

# Quantifying energy fluence and its uncertainty for radio emission from particle cascades in the presence of noise

Sara Martinelli <sup>a,b</sup>,\*, Tim Huege <sup>a,c</sup>, Diego Ravignani <sup>b</sup>, Harm Schoorlemmer <sup>d,e</sup>

<sup>a</sup> Institute for Astroparticle Physics, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

<sup>b</sup> Instituto de Tecnologías en Detección y Astropartículas (CNEA, CONICET, UNSAM), Buenos Aires, Argentina

<sup>c</sup> Astrophysical Institute, Vrije Universiteit Brussel, Brussels, Belgium

<sup>d</sup> IMAPP, Radboud University Nijmegen, Nijmegen, The Netherlands

<sup>e</sup> Nationaal Instituut voor Kernfysica en Hoge Energie Fysica (NIKHEF), Amsterdam, The Netherlands

## ARTICLE INFO

Dataset link: <https://gitlab.iap.kit.edu/saramartlinelli/fluence-and-uncertainty-estimation-base-d-on-rice-distribution>

### Keywords:

Radio detection  
Extensive air showers  
Neutrinos  
Cosmic rays  
Data analysis

## ABSTRACT

Measurements of radio signals induced by an astroparticle generating a cascade present a challenge because they are always superposed with an irreducible noise contribution. Quantifying these signals constitutes a non-trivial task, especially at low signal-to-noise ratios (SNR). Because of the randomness of the noise phase, the measurements can be either a constructive or a destructive superposition of signal and noise. To recover the electromagnetic energy of the cascade from the radio measurements, the energy fluence, i.e. the time integral of the Poynting vector, has to be estimated. Conventionally, noise subtraction in the time domain has been employed for energy fluence reconstruction, yielding significant biases at low signal-to-noise ratios. In several analyses, this bias is mitigated by imposing an SNR threshold cut, though this option is not ideal as it excludes valuable data. Additionally, the uncertainties derived from the conventional method are underestimated, even for large SNR values. To address this known issue, the uncertainties have so far typically been approximated and corrected by using ad-hoc terms. This work tackles these challenges by detailing a method to correctly estimate the uncertainties and lower the reconstruction bias in quantifying radio signals, thereby, ideally, eliminating the need for an SNR cut. The development of the method is based on a robust theoretical and statistical background, and the estimation of the fluence is performed in the frequency domain, allowing for the improvement of further analyses by providing access to frequency-dependent fluence estimation.

## 1. Introduction

Radio detection can be employed to measure cosmic-rays, photons, and neutrinos generating extensive air showers and particle cascades in dense media [1,2]. Since the radio emission is proportional to the number of electrons and positrons in the cascade, by quantifying the underlying signal of a measured radio pulse we can access the electromagnetic component of a shower, and thus, estimate the electromagnetic energy of the originating particle. As the electromagnetic energy is proportional to the area integral of the energy fluence (the energy deposit per unit area in terms of radio waves) [3], a correct reconstruction of the electromagnetic energy and its uncertainty relies on the determination of the energy fluence and its uncertainty.

The total energy fluence  $f_{\text{tot}}$  at a given antenna position is the time integral of the Poynting vector. For discretely sampled measurements,

this means:

$$f_{\text{tot}} = \epsilon_0 c \Delta t \sum_{\text{pol}} \left( \sum_j E_{\text{pol}}^2(t_j) \right), \quad (1)$$

where  $\Delta t$  is the sampling interval of  $\vec{E}(t)$ , the observed three-dimensional electric-field vector (in the equation broken down into its components),  $\epsilon_0$  is the vacuum permittivity, and  $c$  is the speed of light. In an analogous way, we can express the energy fluence as the sum of component-dependent contributions  $f_{\text{pol}}$ :

$$f_{\text{pol}} = \epsilon_0 c \Delta t \sum_j E_{\text{pol}}^2(t_j) \rightarrow f_{\text{tot}} = \sum_{\text{pol}} f_{\text{pol}}. \quad (2)$$

In radio measurements, the noise requires a sophisticated treatment: the measured pulse is either enhanced or diminished compared to the true signal depending on the random phase of the noise. This makes the evaluation of the fluence in the presence of noise non-trivial, especially

\* Corresponding author at: Institute for Astroparticle Physics, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany.

E-mail addresses: [sara.martinelli@kit.edu](mailto:sara.martinelli@kit.edu) (S. Martinelli), [tim.huege@kit.edu](mailto:tim.huege@kit.edu) (T. Huege), [diego.ravignani@iteda.cnea.gov.ar](mailto:diego.ravignani@iteda.cnea.gov.ar) (D. Ravignani), [hschoorlemmer@science.ru.nl](mailto:hschoorlemmer@science.ru.nl) (H. Schoorlemmer).

<https://doi.org/10.1016/j.astropartphys.2025.103091>

Received 26 July 2024; Received in revised form 27 January 2025; Accepted 31 January 2025

Available online 7 February 2025

0927-6505/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>).

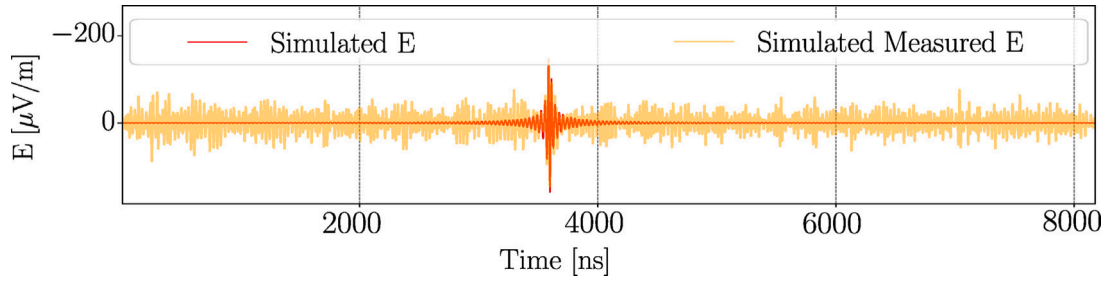


Fig. 1. Example of an electric-field component simulated in the absence of noise (red) and in the presence of noise (orange).

at low signal-to-noise ratios (SNR) [4].

So far, the conventional way of reconstructing the signal energy fluence consisted of estimating and subtracting the noise fluence in the time domain [5]. This way of treating the noise leads to a non-negligible reconstruction bias at low SNR values. To mitigate potential reconstruction bias, especially in the case of electromagnetic energy, a minimum required SNR value can be introduced as a prerequisite to using data from a given antenna. This approach is used, for example, in Refs. [6–8], where a cut is applied at the electric field level for inclusion of a given antenna station. One of the aims of this work is to diminish the reconstruction bias allowing the lowering and, ideally, removal of the need for an SNR cut. Furthermore, the uncertainties derived from the noise-subtraction method are known to be underestimated [7,9,10]. In this work, we detail a method for the quantification of the energy fluence and its uncertainty, exploiting a solid statistical background based on Rice distributions. The same statistical background has already been used within the ANITA experiment to achieve an estimation of the fluence in the frequency domain by employing fitting procedures [11]. Our method is based on estimating the fluence in the frequency domain, too, i.e., providing access to a frequency-dependent fluence estimation. Since the spectral shape can be described more easily than the pulse shape in the time domain, the method presented here can be further developed by combining it with the information derived from spectral modeling [12]. In the following, we validate the Rice-distribution method and compare it to the so-far widely adopted noise subtraction method. The bias of the reconstructed fluence as a function of the SNR will be discussed, as well as the evaluation of the uncertainties.

In the following, both methods are applied to the same data set of simulated air showers initiated by cosmic-rays. Because of its complexity, it is a non-trivial task to artificially generate realistic noise. To realistically simulate radio measurements, we have added the ambient background recorded at a particular site of the Pierre Auger Observatory to the radio-emission simulations. By using these measurements we get a realistic SNR distribution. We simulate the antenna response of the Radio Detector (RD) [13] of the Pierre Auger Observatory, sensitive to the 30–80 MHz frequency bandwidth. Nevertheless, the results of our study are independent of the detector simulation, and the Rice-distribution method can be applied to larger frequency bandwidths than the one tested in this work.

In Appendix B, we also present a method for estimating the spectral amplitudes of the signal and their uncertainties based on maximizing the Rice likelihood function. In Appendix C, the reader can find tables listing most of the notation and variables adopted in the following.

## 2. Simulated data set

Here, we shortly describe the set of simulations used to study the performance of the estimation methods.

### 2.1. Particle simulations and simulated radio-footprints

We exploit the same set of simulated air showers as used in Ref. [14]. The showers are initiated by four different cosmic-ray primaries:

proton, helium, nitrogen, and iron. For each primary, the particle and radio-emission footprints of about 2000 showers are simulated with CORSIKA/CoREAS v7.7401 [15,16], using the high energy interaction model QGSJETII-04 [17] and an optimized thinning level of  $10^{-6}$  [18]. The primaries have energies in the interval between  $10^{18.4}$  eV and  $10^{20.1}$  eV. The arrival directions of the showers are uniformly distributed across azimuth angles, while the zenith angle  $\theta$ , ranging from  $65^\circ$  to  $85^\circ$ , is uniformly distributed according to a  $\cos^2(\theta)$  weighting. The environmental conditions are set to match the Pierre Auger Observatory site, as is the detector layout. The cores of the showers are randomly distributed within this detector layout.

### 2.2. Simulations of electric fields in the absence of noise

The CoREAS simulations are processed through the reconstruction framework of the Pierre Auger Observatory, *Offline* [19]. The RD response simulation is performed and the electric field traces are reconstructed from the raw signal traces by using the Monte Carlo values of the arrival direction and the sensitivity pattern of the RD antennas. This means that the antenna response is applied to the CoREAS radio pulses and then again deconvolved. The electric field traces obtained correspond to the air shower pulses as they would be reconstructed in the absence of noise, in the 30–80 MHz frequency band. Each trace has a length of 8192 ns and  $\Delta t = 1$  ns, upsampled from the nominal 250 MSPS sampling rate with a factor of four.

