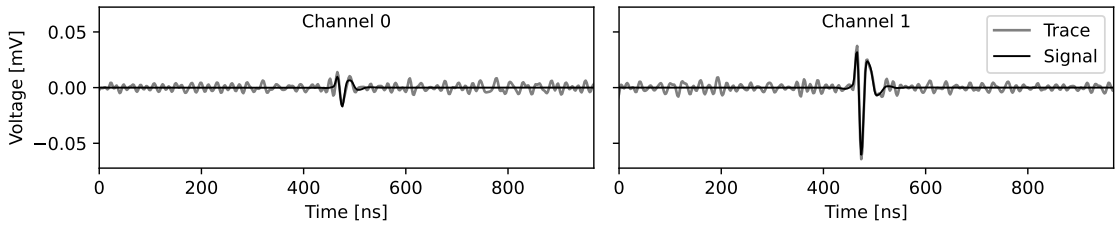

Figure 1: Example air-shower signal from CoREAS simulation as measured by a dual-polarized antenna. The black curve shows the noiseless signal, whereas the gray curve shows the signal with measurement noise.

3.1 Simulation dataset

To evaluate the performance of the reconstruction methods discussed here, we generate a simulation dataset of cosmic-ray air showers using the CoREAS extension [13] of the CORSIKA framework. The dataset consists of a large set of silicon-initiated showers with energies uniformly distributed in $\log_{10}(E)$ between 10^{15} and 10^{19} eV, and arrival directions uniformly distributed over a zenith angle range of 60° to 85°. We simulate the showers for the magnetic field and atmospheric conditions at the South Pole. For each air shower, these simulations produce the time-dependent electric field at 160 observer positions placed on the ground in the standard star-shaped grid around the shower axis. We use these electric fields directly in the subsequent calculations and do not apply any additional randomization of positions and interpolation. As we present all our findings as a function of signal-to-noise ratio, this is the most accurate approach.

Each simulated electric field is observed by a dual-polarized antenna. The antenna is a logarithmic periodic dipole antenna (LPDA) that measures with a sampling rate of 500 MHz in the frequency band from 30 MHz to 80 MHz with two readout channels, one sensitive to north-south polarized electric fields and one sensitive to east-west polarizations. We employ a simple analytical approximation of the LPDA response available in NuRadioReco [8], which we use to simulate the response of the antenna to the electric field.

Finally, we add noise to the traces and apply several filters to simulate hardware- and analysis-level filters, which remove environmental noise at low and high frequencies. The generated noise is Gaussian with an amplitude corresponding to a temperature of $T_{\text{noise}} = 300\,\text{K}$, and the filter is a third-order Butterworth high-pass filter at 30 MHz and an eighth-order low-pass Butterworth filter at 80 MHz. Additionally, as very low SNR observers are overrepresented, we randomly skip some of them to make the distribution of datasets over SNR more uniform. The SNR is calculated using the maximum absolute amplitude of the noiseless traces relative to the standard deviation of the noise. This leaves us with 73,459 cosmic-ray electric-field waveforms with SNR values of below 20, which we will reconstruct in the following sections. An example of the resulting traces as measured by one of the dual-polarized antennas is shown in Fig. 1.

For the purpose of this study, we assume the cosmic-ray arrival direction is known from external reconstruction (e.g., from triangulation using the signal times in several observers), and the signal arrival time is known to within ±30 ns, which we call the *search window*. This is a reasonable assumption since the arrival direction can often be well constrained by fitting a wavefront to the arrival times of the signal in antennas with high SNR. The wavefront can then be extrapolated to estimate the arrival time in antennas with lower SNR.

