



OBSERVATORY





DPG 2024, Karlsruhe 05.03.24

Radio signal and uncertainty estimation

S. Martinelli^{*}, Dr. T. Huege, Dr. D. Ravignani, Dr. H. Schoorlemmer for the Pierre-Auger collaboration





This talk: how to determine the signal in the presence of noise

Radio measurements of CR pulses have both an amplitude and a phase

- The CR signal and the random noise can add up constructively or destructively
- Recovering of the underlying signal for small signal-to-noise ratio (SNR) is non-trivial











This talk: how to determine the signal in the presence of noise

Radio measurements of CR pulses have both an amplitude and a phase

4

- The CR signal and the random noise can add up constructively or destructively
- Recovering of the underlying signal for small signal-to-noise ratio (SNR) is non-trivial

Energy fluence

- Energy deposit per unit area [eV/m²]
- Typically reconstructed by noise subtraction (Auger, LOFAR...)













• Method unbiased for SNR>10 (cut)

How does the method work?

• Uncertainties underestimated







Energy Fluence Reconstruction: Noise Subtraction Method

Definition of the search-signal

Definition of the noise window

Find the peak in the search-signal window

Definition of the signal window

 Fluence in the signal window

n-th pol.

The estimator can be negative!

 $f \sim sum over the amplitudes squared$

n-th pol.

+

n-th pol. Definition of the search-signal

Definition of the noise window

Find the peak in the search-signal window

Definition of the signal window

Fluence in the signal window

Subtraction of the noise fluence

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

 $\begin{array}{c} & \\ \hline \\ station \end{array} \right] \qquad \hat{f}^{s}$

$$\hat{f}^{\text{station}} = \sum_{\text{pol}} \hat{f}^{\text{pol}}$$

The estimator can be negative!

Can we do better than this?

- Radio measurements have both an amplitude and a phase
 - Our **measurement** can be expressed as the sum of **constant known phasor s** and a **random phasor sum** (Rayleigh-distributed noise).

random phasor sum =

- Radio measurements have both an amplitude and a phase $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$
 - Our **measurement** can be expressed as the sum of **constant known phasor s** and a **random phasor sum** (Rayleigh-distributed noise).

• The marginal p.d.f. of the measured **amplitude** is the Rice distribution.

$$p_A(a|s,\sigma) = \begin{cases} \frac{a}{\sigma^2} \cdot \exp\left(-\frac{a^2+s^2}{2\sigma^2}\right) \cdot I_0\left(\frac{as}{\sigma^2}\right) & a > 0\\ 0 & \text{otherwise} \end{cases}$$

 $a \sim \mathcal{R}ice(s, \sigma)$

Reference: Chapter 2.9 from J. W. Goodman, Statistical Optics (2015)

 $a \sim \mathcal{R}ice(s, \sigma)$

Reference: Chapter 2.9 from J. W. Goodman, Statistical Optics (2015)

Formalism valid for both time and frequency domains

How can we use this theory to build a better estimator of the fluence?

Fourier Transform

Ó

 $\dot{4}$

a [a.u.]

Derivation:

 $b=a/\sigma$ Change of variable by scaling the measured amplitude by the noise level of the bin | Adimensional

 $b=a/\sigma$ $snr=(s/\sigma)^2$ snr at bin level, to not be confused with the SNR of the full measurement!

Derivation:

$$E(a^2) = 2\sigma^2 + s^2 \qquad \longrightarrow \qquad \hat{f} = \begin{cases} a^2 - 2\sigma^2 & a^2 \ge 2\sigma^2 \\ 0 & \text{otherwise} \end{cases}$$
$$Var(a^2) = 2\sigma^4(2 + 2snr) \qquad \qquad Var(\hat{f}) = Var(a^2)$$

We have all the ingredients at bin level

We can now evaluate the fluence at trace level ...