### 2.3. Simulations of measured electric fields in the presence of noise

To get the simulated measurements of the cosmic-ray pulses in the presence of noise, we apply the detector simulation to the CoREAS simulations and add measured noise data. The measured traces of the ambient background recorded at the Pierre Auger Observatory site operating over one year are added to the digitized simulated traces. Each noise measurement is used multiple times by rolling the trace. In Fig. 1, we show an example of a simulated trace in the absence of noise and the corresponding trace simulated in the presence of noise, obtained as just described.

## 3. The noise subtraction method

The noise subtraction method is based on the assumption that the measured amplitude of the electric field is given by the sum of the cosmic-ray radio pulse and Gaussian-distributed noise  $e(t)$ , having a certain standard deviation  $\sigma_e$  and centered on zero [5]. To estimate the energy fluence, we define a **noise window**, delimited by  $t_1$  and  $t_2$ , and a **signal window**, delimited by  $t_3 = t_{\text{peak}} - \Delta$  and  $t_4 = t_{\text{peak}} + \Delta$ , where  $t_{\text{peak}}$  indicates the position of the radio pulse in the time trace, and  $\Delta$  has to be chosen such as to contain most of the pulse. As shown in Fig. 2, the noise window has to be far away from the signal window to contain mainly only noise contribution. For each component of the measured electric field at a given antenna position  $E_{\text{pol}}(t)$ , the energy

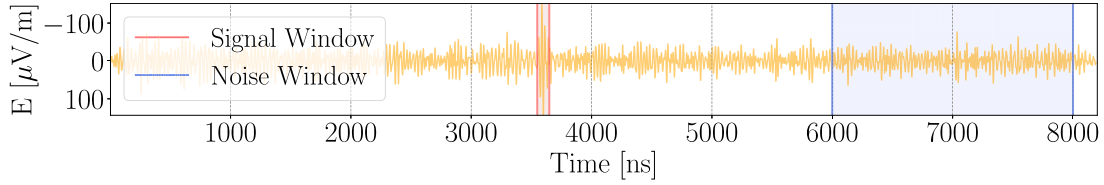


Fig. 2. Signal window (between red lines) and noise window (between blue lines) as used in the noise subtraction method.

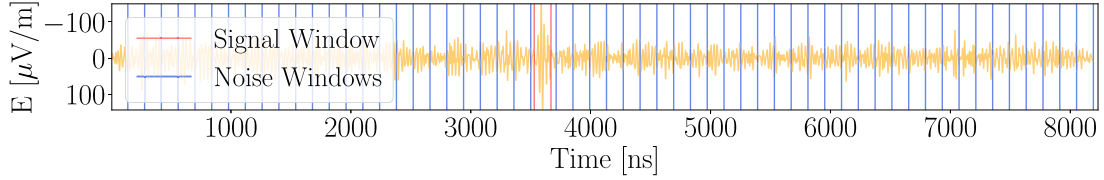


Fig. 3. Signal window (between red lines) and noise windows (between blue lines) as used in the Rice-distribution method. In the example,  $N = 57$  noise windows are used.

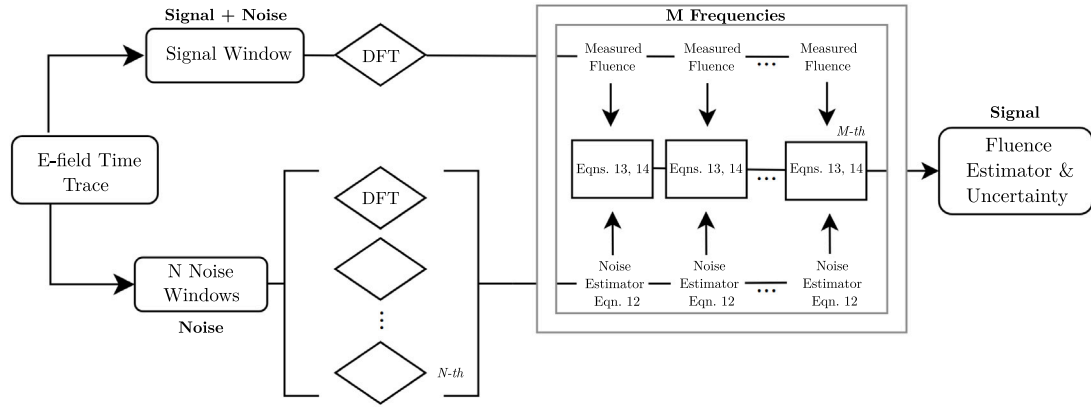


Fig. 4. Flowchart representing the logic of the energy fluence estimation for a single polarization of the electric field. Combining the  $M$  estimators of the noise fluence obtained from the  $N$  noise windows, and the fluence measured in the signal window, we get the polarization fluence estimator and its uncertainty.

fluence is estimated as:

$$\hat{f}_{\text{pol}} = \epsilon_0 c \Delta t \left( \sum_{t_j=t_3}^{t_4} E_{\text{pol}}^2(t_j) - \frac{t_4 - t_3}{t_2 - t_1} \sum_{t_j=t_1}^{t_2} E_{\text{pol}}^2(t_j) \right), \quad (3)$$

where the normalized fluence in the noise window is subtracted from the fluence calculated in the signal window. When dealing with very noisy pulses, the above equation can yield negative values. To prevent unphysical results, negative-valued estimators are typically set to zero, and we do the same in the analysis presented here. The uncertainty on the fluence estimator, as derived in [5], consists of:

$$\delta(\hat{f}_{\text{pol}}) = \sqrt{4 \epsilon_0 c \Delta t \hat{f}_{\text{pol}} \sigma_e^2 + 2 (\epsilon_0 c)^2 \Delta t \sigma_e^4}, \quad (4)$$

where  $\sigma_e$  can be approximated with the root-mean-square of the trace computed in the noise window. The estimator of the total energy fluence at the antenna position is given by  $\hat{f}_{\text{tot}} = \sum_{\text{pol}} \hat{f}_{\text{pol}}$ , and its uncertainty is obtained by propagating the errors  $\delta(\hat{f}_{\text{pol}})$ .

#### 4. Rice-distribution method

Unlike the relatively straightforward noise subtraction method, the Rice-distribution method comprises several steps. Here, we summarize the general logic behind the method, with detailed explanations and discussion provided in the subsequent sections.<sup>1</sup>

For each component of the measured electric field at a given antenna position, we define a signal window and  $N$  noise windows along the time trace (see Fig. 3). Since our method relies on estimating the energy fluence in the frequency domain, we ensure a correct and efficient evaluation by applying *windowing* before performing any Fourier transform (Section 4.1). This results in a signal-window frequency spectrum of  $M$  Rice-distributed spectral amplitudes, and  $N$  noise frequency spectra, each having  $M$  spectral amplitudes assumed to be Rayleigh-distributed (Section 4.2). We express the measured fluence as the sum of  $M$  frequency-dependent contributions. For each frequency, we estimate the noise fluence over the  $N$  windows and, finally, estimate the signal fluence and its uncertainty. A summary of the formulas used is provided in Section 4.3, while the derivation of these formulas - obtained by exploiting the statistical background based on the Rice distribution - can be found in Appendix A. In Section 4.4, we study the bias of the signal fluence estimator and the coverage of its error by performing a toy Monte Carlo for a single frequency. Finally, by summing up the  $M$  signal estimators and propagating the errors, we obtain the estimator of the polarization fluence and its uncertainty. The logic is illustrated in the flowchart of Fig. 4.

To finally estimate the total fluence at the antenna position and its uncertainty, we repeat this process for the remaining polarizations.

##### 4.1. Signal and noise windows

When calculating a discrete Fourier transform (DFT), we implicitly assume that the finite sequence of considered samples is periodic in time. If the start and end points of the sequence do not match each

<sup>1</sup> The implementation of the algorithm can be found at <https://gitlab.iap.kit.edu/saramartinelli/fluence-and-uncertainty-estimation-based-on-rice-distribution>.

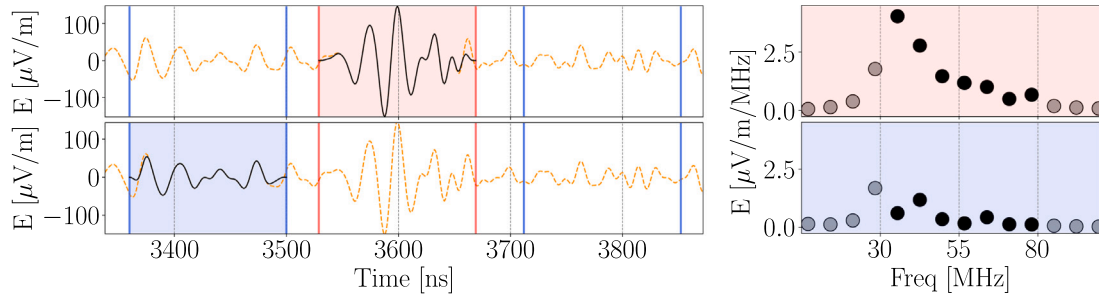


Fig. 5. Examples of signal window (upper plots) and noise window (lower plots) as defined for the Rice-distribution method. On the left, we show the time trace before (orange) and after (black) applying the Tukey function in the signal window (red) and the noise window (blue). On the right, the corresponding frequency spectra are shown using the same color code, with the black dots indicating the spectral amplitudes included in the 30–80 MHz frequency band.

other, artificial spectral contributions are introduced in the frequency spectrum. To reduce the spectral leakage, it is recommended to apply a tapering function [20]. In the implementation of our method, anytime we perform a DFT of a sequence of samples, we employ a fast Fourier transform (FFT) algorithm, and first we apply in the time domain the so-called *Tukey window* [21]. This is a symmetrical function consisting of a rectangular function combined with two halves of a *Hann window*, also known as *raised cosine bell window* [20]. For this reason, the Tukey window is often referred to as *split cosine bell* or *cosine-tapered window*. Given a Tukey window of total length  $L_{\text{tot}} = N_{\text{tot}} \Delta t$ , where  $N_{\text{tot}}$  is the total number of the samples of the considered sequence, one can require the proportion  $p$  of the sequence to be tapered by the two halves of the Hann window. In other words:

$$L_H = p L_{\text{tot}} = N_H \Delta t, \quad \text{with } N_H = p N_{\text{tot}}, \quad (5)$$

where  $L_H$  is the length of the Hann window and  $N_H$  is the total number of samples covered by its two halves.