3.2 Noise subtraction method

The most commonly used method to estimate the fluence and polarization of an electric field in a dual-polarized antenna is often called the *noise subtraction* method. In this method, the trace is unfolded with the antenna response for the known arrival direction, and the maximum of the Hilbert envelope (inside the search window) is used to define the signal arrival time. Then, the energy fluence is calculated in a ±30 ns window around the signal time. The average energy fluence of pure noise is calculated in a separate window of the trace that does not contain signal, and subtracted from the signal energy fluence. Additionally, the uncertainty on the fluence estimate is calculated using the standard deviation of the noise.

The basis vectors of the coordinate system chosen for three polarizations of the electric field are along the shower axis, $\hat{r}$ (with $\boldsymbol{E}_r$ assumed to be 0), perpendicular to the shower axis in the vertical plane, $\hat{\theta}$, and perpendicular to the shower axis in the horizontal plane $\hat{\phi}$. The estimate of the electric field polarization is then calculated as $P = \arctan(f_\phi / f_\theta)$.

3.3 Forward folding reconstructions

In forward folding methods, an electrical field is folded through the antenna response and compared to the measured traces. We use a simple frequency domain parameterization of the electrical field, which has been shown to describe radio pulses from cosmic-ray air showers well, given by:

$$\begin{pmatrix} \mathcal{E}_{\theta,k} \\ \mathcal{E}_{\phi,k} \end{pmatrix} = \begin{pmatrix} A_\theta \\ A_\phi \end{pmatrix} 10^{\alpha \cdot f_k + \beta (f_k - f_{\text{offset}})^2} e^{-i2\pi (f_k \cdot t_{\text{offset}} - \psi)} \tag{3}$$

where $\mathcal{E}$ is the electric field in the frequency domain, f_k are the frequencies, and A_{pol} are the amplitudes. The frequency dependence of $\mathcal{E}$ is controlled by α, which sets the slope of the electric field in the frequency domain and is often constrained to negative values, and β, which controls a second order correction where f_{offset} is fixed to 30 MHz. The last exponential controls the complex phases, where t_{offset} is the time of the pulse relative to the start of the trace, and ψ is a phase offset. Additionally, we numerically re-normalize the electric field, such that the parameters we reconstruct are directly the fluences of the two components, f_θ and f_ϕ. The electric field is then multiplied with the vector effective length (VEL) of the antenna for the known arrival direction and filtered accordingly. By applying the inverse Fourier transform, the estimate of the noiseless signal, $\boldsymbol{\mu}$, is obtained, which can then be compared to the measured trace in an objective function. The resulting signal depends on six parameters to be reconstructed, $\boldsymbol{\theta} = (f_\theta, f_\phi, \alpha, \beta, t_{\text{offset}}, \psi)$.

Traditionally, the objective function to be minimized is a χ^2, as expressed in Eq. 1, which does not consider the correlation in the noise. We propose minimizing the $-2 \ln \mathcal{L}$ expressed in Eq. 2, which correctly describes the band-limited nature of the noise. In this work, the minimization algorithms for both objective functions were improved to achieve better stability. For the $-2 \ln \mathcal{L}$ minimization, the algorithm consists of two steps. First, a matched filter is used to analytically profile over amplitudes and efficiently profile over a grid of t_0 in the search window:

$$\hat{s}_0(\boldsymbol{\theta}) = \sum_{i=0}^{n_{\text{ant}}} \frac{\boldsymbol{\mu}_i(\boldsymbol{\theta})^{\mathrm{T}} \boldsymbol{\Sigma}_i^{-1} \boldsymbol{x}_i}{\boldsymbol{\mu}_i(\boldsymbol{\theta})^{\mathrm{T}} \boldsymbol{\Sigma}_i^{-1} \boldsymbol{\mu}_i(\boldsymbol{\theta})}, \tag{4}$$

By choosing the t_0 that maximizes $\hat{s}_0(\boldsymbol{\theta})$, for any step in the minimization process, we automatically find the optimal time and amplitude of the signal. This is then used to calculate the $-2\ln\mathcal{L}$ which is the objective function of the first minimization. The method allows us to fit the parameters that control the shape of the signal, α, β, ψ, and the ratio of the fluences $r_{\theta/\phi} = f_\theta/f_\phi$, which dramatically decreases the number of local optima of the likelihood. In the second step, the $-2\ln\mathcal{L}$ is minimized with all parameters free to fine-tune the result. Since the $-2\ln\mathcal{L}$ has local minima for opposite polarizations, we run the algorithm four times, initialized in each quadrant to guarantee convergence.

The χ^2 minimization was improved in a corresponding way, by using a normalized cross-correlation to profile over time and amplitude in the first minimization. The normalized cross-corelation is equivalent to a matched filter with a diagonal covariance matrix, i.e., no correlations taken into account. This allows us to evaluate the impact of the objective function directly, i.e. $-2\ln\mathcal{L}$ or χ^2, with the same improvements to the minimization process.

For both objective functions, the uncertainties on the reconstructed parameters are extracted from the Hessian matrix at the minimum of the $-2\ln\mathcal{L}$ or χ^2.

3.4 Recontruction performances

All events in the simulation dataset were reconstructed using the three methods discussed. The relative uncertainties of the reconstructed fluences and the uncertainties on the reconstructed polarizations are shown in Fig. 2. The results are binned in SNR, and the medians and central 68% quantiles, $\pm\sigma_{68\%}$, of each bin are shown in the figure. The noise subtraction method shows two competing biases in the fluence at low SNR. At SNR<3, there is a positive bias originating from the peak-finding, which will always find the noisiest part of the unfolded electric field in the search window when almost no signal is present. The negative bias at SNR>3 is caused by the fluence estimator not taking into account that the signal and noise can be destructively aligned. Additionally, since f_θ and f_ϕ are not allowed to be negative, the reconstructed fluences pile up at 0 and the polarizations pile up at 0 degrees and 90 degrees, which causes very large uncertainties at low SNR.

The χ^2 minimization method reduced the bias at low SNR for both the fluence and polarization reconstruction. The $-2\ln\mathcal{L}$ minimization further improves on the bias and remains unbiased in the fluence down to SNR>2 and is virtually unbiased in the polarization. The remaining positive bias in the fluence at very low SNR is caused by the minimization fitting the pulse to the largest fluctuation in the noise when no signal is present.

In order to compare the uncertainties of the three methods, we plot half the width of the central 68% quantiles of each bin in the left column of Fig. 3. The fluence relative uncertainties of the likelihood method are decreased by a factor of 1.6 to 2.3 in the $5 < SNR < 15$ range compared to the noise subtraction method, and the polarization uncertainties are decreased by 1.7 to 13.1 in the same range. Compared to the χ^2 method, the likelihood fluence relative uncertainties are decreased by a factor of 1.2 to 1.3, and the polarization uncertainties by 1.2 to 1.5 in the same range. The latter improvement arises purely from changing the objective function, which underlines the importance of taking bin-to-bin correlations into account.

All methods estimate the uncertainties on the reconstructed parameters, which are shown in the right column of Fig. 3. The coverage was calculated by counting how often the true value