To properly define the **signal window**, we employ the Tukey window just introduced. First, we fix  $L_{\text{tot}} = 140$  ns, and  $L_H = 40$  ns, then, we clip the time trace around the pulse position  $t_{\text{peak}}$  such that the resulting clipped trace has a length equal to  $L_{\text{tot}}$ . Finally, we apply the Tukey window and we perform the FFT of the windowed trace. Because of the usage of windowing, the time trace is damped at the edges of the signal window, where usually the noise is stronger than the signal (see Fig. 5). The resulting spectrum could present a non-negligible contribution outside the sensitive frequency band of interest, mainly due to the noise. For this reason, these spectral amplitudes are not considered. To evaluate the noise level of the measurement, we now define  $N$  **noise windows**. Starting from the beginning of the trace, we apply the Tukey window every multiple of  $L_{\text{tot}}$ , until we cover the entire trace length. We take care of skipping the signal window and applying additional spacing to reduce the signal contribution to the noise evaluation. In this work, we use a spacing of  $\pm 20$  ns and the same  $L_{\text{tot}}$  and  $L_H$  employed for the signal window. This results in having about  $N \approx 60$  noise windows per measured trace. We perform the FFT in each noise window and we exclude the spectral amplitudes outside the sensitive frequency band of interest, as done in the signal window. As shown in Fig. 5, we get  $M = 7$  spectral amplitudes within the 30–80 MHz bandwidth both for the signal window and each of the noise windows.

#### 4.2. Rice distribution

After defining the signal window and performing the FFT, we obtain  $M$  measured spectral amplitudes  $a(v_j)$ , where the index  $j$  runs over  $M$ . These amplitudes are given by the superposition of the true signal and the measurement noise of each frequency bin. To recover the unknown spectral amplitudes of the true signal  $s(v_j)$ , we need to use a formalism that allows us to properly disentangle the phase information from the amplitude. To do this, the phasors formalism can be adopted.

Let us first focus on a single-frequency bin, where  $\vec{s}$  is a constant phasor representing the true signal, and  $\vec{m} = a e^{i\theta}$  is the measurement phasor, having amplitude  $a$  and phase  $\theta$ . Following the derivation of Ref. [22], we express the noise contribution to the measurement as the sum of  $\mathcal{N}$  random distributed phasors. The random phasors represent the elementary and monochromatic disturbances to the true signal we aim to estimate. Since many noise sources can contribute to the disturbances (e.g. narrow-band TV transmitters, electronics noise, atmospheric disturbances, broadband galactic background, and so on), it is reasonable to assume  $\mathcal{N}$  to be large. We make further assumptions on the noise phasors. We assume that the amplitude and phase characterizing them are statistically independent. Furthermore, we assume that their amplitudes are identically distributed and that their phases are uniformly distributed between  $[-\pi, \pi]$ . According to these conditions, it can be demonstrated that the marginal probability density function (PDF) of the length of the resultant phasor  $\vec{n}$ , given by the noise random phasors' sum, is Rayleigh<sup>2</sup> distributed. According to [22], we can express our measurement  $\vec{m}$  as the sum of  $\vec{s}$  and  $\vec{n}$  (as depicted in Fig. 6), whose amplitude  $a$  follows the marginal PDF given by the Rice distribution:

$$p_a(s, \sigma) = \begin{cases} \frac{a}{\sigma^2} \cdot \exp\left(-\frac{a^2 + s^2}{2\sigma^2}\right) \cdot I_0\left(\frac{as}{\sigma^2}\right) & a > 0 \\ 0 & a \leq 0, \end{cases} \quad (6)$$

In Eq. (6),  $s$  is the amplitude of the signal,  $\sigma$  is the scale parameter of the function and represents the noise level of the measurement, and  $I_0$  is the modified Bessel function of the first kind and order zero. When the signal is large compared to the noise, the Rice distribution can be approximated with a normal distribution having a standard deviation equal to  $\sigma$  and centered on  $\sqrt{s^2 + \sigma^2}$ . For low signals, the Rice distribution approaches a Rayleigh distribution. In our case, each of the  $M$  amplitudes measured in the signal window will follow a different Rice distribution depending on the unknown underlying signal and noise level of the frequency bin. Note that the described formalism is valid both in the time and frequency domains, thus the development of further applications is possible in both domains. We refer the reader to Appendix B if interested in the method we have developed to estimate the amplitudes  $s(v_j)$  and their errors based on maximizing the Rice likelihood function. In Appendix B, we also describe how to estimate the noise level of the bin over the  $N$  noise windows.

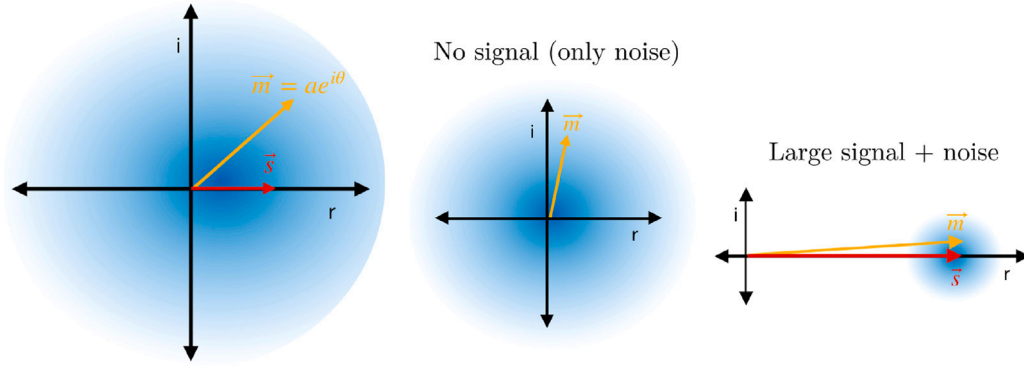
#### 4.3. Energy fluence estimator and its uncertainty

To evaluate the energy fluence in the frequency domain, we recall Parseval's theorem:

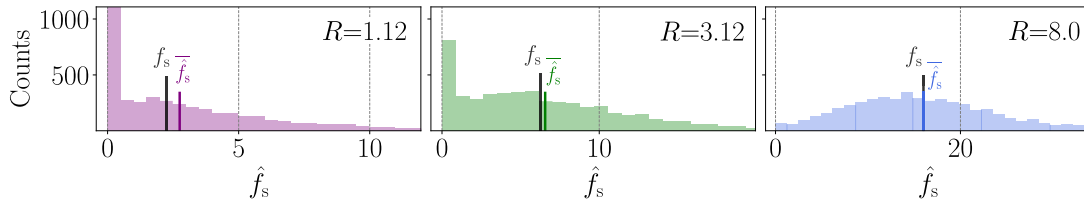
$$f_{\text{pol}} = \epsilon_0 c \Delta t \sum_{j=0}^{N_s-1} E_{\text{pol}}^2(t_j) \approx 2 \epsilon_0 c \Delta v \Delta t^2 \sum_{j=0}^{M-1} |D(v_j)|^2, \quad (7)$$

<sup>2</sup> The Rayleigh distribution is a  $\chi$  distribution with two degrees of freedom after the rescaling of the random variable by a constant factor.





**Fig. 6.** Representation of radio measurements using the phasor formalism. In the images, the signal phasor  $\vec{s}$  is depicted on the real axis (red). The measurement  $\vec{m}$  (orange) is given by the sum of  $\vec{s}$  and the sum of the  $\mathcal{N}$  random noise phasors. The probability density function of the noise phasor is illustrated as a blue cloud.



**Fig. 7.** Histograms of the signal fluence estimators obtained by running a toy Monte Carlo for three fixed values of  $R$ . The vertical black lines indicate the underlying true signal fluence used in the toy Monte Carlo. For comparison, the mean values of the histograms are shown through the colored vertical lines.

where  $N_s$  is the number of samples considered belonging to the time trace,  $D$  are the corresponding complex-valued DFTs,  $M = \frac{N_s}{2}$  is the number of positively valued frequencies,  $\Delta\nu$  is the frequency sample size, and the time sample size is given by  $\Delta t = \frac{1}{N_s \Delta\nu}$ . We express Eq. (7)<sup>3</sup> as a sum of the fluences for each frequency:

$$f(v_j) = K |D(v_j)|^2 \rightarrow f_{\text{pol}} = \sum_{j=0}^{M-1} f(v_j), \quad (8)$$

where  $K = 2 \epsilon_0 c \Delta\nu \Delta t^2$ . Let us now consider the ideal case of a signal in the complete absence of noise. The energy fluence of the unknown signal is:

$$f_s = K \sum_{j=0}^{M-1} s^2(v_j) = \sum_{j=0}^{M-1} f_s(v_j). \quad (9)$$

The estimator of the polarization fluence and the estimator of the total fluence at the antenna position will be of the form:

$$\hat{f}_{\text{pol}} = \sum_{j=0}^{M-1} \hat{f}_s(v_j) \rightarrow \hat{f}_{\text{tot}} = \sum_{\text{pol}} \hat{f}_{\text{pol}}, \quad (10)$$

where  $\hat{f}_s(v_j)$  are the signal fluence estimators in the  $j$ th frequency. The fluence of the same signal measured in the presence of random noise will be:

$$\hat{f}_a = K \sum_{j=0}^{M-1} a^2(v_j) = \sum_{j=0}^{M-1} \hat{f}_a(v_j), \quad (11)$$

with  $\hat{f}_a(v_j)$  being the frequency-dependent contributions to the amplitude fluence measured in the signal window.

For simplicity, let us analyze a single-frequency bin. We estimate the noise fluence of the  $j$ th frequency by exploiting the  $N$  noise windows:

$$\hat{f}_n(v_j) = \frac{K}{N} \sum_{i=0}^{N-1} n_i^2(v_j), \quad (12)$$

<sup>3</sup> In Eq. (7), we approximate  $\epsilon_0 c \Delta\nu \Delta t^2 (|D(v_0)|^2 + 2 \sum_{j=1}^{M-2} |D(v_j)|^2 + |D(v_{M-1})|^2)$  by double counting the contribution from the first and the last bin, which are typically zero due to bandpass filtering.

where  $n_i(v_j)$  indicates the Rayleigh-distributed random noise of the  $j$ th bin measured in the  $i$ th noise window. In Eq. (12), we use the sample mean over the windows, but depending on the characteristic noise of the measurements, one could consider excluding some of the noise windows from the evaluation, or using a more robust estimator as the one of Eq. (24). Finally, we define the fluence estimator of the signal in the  $j$ th frequency by subtracting the noise fluence from the amplitude fluence estimators:

$$\hat{f}_s(v_j) = \begin{cases} \hat{f}_a(v_j) - \hat{f}_n(v_j) & \hat{f}_a(v_j) \geq \hat{f}_n(v_j) \\ 0 & \hat{f}_a(v_j) < \hat{f}_n(v_j). \end{cases} \quad (13)$$

The statistical uncertainty can be estimated as:

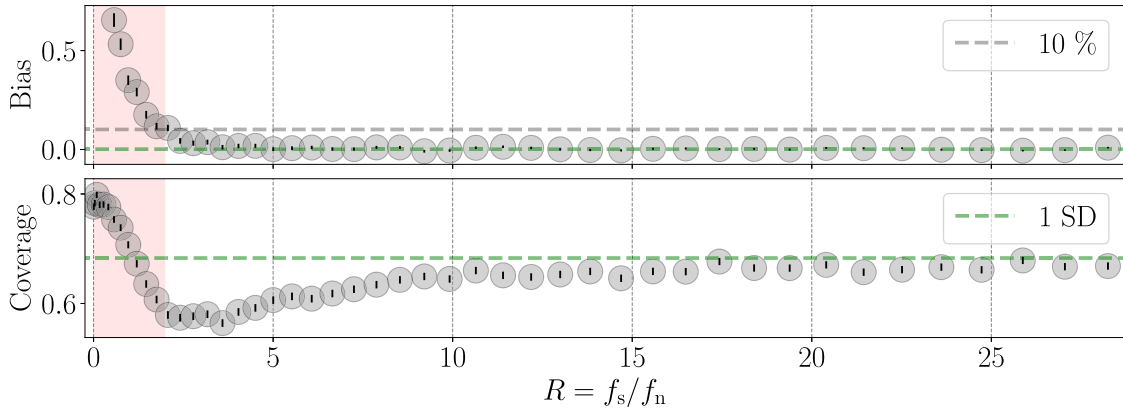
$$\delta(\hat{f}_s(v_j)) = \sqrt{\hat{f}_n(v_j) (\hat{f}_n(v_j) + 2 \hat{f}_s(v_j))} \quad (14)$$

Finally, the fluence estimator of the polarization and the total fluence estimator can be calculated, as well as their uncertainties.

#### 4.4. Toy Monte Carlo: Bias and error coverage for single-frequency estimator

In Eq. (13), we included the second condition to prevent the signal estimator of a frequency bin from assuming negative unphysical values. However, including this condition introduces a bias, that we evaluated with a toy Monte Carlo. Exploiting the fact that the random variables  $a$  are Rice-distributed, we generate  $N_{\text{MC}} = 5000$  random variables by setting  $s$ , and thus  $f_s$ , to fixed values. Without loss of generality, we keep  $\sigma = 1$  for simplicity and we define the noise parameter  $f_n := 2 \sigma^2$ . For each amplitude  $a$ , we evaluate the noise fluence estimator  $\hat{f}_n$  as the mean over  $N$  random variables following a normal distribution having mean  $\mu = 2 \sigma^2$  and variance  $\text{Var} = 4 \sigma^4 / N$ , where  $N$  is the number of noise windows used in this work (see Appendix A for a detailed explanation). Finally, we evaluate the signal fluence estimator  $\hat{f}_s$  and its uncertainty.

In Fig. 7, we show some examples of estimator distributions generated with the toy Monte Carlo for three fixed values of  $R = 1.12, 3.12, 8.0$ , the signal-to-noise ratio of the frequency bin given by  $R = f_s / f_n$ . In the histograms relative to  $R = 1.12, 3.12$ , the presence of a



**Fig. 8.** Intrinsic bias of the signal fluence estimator (top) and errors coverage (bottom) obtained by running a toy Monte Carlo for several values of  $R$ . The orange shadowed area indicates the  $R$  range which is strongly biased. In both plots, the error bars are divided by  $\sqrt{N_{\text{bin}}}$ , with  $N_{\text{bin}}$  being the number of entries in the bin.

peak in correspondence of the bin including the estimators  $\hat{f}_s = 0$  can be noticed. On the contrary, in the histogram relative to  $R = 8.0$ , the number of entries in the same bin decreases significantly. This reflects a bias that gets less prominent with increasing values of  $R$ . We tackled the dependence on  $R$  by scanning several values and computing the relative intrinsic bias and its standard deviation as:

$$\psi = \overline{\hat{f}_s} / f_s - 1, \quad \sigma_\psi = \sqrt{\sum_{i=0}^{N_{\text{MC}}-1} \frac{(\hat{f}_{s,i} - \overline{\hat{f}_s})^2}{N_{\text{MC}} - 1}}, \quad (15)$$

where  $\overline{\hat{f}_s}$  is the mean value of the estimators obtained through the toy Monte Carlo. As shown in the upper plot of Fig. 8, the relative bias is larger than 10% up to  $R = 2$ . For higher values of  $R$ , the bias decreases and can be neglected. The bias of the single-frequency estimator seems to affect the bias of the fluence estimator of a single polarization mostly up to  $\text{SNR} \approx 2.5$  (see Section 5.1). We studied also the coverage of the errors as a function of  $R$ . We define the coverage as the percentage of data satisfying the condition:

$$f_s \in [\hat{f}_s - \delta(\hat{f}_s), \hat{f}_s + \delta(\hat{f}_s)]. \quad (16)$$

As shown in the lower plot of Fig. 8, in correspondence of large bias, we overestimate the errors, while for  $R > 2$  the coverage approaches the classical definition of standard deviation.

## 5. Validation of the method

In this section, we validate the Rice-distribution method and compare it to the noise subtraction method. Both methods are applied to simulated noisy electric field traces (Section 2.3). No additional signal cleaning or RFI suppression is applied. The fluence estimators are compared to the reference values evaluated from the same electric fields simulated in the absence of noise (Section 2.2). In this way, the analysis is not affected by biases potentially introduced by the unfolding of the antenna response. Since the methods differ, we compare the estimators to different reference values. Concerning the noise subtraction method, the reference values are calculated in the time domain by evaluating Eq. (2) in the signal window. Because of the usage of windowing in the Rice-distribution method, this equation would yield a different value.<sup>4</sup> Instead, the reference values are obtained evaluating Eq. (8) over the frequency spectrum of the signal window after applying the Tukey window and excluding the frequencies outside the 30–80 MHz band. In the following, error coverage, bias, and uncertainty for both methods are analyzed as functions of the SNR.

<sup>4</sup> For the window function adopted, we estimate an average difference of 2% for antennas within 1.5 Cherenkov radii.

For a single polarization, we define the signal-to-noise ratio as:

$$\text{SNR}_{\text{pol}} = \left( \frac{|E_{\text{pol}}^{\text{Hilb}}(t_{\text{peak}})|}{\text{RMS}_{\text{pol}}} \right)^2, \quad (17)$$

where  $t_{\text{peak}}$  is the estimated pulse position,  $E_{\text{pol}}^{\text{Hilb}}(t_{\text{peak}})$  is the amplitude of the Hilbert envelope of the electric-field component at the pulse position, and  $\text{RMS}_{\text{pol}}$  is the root-mean-square of the Hilbert envelope in the noise window. Similarly, we define the signal-to-noise ratio over the three polarizations as:

$$\text{SNR}_{\text{tot}} = \left( \frac{A_{\text{tot}}^{\text{Hilb}}(t_{\text{peak}})}{\text{RMS}_{\text{tot}}} \right)^2, \quad (18)$$

with  $\text{RMS}_{\text{tot}}$  being the root-mean-square in the noise window of the trace:

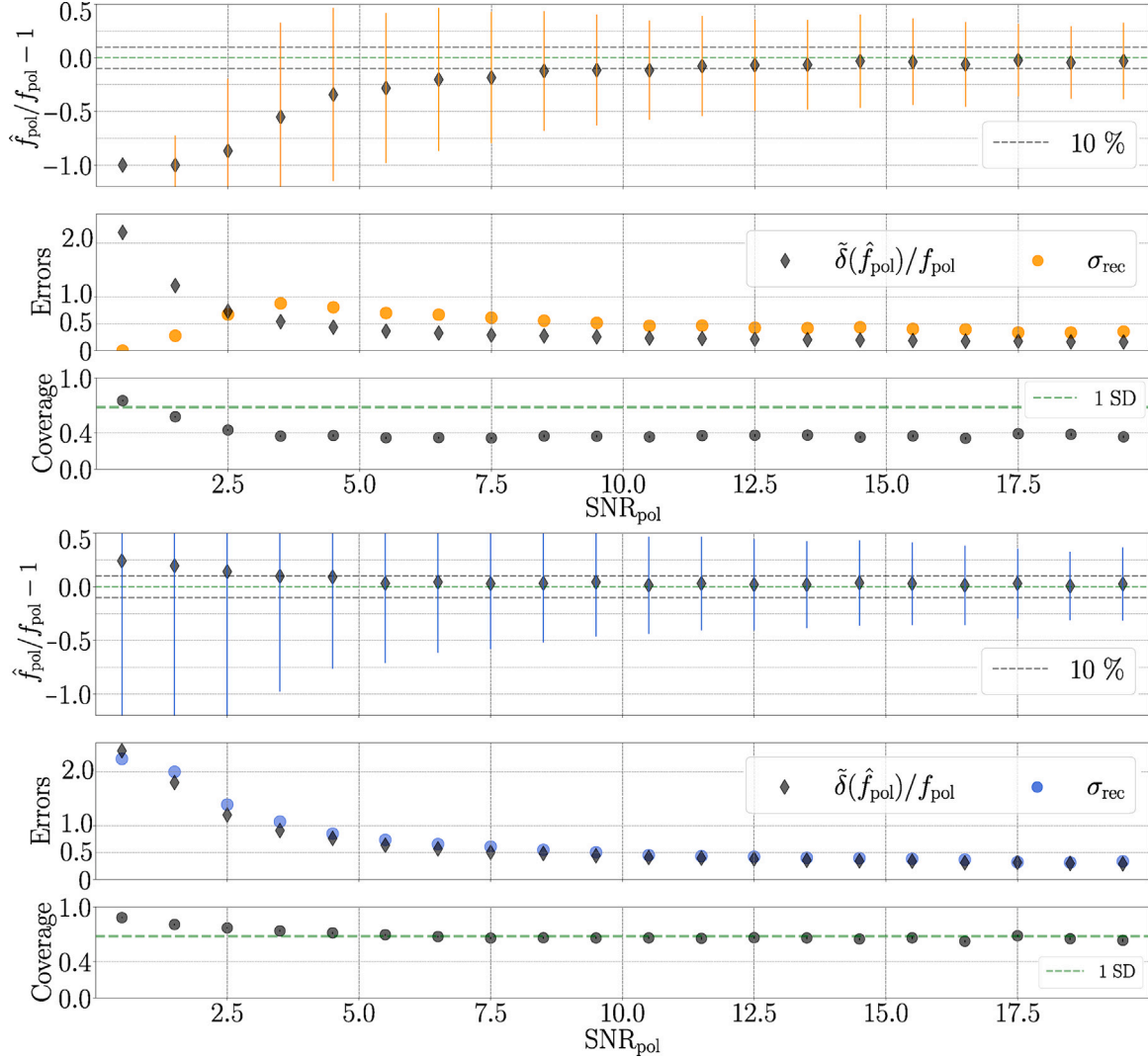
$$A_{\text{tot}}^{\text{Hilb}}(t) = \sqrt{\sum_{\text{pol}} |E_{\text{pol}}^{\text{Hilb}}(t)|^2}. \quad (19)$$

In real data, the position of pulses with a small signal-to-noise ratio is not known. It can, however, be estimated, for example by fitting a radio wavefront model [23] to pulses with sufficient signal-to-noise ratio. In this work, we assume that the approximate pulse position is known. To this end, we start from the Monte Carlo position  $t_{\text{peak}}^{\text{MC}}$  and then determine  $t_{\text{peak}}$  by evaluating the maximum of the Hilbert envelope around  $t_{\text{peak}}^{\text{MC}} \pm 30$  ns. Simulated traces strongly affected by thinning artifacts are excluded by applying a cut based on antenna position. Following [24], we exclude antennas with a distance from the shower axis greater than 2 Cherenkov radii.

### 5.1. Polarization fluence $f_{\text{pol}}$

For each electric field component, we evaluate the fluence estimators  $\hat{f}_{\text{pol}}$  and their uncertainties  $\delta(\hat{f}_{\text{pol}})$ ,<sup>5</sup> along with the reference values  $f_{\text{pol}}$  and the signal-to-noise ratios  $\text{SNR}_{\text{pol}}$ . After calculating the ratios  $\psi = \hat{f}_{\text{pol}} / f_{\text{pol}}$ , we examine different signal-to-noise ratio intervals. Electric fields are broken down into East-West (EW), North-South (NS), and Vertical (V) components. We combine the data irrespective of the polarization, collecting all ratios  $\psi$  within a given SNR range without distinguishing between EW, NS, and V components. To reduce the impact of outliers, in each SNR interval, we calculate  $\tilde{\psi}$ , the median value of  $\psi$ . Finally, we define the average reconstruction bias in the bin as  $\text{Bias} = \tilde{\psi} - 1$ . The dispersion of each bin is described by the interquartile range (i.q.r.) of the distribution. We calculate the equivalent standard deviation as  $\sigma_{\text{rec}} \approx \frac{\text{i.q.r.}}{1.35}$ .

<sup>5</sup> We used the estimator of Eq. (24) to estimate  $\hat{f}_n(v_j)$ , the noise fluence of the  $j$ th frequency bin.



**Fig. 9.** Estimation of  $f_{\text{pol}}$  for the noise subtraction method (upper figure) and the Rice-distribution method (lower figure) up to  $\text{SNR}_{\text{pol}} = 20$  (as defined in Eq. (17)). In the upper plots, we show the bias together with the error bar given by  $\sigma_{\text{rec}}$ . In the middle plots, the resolution of each bin is compared with  $\delta(\hat{f})/f$  of the same bin on a logarithmic scale. The coverage of the uncertainties (Eq. (20)), together with its error normalized to the number of entries of the bins, are shown in the lower plots.

As shown in Fig. 9, where SNR values up to 20 are investigated, we found that for the noise subtraction method, SNR values below 2.5 are characterized by an extremely large bias (exceeding 75%). In the first bin, the bias reaches 100%, as the negative-valued estimators have been set to zero, significantly influencing the result. For larger values, the bias decreases: starting from  $\text{SNR} \approx 9$ , the average bias is contained within 10%. In many works, SNR values of at least 10 are required for higher-level analyses [6–8]. In comparison, the Rice-distribution method exhibits a smaller bias. On average, up to  $\text{SNR} \approx 2.5$ , the bias fluctuates between 25% and 15%. The bias in this region is probably due to a combination of the estimator intrinsic bias discussed in Section 4.4, and the lack of RFI suppression. For increasing SNR values, the bias fluctuates between 10 and 5%, and can be neglected. We now focus on the uncertainty estimation of the two methods. We compare  $\tilde{\delta}(\hat{f}_{\text{pol}})/f_{\text{pol}}$ , the median value of the relative uncertainties of each bin, with  $\sigma_{\text{rec}}$ . For unbiased estimators, we would expect the uncertainties to reflect the reconstruction resolution. The uncertainties of the noise subtraction method result in being systematically underestimated, as shown in Fig. 9. To further corroborate this conclusion, we investigate

the uncertainty coverage of the method. We define the coverage as the percentage of data points satisfying the condition:

$$f_{\text{pol}} \in [\hat{f}_{\text{pol}} - \delta(\hat{f}_{\text{pol}}), \hat{f}_{\text{pol}} + \delta(\hat{f}_{\text{pol}})]. \quad (20)$$

The error coverage of the noise subtraction method oscillates between 30% and 40%. The coverage of the Rice-distribution method aligns with its bias. For SNR values above 2.5, where the bias is less significant, the coverage fluctuates around 68%, close to the classical definition of one standard deviation (SD) coverage.

We found that for higher values of SNR, both methods converge to a reconstruction bias below 5%, as shown in Fig. 10, where the region up to  $\text{SNR} = 400$  is studied. For a fair comparison, in all of the plots of the figure, we exclude the data points corresponding to  $\text{SNR} < 10$ , where the noise subtraction method is strongly biased. Once more, the noise subtraction method systematically underestimates the uncertainties. The average coverage is about 40%, while our method provides a more consistent way of uncertainty estimation, with an error coverage slightly below the desired 68%.

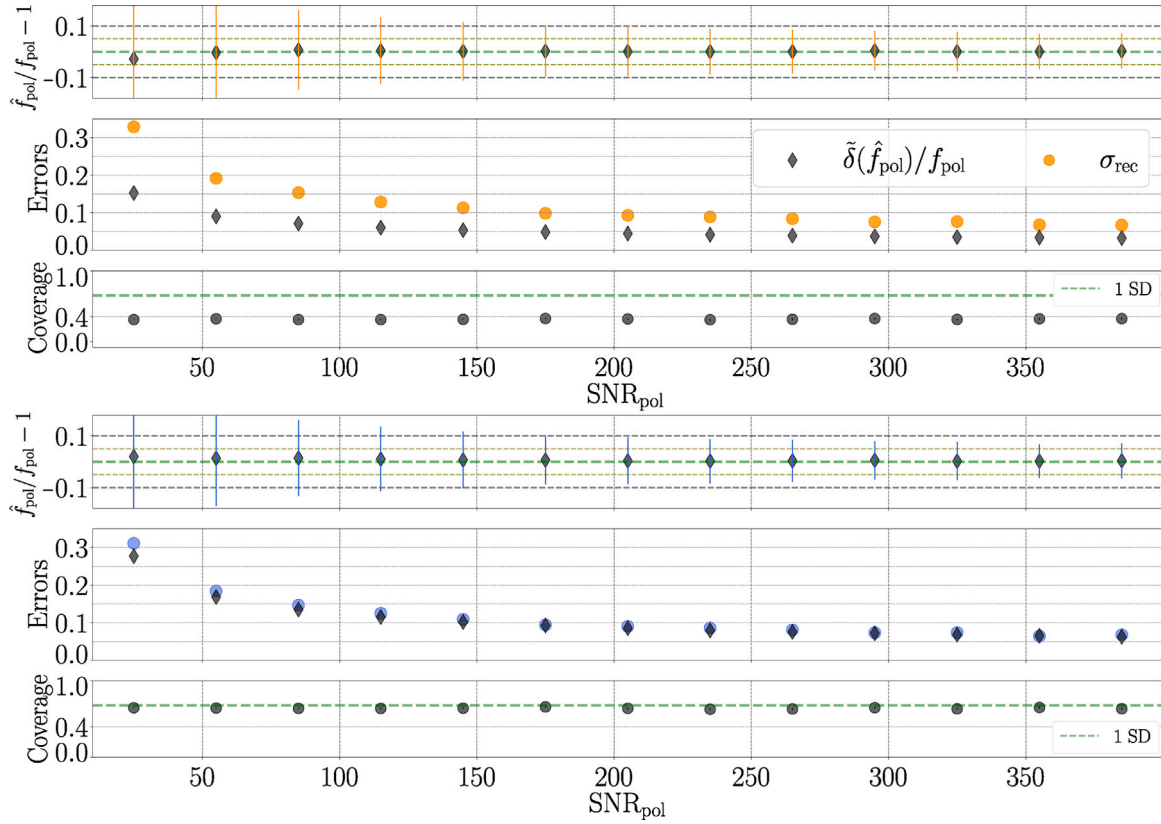


Fig. 10. Estimation of  $f_{\text{pol}}$  for the noise subtraction method (upper figure) and the Rice-distribution method (lower figure) up to  $\text{SNR}_{\text{pol}} = 400$  (as defined in Eq. (17)).