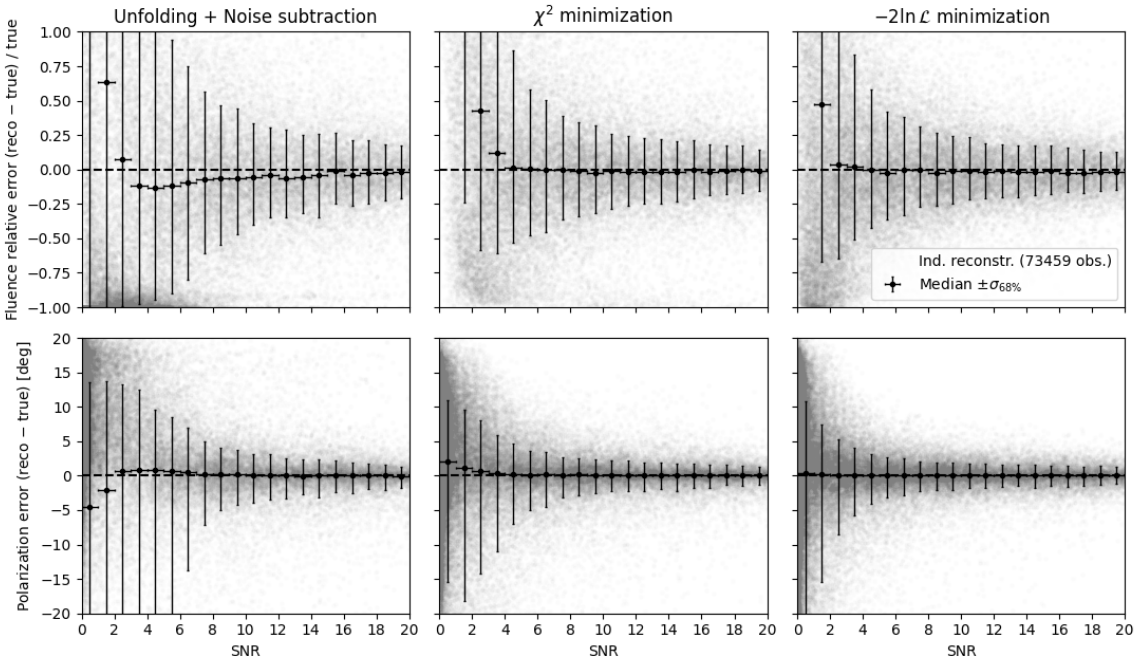

Figure 2: Reconstruction results for the three methods. The relative uncertainty (reconstructed minus true divided bye true) of the fluence and uncertainty (reconstructed minus true) of the polarization are shown for individual events, and they have been binned in SNR to calculate the median (bias) and asymmetric central 68% quantiles, $\pm\sigma_{68\%}$, of each bin.

was within the estimate plus/minus the estimated 68% uncertainty, which should ideally be 68% of the time. Both the noise subtraction and χ^2 methods severely under-predict the uncertainties, whereas the likelihood method predicts the uncertainties well across most SNRs. The remaining slight under-covarage is likely caused by the simplicity of the electric field model, which can not perfectly model the cosmic ray signals at high SNR.

Another method of estimating the energy fluence from noisy electric fields based on Rice distributions was recently proposed in Ref. [14], which takes into account the band-limited nature of the noise but does not constrain the electric field to a physical model. When applied to our simulation dataset, we observed an equivalent bias to the noise subtraction method at low SNR. However, the method correctly estimated the uncertainties across all SNRs, and it may hence be beneficial to use it at high SNRs.

4. Conclusion

The likelihood reconstruction is a promising technique for reconstructing particle signals in radio detectors. We have previously demonstrated that the method works well for in-ice radio neutrino reconstruction, and that it gives correct uncertainty estimates and improves resolution. In this work, we applied the technique to a cosmic-ray dataset and show that it gives a similar improvement in performance, and that the estimated uncertainties are correct. Additionally, we improved the minimization algorithm for forward-folding methods using a matched filter in the minimization process to guarantee convergence. These improvements can potentially unlock the

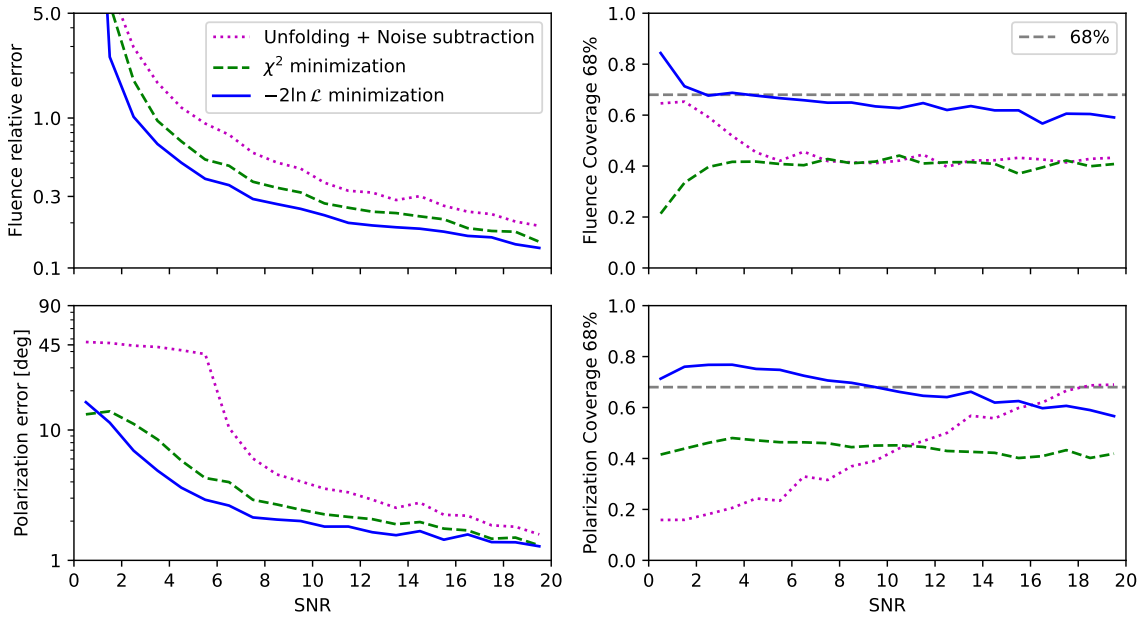

Figure 3: Left: Half the width of the of the asymmetric central 68% quantiles, $\pm\sigma_{68\%}$, of the fluence relative uncertainties (reconstructed minus true divided by true) and polarization uncertainties (reconstructed minus true) of the three methods. Right: Coverage of the event-by-event estimated uncertainties of the fluence and polarization for the three methods, i.e., how often the true value is within the estimate plus/minus the estimated 68% uncertainty, which should ideally be 68%.

use of antennas with low signal amplitudes in full cosmic-ray air shower reconstruction, to further confine key observables such as the energy of the shower and $X_{\max}$.

References

[1] T. Huege, *Phys. Rept.* **620** (2016) 1–52.

[2] **Pierre Auger** Collaboration, A. Abdul Halim *et al.*, *PoS* **ICRC** (2023) 344.

[3] **IceCube-Gen2** Collaboration, "IceCube-Gen2 Technical Design Report." https://icecube-gen2.wisc.edu/science/publications/TDR.

[4] S. Buitink *et al.*, *PoS* **ICRC** (2023) 503.

[5] **ARA** Collaboration, P. Allison *et al.*, *Astropart. Phys.* **35** (2012) 457–477.

[6] **ARIANNA** Collaboration, A. Anker *et al.*, *Adv. Space Res.* **64** (2019) 2595–2609.

[7] **RNO-G** Collaboration, J. A. Aguilar *et al.*, *JINST* **16** (2021) P03025.

[8] C. Glaser *et al.*, *Eur. Phys. J. C* **79** (2019) 464.

[9] **Arianna** Collaboration, A. Anker *et al.*, *JCAP* **04** (2022) 022.

[10] M. Ravn, C. Glaser, T. Glüsenkamp, and A. Coleman, *PoS(ARENA2024)054* .

[11] **Pierre Auger** Collaboration, A. Aab *et al.*, *Phys. Rev. D* **93** (2016) 122005.

[12] S. Buitink *et al.*, *Phys. Rev. D* **90** (2014) 082003.

[13] T. Huege, M. Ludwig, and C. W. James, *AIP Conf. Proc.* **1535** (2013) 128.

[14] S. Martinelli *et al.*, *Astropart. Phys.* **168** (2025) 103091.