## 5.2. Total fluence $f_{\text{tot}}$

We repeat the analysis presented in the previous section, adapting it to the estimation of the total fluence evaluated at the antenna position. The noise subtraction method exhibits a large bias at the lowest SNR values, still exceeding 10% up to  $\text{SNR} \approx 6.5$ . From this value, the bias starts fluctuating around 10% and then stabilizes within 5% with increasing SNR (see Fig. 11). The uncertainties derived from the method are underestimated, with an average coverage of about 35%. We can draw the same conclusion for higher values of SNR, despite the bias remaining contained within 5% (see Fig. 12, where we show the SNR range from 10 up to 400). In contrast, the bias of the method we implemented fluctuates around 10% up to  $\text{SNR} \approx 7.5$ , and around 5% with increasing SNR (see Figs. 11, 12). On average, the error coverage fluctuates around 60% at any SNR value.

## 6. Discussion

In this article, we focused on an unbiased comparison of signal-estimation methods, using a generic approach and not tailoring their application to any specific experiment. The Rice distribution method shows promising improvements, especially for incorporating low-SNR signals. The extent of improvement achievable in the reconstruction of shower observables, such as primary particle energy or depth of shower maximum, will be dependent on the specifics of said reconstruction algorithm. For example, it may be limited by the reliability of the lateral distribution function at low SNR values. Other technical aspects should also be considered in higher-level analyses. For example, due to the power loss associated with windowing, careful selection of the window function, length, and tapering percentage is necessary. This step may require renormalizing fluences and radiation energies when relating them to particle energies. Lastly, if pulse position determination is inaccurate, it could introduce additional biases in the fluence

estimation. We note that none of these limitations are specific to the Rice method for signal estimation, which will yield a minimally biased estimate of the energy fluence and an adequate quantification of its uncertainty. The Rice method thus provides the potential to incorporate data from lower-SNR measurements in reconstruction procedures based on fluence estimation, when proper care is taken regarding limitations such as the ones described above.

## 7. Conclusions

In this work, we quantified the bias at low SNR values in the noise-subtraction method, which is conventionally used to reconstruct the energy fluence. We also quantified the underestimation of the uncertainties associated with this method. To address these known limitations, we developed a method based on Rice distributions for a more accurate estimation of the fluence at low SNR values and a correct evaluation of its uncertainty. At the antenna level, we achieve an average bias of 10% up to signal-to-noise ratios of about 7.5. Above this value, the bias fluctuates around 5%. We also significantly improved the estimation of the uncertainties, reaching a reliable coverage close to 68%. With the Rice-distribution method, the estimation of energy fluence and its uncertainty can thus be significantly improved, and signal-to-noise cuts can be lowered or possibly completely avoided in higher-level analyses.

The Rice-distribution method is generic and can be applied to any kind of radio observations of air showers and in-ice cascades. Since the method is not constrained by the frequency bandwidth of sensitivity, radio experiments in the field of astroparticle physics can make use of it. Among others, the Pierre Auger Observatory, RNO-G [25], LOFAR, ANITA, GRAND [26], SKA [27], and the IceCube Neutrino Observatory [28] could benefit from adopting this approach. The statistical background employed to develop the fluence estimation method can be further exploited, as we show in Appendix B. There, we describe



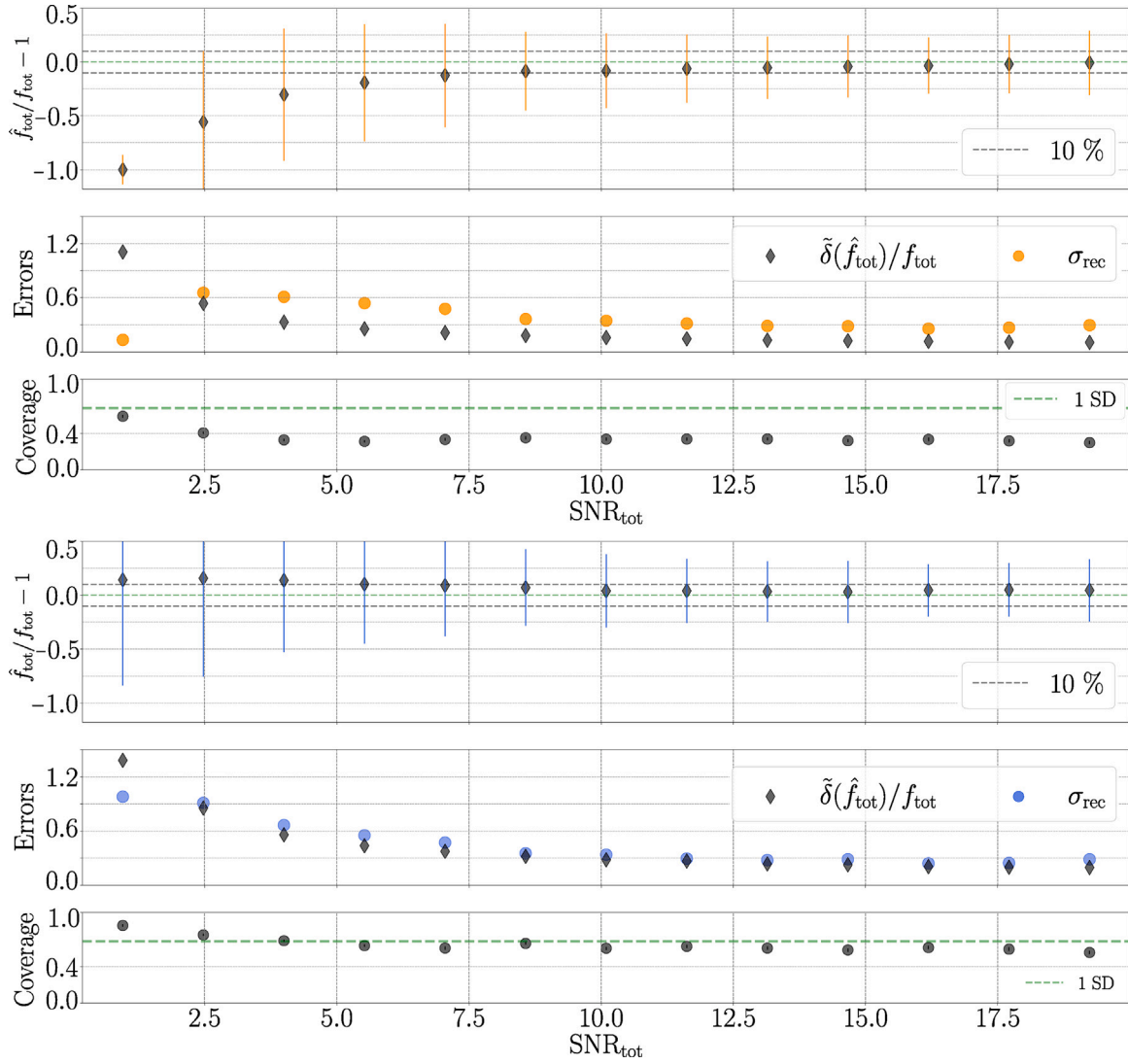


Fig. 11. Estimation of  $f_{\text{tot}}$  for the noise subtraction method (upper figure) and the Rice-distribution method (lower figure) up to  $\text{SNR}_{\text{tot}} = 20$  (as defined in Eq. (18)).

how to estimate the signal amplitude employing the Rice likelihood function. Further improvements and applications are still feasible by combining this statistical approach and the signal modeling in the frequency domain.

#### CRediT authorship contribution statement

**Sara Martinelli:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Tim Huege:** Writing – review & editing, Supervision, Methodology, Investigation, Funding acquisition, Conceptualization. **Diego Ravignani:** Writing – review & editing, Validation, Supervision, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Harm Schoorlemmer:** Writing – review & editing, Software, Methodology, Investigation, Conceptualization.

#### Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Tim Huege is a receiving editor for the journal, but will not be involved in the reviewing of this article. If there are other authors, they declare that they have no known competing financial interests or personal

relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgments

The simulations used in this work were processed through Offline, the reconstruction framework of the Pierre Auger collaboration. We thank the collaboration for allowing us to use the detector simulation implemented within the framework and for providing us access to their radio background measurements. We would also like to thank our colleagues in the Pierre Auger Collaboration for fruitful discussions and helpful input. We thank the Hermann von Helmholtz-Gemeinschaft Deutscher Forschungszentren e.V. for supporting Sara Martinelli through the Helmholtz International Research School for Astroparticle Physics (HIRSAP), part of *Initiative and Networking Fund*, Germany (grant number HIRS-0009).

#### Appendix A. Derivation of the fluence estimator and its uncertainty

We derive the energy fluence estimator and its uncertainty as presented in Section 4.3. For a monochromatic signal with  $K = 1$ , we rewrite Eqs. (9), (11) as:

$$f_s = s^2(v_0) \equiv s^2, \quad \hat{f}_a = a^2(v_0) \equiv a^2. \quad (21)$$

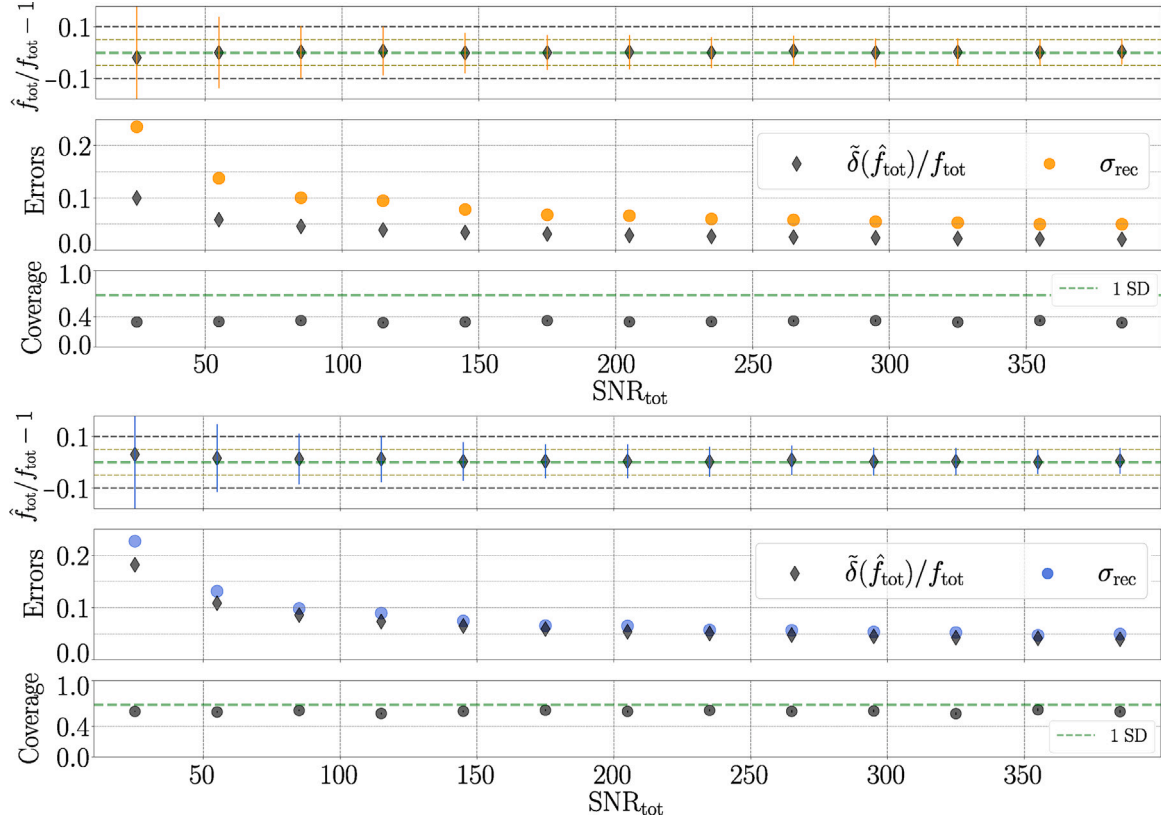


Fig. 12. Estimation of  $f_{\text{tot}}$  for the noise subtraction method (upper figure) and the Rice-distribution method (lower figure) up to  $\text{SNR}_{\text{tot}} = 400$  (as defined in Eq. (18)).

We aim for an estimator of  $f_s$  proportional to  $a^2$ . We standardize  $a$  as  $b = a/\sigma$ , where  $b \sim p_a(s/\sigma, 1)$ , equivalent to a non-central  $\chi^2$ -distribution with two degrees of freedom (DF) and non-centrality parameter  $\lambda = s/\sigma$ . Thus,  $b^2$  follows a non-central  $\chi^2$ -distribution:

$$b^2 \sim \chi_{\text{nc}}^2(\text{DF} = 2, \lambda = (s/\sigma)^2), \quad (22)$$

with  $E(b^2) = 2 + (s/\sigma)^2$  and  $\text{Var}(b^2) = 2(2 + 2(s/\sigma)^2)$ .

#### Noise fluence estimator

Estimating the signal fluence requires a noise fluence estimator. We sample the noise trace in  $N$  windows, each yielding a noise amplitude  $n_i$  assumed to be Rayleigh-distributed. Each  $(n_i/\sigma)^2$  follows a chi-square distribution with two DF, so the sum  $T = \sum_{i=0}^{N-1} (n_i/\sigma)^2$  follows a chi-square distribution with  $2N$  DF, approximated by a normal distribution with mean  $\mu = 2N$  and variance  $\text{Var} = 4N$  for large  $N$ . The noise fluence estimator is the sample mean of the squared noise amplitudes:

$$\hat{f}_n = \frac{1}{N} \sum_{i=0}^{N-1} n_i^2 = \frac{\sigma^2}{N} T, \quad (23)$$

which is normally distributed with mean  $\mu = 2\sigma^2$  and variance  $\text{Var} = \sigma^4/N$ . Defining  $f_n := 2\sigma^2$ , we find  $\hat{f}_n$  to be an unbiased estimator of  $f_n$ . For  $N \approx 60$ , its relative uncertainty is around 13%. We derive a more robust estimator in the presence of outliers using the population median, as:

$$\hat{f}_n = \frac{2}{1.405} \cdot \text{median}[n_i^2], \quad (24)$$

where we exploit that the median value of a  $\chi_{\text{DF}}^2$  distribution can be approximated as  $\text{DF} \cdot (1 - \frac{2}{9 \cdot \text{DF}})^3$ , and that  $n_i^2 \sim \sigma^2 \cdot \chi^2$ .

#### Signal fluence estimator

We derive the expected value of  $\hat{f}_a$  as:

$$E(\hat{f}_a) = E(a^2) = \sigma^2 E(b^2) = 2\sigma^2 + s^2 \quad (25)$$

By setting  $s = 0$ , for the noise estimator we get:

$$E(\hat{f}_n) = 2\sigma^2 := f_n. \quad (26)$$

Estimating the signal fluence as  $\hat{f}_s = \hat{f}_a - \hat{f}_n$ , its mean value will be:

$$E(\hat{f}_s) = E(\hat{f}_a) - E(\hat{f}_n) = s^2 = f_s, \quad (27)$$

meaning that such an estimator would be unbiased. Since the energy fluence has to be positively defined, we have to introduce the condition:

$$\hat{f}_s = 0 \quad \hat{f}_a < \hat{f}_n. \quad (28)$$

As discussed in Section 4.4, the above condition introduces a positive bias that can be neglected for  $R > 2$ .

#### Signal fluence uncertainty

We derive the variance of  $a^2$ :

$$\text{Var}(a^2) = \sigma^4 \text{Var}(b^2) = 2\sigma^2(2\sigma^2 + 2s^2). \quad (29)$$

To derive the variance of  $\hat{f}_s$ , we make the assumptions:

- The noise fluence is measured with much better precision than  $a^2$ . This can be considered a valid assumption since the variable  $a$  is a single measurement (from the signal window), while the noise fluence is estimated through  $N \approx 60$  noise windows;
- The probability of hitting the physical limit of Eq. (28) is negligible, i.e. we are assuming  $R > 2$ ;
- $(s/\sigma^2)$  is large, meaning that  $a^2$  and  $b^2$  are normal in a good approximation, and, again  $R$  is large.

Under these assumptions, the variance of  $\hat{f}_s$  is:

$$\text{Var}(\hat{f}_s) = \text{Var}(a^2) \approx \hat{f}_n (\hat{f}_n + 2 \hat{f}_s), \quad (30)$$

where we approximate the parameters as  $s^2 \approx \hat{f}_s$  and  $2\sigma^2 \approx \hat{f}_n$ . We define the confidence interval as  $\hat{f}_s \pm \delta(\hat{f}_s)$ , with:

$$\delta(\hat{f}_s) = \sqrt{\text{Var}(\hat{f}_s)} = \sqrt{\hat{f}_n (\hat{f}_n + 2 \hat{f}_s)}. \quad (31)$$

## Appendix B. Spectral amplitude estimator and its uncertainty

The theoretical background described in Section 4.2 has much more potential to exploit than what has been shown so far. Here we introduce the Rice likelihood function and provide an application example. In particular, we present a method to estimate the signal spectral amplitude and its uncertainty.

Let us consider the case of a single-frequency signal  $s_0 \equiv s(v_0)$ , with  $a_0 \equiv a(v_0)$  being the measured spectral amplitude, and  $\hat{\sigma}_0$  the estimator of the noise level. The Rice likelihood function will be:

$$L(s | a = a_0, \sigma = \hat{\sigma}_0) = \frac{a}{\sigma^2} \exp\left(-\frac{a^2 + s^2}{2\sigma^2}\right) I_0\left(\frac{as}{\sigma^2}\right), \quad (32)$$

where  $s \geq 0$ . We evaluate  $\hat{\sigma}_0$  over the  $N$  noise windows:

$$\hat{\mu}_0 = \frac{1}{N} \sum_{i=0}^{N-1} a_i(v_0) \rightarrow \hat{\sigma}_0 \equiv \mu_0 \sqrt{2/\pi}, \quad (33)$$

where in the last equivalence we assumed the noise to be Rayleigh-distributed with scale parameter  $\sigma_0$  and mean value  $\mu_0$ . We define the estimator of the parameter  $s_0$  as the amplitude  $\hat{s}_{\text{ML}}$  that maximizes  $L(s)$ . Since it is not possible to solve analytically  $\left.\frac{dL(s)}{ds}\right|_{s=\hat{s}_{\text{ML}}} = 0$ , we recommend the reader to run a scalar minimization algorithm to find the minimum of the *cost* function  $J(s) = -\ln(L(s))$ . To avoid unphysical solutions, such as negative results, and to make the minimization process reliable, we exploit a bounded solver. We set as bounds  $[a_0 - 1.5 \hat{\sigma}_0, a_0 + 1.5 \hat{\sigma}_0]$ .<sup>6</sup> There are some intervals where it is not strictly necessary to run the minimizer solver. In particular, for  $a_0/\hat{\sigma}_0 \leq \sqrt{2}$ ,  $\hat{s}_{\text{ML}} = 0$ . For larger ratios, as  $a_0/\hat{\sigma}_0 \geq 4$ , we can approximate  $L(s)$  by a normal distribution having SD =  $\hat{\sigma}_0$  and centered in  $\sqrt{a_0^2 - \hat{\sigma}_0^2}$ , thus  $\hat{s}_{\text{ML}} \approx \sqrt{a_0^2 - \hat{\sigma}_0^2}$ . To evaluate the estimator uncertainty  $\delta$ , we work in Gaussian approximation. First, we approximate the likelihood *cost* function as:

$$J(s) \approx J(\hat{s}_{\text{ML}}) + (s - \hat{s}_{\text{ML}}) \cdot \left.\frac{\partial J(s)}{\partial s}\right|_{s=\hat{s}_{\text{ML}}} + \frac{(s - \hat{s}_{\text{ML}})^2}{2} \cdot \left.\frac{\partial^2 J(s)}{\partial s^2}\right|_{s=\hat{s}_{\text{ML}}}, \quad (34)$$

where the first two terms are null by definition. Since the *cost* function of a Gaussian distribution is a parabolic function, we can approximate  $J(s)$  to a parabola as well. This would lead to the following system of equations:

$$\begin{cases} J(s) = (s - \hat{s}_{\text{ML}})^2 \cdot \frac{1}{2} \left.\frac{\partial^2 J(s)}{\partial s^2}\right|_{s=\hat{s}_{\text{ML}}} \\ \quad = \delta^2 \cdot \frac{1}{2} \left.\frac{\partial^2 J(s)}{\partial s^2}\right|_{s=\hat{s}_{\text{ML}}} \\ J(s) \approx k^2 \rightarrow \delta = k \cdot \text{SD} \end{cases} \quad (35)$$

Finally, by requiring a one-sigma interval, i.e.  $\delta = 1 \text{ SD}$ , the solution of Eq. (35) is given by:

$$\delta = 1 \sqrt{\frac{1}{2} \left.\frac{\partial^2 J(s)}{\partial s^2}\right|_{s=\hat{s}_{\text{ML}}}}. \quad (36)$$

<sup>6</sup> The bounds were found by studying the first and second derivative spaces of the likelihood function.

## Appendix C. Table of variables and notation

Generic variables	
$E_{\text{pol}}$	Electric field component in the chosen coordinates system
$f_{\text{pol}}$	Energy fluence relative to the considered electric field component
$f_{\text{tot}}$	Total energy fluence at the antenna position
$\text{SNR}_{\text{pol}}$	Signal-to-noise ratio relative to the considered electric field component
$\text{SNR}_{\text{tot}}$	Signal-to-noise ratio calculated over all the electric field components
Notation	
$y$	Unknown parameter (true value)
$\hat{y}$	Estimator of the parameter (or measurement)
$\delta(\hat{y})$	Uncertainty on the estimator (or measurement)
$\bar{h}$	Mean value of the variable
$\tilde{h}$	Median value of the variable
Rice-distribution variables (fixed frequency or time)	
$s$	Signal amplitude in the absence of noise
$a$	Measured amplitude of the signal in the presence of noise
$f_s, \hat{f}_s$	Signal energy fluence, its estimator
$f_n, \hat{f}_n$	Noise energy fluence, its estimator
$\hat{f}_a$	Signal energy fluence measured in the presence of noise

## Data availability

Code of the algorithm available at: <https://gitlab.iap.kit.edu/sar martinelli/fluence-and-uncertainty-estimation-based-on-rice-distribution>.

## References

- [1] T. Huege, Radio detection of cosmic ray air showers in the digital era, *Phys. Rep.* 620 (2016) 1–52.
- [2] F.G. Schröder, Radio detection of cosmic-ray air showers and high-energy neutrinos, *Prog. Part. Nucl. Phys.* 93 (2017) 1–68, <http://dx.doi.org/10.1016/j.pnpnp.2016.12.002>.
- [3] A. Aab, et al., Pierre Auger Collaboration Collaboration Collaboration, Measurement of the radiation energy in the radio signal of extensive air showers as a universal estimator of cosmic-ray energy, *Phys. Rev. Lett.* 116 (2016) 241101, <http://dx.doi.org/10.1103/PhysRevLett.116.241101>.
- [4] F. Schröder, On noise treatment in radio measurements of cosmic ray air showers, *Nucl. Instr. Methods A* 662 (2012) S238–S241, <http://dx.doi.org/10.1016/j.nima.2010.11.009>.
- [5] J. Glaser, Absolute Energy Calibration of the Pierre Auger Observatory using Radio Emission of Extensive Air Showers (Ph.D. thesis), RWTH Aachen University, 2017.
- [6] A. Abdul Halim, et al., Pierre Auger Collaboration Collaboration Collaboration, Radio measurements of the depth of air-shower maximum at the pierre auger observatory, *Phys. Rev. D* 109 (2024) 022002, <http://dx.doi.org/10.1103/PhysRevD.109.022002>, URL <https://link.aps.org/doi/10.1103/PhysRevD.109.022002>.
- [7] M. Gottowik, Pierre Auger Collaboration Collaboration, Measuring the muon content of inclined air showers using AERA and the water-cherenkov detector of the pierre auger observatory, in: Proceedings of 10th International Workshop on Acoustic and Radio EeV Neutrino Detection Activities — PoS, ARENA2024, Vol. 470, 2024, p. 033, <http://dx.doi.org/10.22323/1.470.0033>.
- [8] M. Büskens, Pierre Auger Collaboration Collaboration, Towards a cosmic-ray energy scale with the auger engineering radio array, in: Proceedings of 10th International Workshop on Acoustic and Radio EeV Neutrino Detection Activities — PoS, ARENA2024, Vol. 470, 2024, p. 035, <http://dx.doi.org/10.22323/1.470.0035>.

- [9] A. Aab, et al., Pierre Auger Collaboration Collaboration Collaboration, Energy estimation of cosmic rays with the engineering radio array of the pierre auger observatory, *Phys. Rev. D* 93 (2016) 122005, <http://dx.doi.org/10.1103/PhysRevD.93.122005>.
- [10] F. Schlüter, Expected Sensitivity of the AugerPrime Radio Detector to the Masses of Ultra-High-Energy Cosmic Rays Using Inclined Air Showers (Ph.D. thesis), Karlsruher Instituts für Technologie (KIT), 2022.
- [11] H. Schoorlemmer, et al., Energy and flux measurements of ultra-high energy cosmic rays observed during the first ANITA flight, *Astropart. Phys.* 77 (2016) 32–43, <http://dx.doi.org/10.1016/j.astropartphys.2016.01.001>.
- [12] N. Karastathis, et al., Using pulse-shape information for reconstructing cosmic-ray air showers and validating antenna responses with LOFAR and SKA, in: 38th International Cosmic Ray Conference (ICRC2023) - Cosmic-Ray Physics (Indirect, CRD), in: *Proceedings of Science*, vol. 444, Scuola Internazionale Superiore di Studi Avanzati (SISSA), 2023, pp. Art.-Nr.: 487, <http://dx.doi.org/10.22323/1.444.0487>, 51.13.04; LK 01.
- [13] T. Huege, Pierre Auger Collaboration Collaboration, The radio detector of the pierre auger observatory – status and expected performance, in: *EPJ Web Conf.*, Ultra High Energy Cosmic Rays, UHECR 2022, Vol. 283, 2023, <http://dx.doi.org/10.1051/epjconf/202328306002>.
- [14] F. Schlüter, Pierre Auger Collaboration Collaboration, Expected performance of the AugerPrime radio detector, in: *Proceedings of 9th International Workshop on Acoustic and Radio EeV Neutrino Detection Activities — PoS, ARENA2022*, Vol. 424, 2023, p. 028, <http://dx.doi.org/10.22323/1.424.0028>.
- [15] T. Huege, M. Ludwig, C.W. James, Simulating radio emission from air showers with CoREAS, *AIP Conf. Proc.* 1535 (1) (2013) 128.
- [16] D. Heck, J. Knapp, J.N. Capdevielle, G. Schatz, T. Thouw, CORSIKA: a Monte Carlo Code to Simulate Extensive Air Showers, FZKA Report 6019, Forschungszentrum Karlsruhe, 1998.
- [17] S. Ostapchenko, Monte Carlo treatment of hadronic interactions in enhanced pomeron scheme: QGSJET-II model, *Phys. Rev. D* 83 (014018) (2011) 128.
- [18] M. Kobal, A thinning method using weight limitation for air-shower simulations, *Astropart. Phys.* 15 (2001) 259–273.
- [19] P. Abreu, et al., Advanced functionality for radio analysis in the offline software framework of the pierre auger observatory, *Nucl. Instrum. Methods Phys. Res. Sect. A: Accel. Spectrometers Detect. Assoc. Equip.* 635 (1) (2011) 92–102, <http://dx.doi.org/10.1016/j.nima.2011.01.049>.
- [20] F. Harris, On the use of windows for harmonic analysis with the discrete Fourier transform, *Proc. IEEE* 66 (1) (1978) 51–83, <http://dx.doi.org/10.1109/PROC.1978.10837>.
- [21] J.W. Tukey, An introduction to the calculations of numerical spectrum analysis, in: B. Harris (Ed.), *Spectral Analysis of Time Series*, Wiley, New York, 1967, pp. 25–46.
- [22] J.W. Goodman, *Statistical Optics*, Wiley Classics Library, 2000.
- [23] A. Corstanje, et al., The shape of the radio wavefront of extensive air showers as measured with LOFAR, *Astropart. Phys.* 61 (2015) 22–31, <http://dx.doi.org/10.1016/j.astropartphys.2014.06.001>.
- [24] F. Schlüter, T. Huege, Signal model and event reconstruction for the radio detection of inclined air showers, *JCAP* 01 (2023) 008, <http://dx.doi.org/10.1088/1475-7516/2023/01/008>, [arXiv:2203.04364](https://arxiv.org/abs/2203.04364).
- [25] J. Aguilar, et al., Design and sensitivity of the radio neutrino observatory in greenland (RNO-G), *J. Instrum.* 16 (03) (2021) P03025, <http://dx.doi.org/10.1088/1748-0221/16/03/p03025>.
- [26] J. Álvarez-Muñiz, et al., The giant radio array for neutrino detection (GRAND): Science and design, *Sci. China Phys. Mech. Astron.* 63 (1) (2019) <http://dx.doi.org/10.1007/s11433-018-9385-7>.
- [27] E. de Lera Acedo, A. Faulkner, J. Bij de Vaate, SKA low frequency aperture array, in: 2016 United States National Committee of URSI National Radio Science Meeting, USNC-URSI NRSM, 2016, pp. 1–2, <http://dx.doi.org/10.1109/USNC-URSI-NRSM.2016.7436228>.
- [28] M. Oehler, R. Turcotte-Tardif, Development of a scintillation and radio hybrid detector array at the south pole, 2021, [arXiv:2107.09983](https://arxiv.org/abs/2107.09983).