... and at the antenna level

Now we can compare the new method and the noise-subtraction method

RD Simulations

- 8000 proton/iron/nitrogen/helium CORSIKA/CoREAS simulations
- RD simulation (antenna response folded and unfolded back)

RD Simulations

- RD traces recorded over one year from the stations of the Engineering Array
 - Cleaned from the showers signals and corrupted traces

RD Simulations

Simulated RD Measurements

> RD Noise Library

Different reference values for different methods

Quality cuts:

- Stations affected by thinning artifacts from simulation
- Stations where the pulse-finding algorithm "fails" (>2ns)

On average the new method is unbiased even at small SNR

Noise subtraction method

The relative errors are smaller than the reconstruction resolution of the same bin

The relative errors of the new method reflect better the reconstruction resolution

The noise subtraction method underestimates the uncertainties

Coverage of the errors fluctuates around 68%

52

New method

Noise subtraction method

The noise subtraction method strongly underestimates the uncertainties at any SNR.

53

Conclusions

- The new fluence estimation shows a smaller bias than the conventional method for small SNR values (on average less than 10%)
- At higher SNRs, the bias is comparable (on average less than 5%)
- The new method correctly estimate the uncertainties at any SNR (coverage about 68%)

... & Outlooks

- Paper soon
- SNR cut can be lowered/removed
- Improvements on the elm. Energy reconstruction
- Possible employment of frequency spectrum modeling

 $\mathbf{N=5000}$ $\left(snr_{\mathrm{MC}}, (a_{\mathrm{MC}}^{k})^{2}\right) \rightarrow \left(\hat{f}^{k}, \delta(\hat{f}^{k})\right)$

Repeating by fixing snr to several values

Noise subtraction method

Noise subtraction method

CR initiating EAS

elm. component => radio emission

Pierre Auger Observatory

AugerPrime Upgrade, Radio Detector (RD) Not in scale ;)

[eV m⁻²] to estimate the elm. Energy of the shower

Reference: C. Glaser's PhD thesis

5.5.2 Uncertainty of the energy fluence

We estimate the uncertainty of the energy fluence by assuming that the measured electric-field amplitude A(t) is the sum of the cosmic-ray radio pulse S(t) and noise e(t). Furthermore, we assume that the noise e(t) is Gaussian distributed with mean $\mu = 0$ and standard deviation $\sigma = \sigma_e$. The energy fluence of A is then given by the equation

$$f(A) = \epsilon_0 c \Delta t \sum_{t_1}^{t_2} A(t_i)^2 = \epsilon_0 c \Delta t \sum_{t_1}^{t_2} \left[S(t_i) + e(t_i) \right]^2 = \epsilon_0 c \Delta t \sum_{t_1}^{t_2} \left[S(t_i)^2 + 2S(t_i)e(t_i) + e(t_i)^2 \right]$$
(5.16)

and the expectation value of f(A) is

$$\langle f(A) \rangle = \epsilon_0 c \Delta t \sum_{t_1}^{t_2} A(t_i)^2 = \epsilon_0 c \Delta t \sum_{t_1}^{t_2} \left[\langle S(t_i)^2 \rangle + 2 \langle S(t_i)e(t_i) \rangle + \langle e(t_i)^2 \rangle \right]$$

$$= \epsilon_0 c \Delta t \sum_{t_1}^{t_2} \left[\langle S(t_i) \rangle^2 + \underbrace{Var(S(t_i))}_{=0} + 2 \langle S(t_i) \rangle \underbrace{\langle e(t_i) \rangle}_{=0} \right]$$

$$+ 2 \underbrace{Cov(S(t_i), e(t_i))}_{=0} + \underbrace{\langle e(t_i) \rangle^2}_{=0} + \underbrace{Var(e(t_i))}_{\sigma_e^2} \right]$$

$$(5.17)$$

$$= \epsilon_0 c \Delta t \sum_{t_1}^{t_2} \left[\langle S(t_i) \rangle^2 + \sigma_e^2 \right] \, .$$

Hence, the best estimate of the energy fluence of the radio signal S is indeed

$$f(S) = \epsilon_0 c \Delta t \sum_{t_1}^{t_2} \left[A(t_i)^2 - \sigma_e^2 \right]$$
(5.18)

as defined in Eq. (5.8) where σ_e^2 is also calculated from the electric-field trace in a part where no signal is present. Following a similar calculation we can estimate the uncertainty of f(S) by computing $\sigma_f^2 = Var(f) = \langle f^2 \rangle - \langle f \rangle^2$. After several lines of calculation it follows that

$$\sigma_f^2 = 4f \,\epsilon_0 c \,\Delta t \,\sigma_e^2 + 2 \,(\epsilon_0 c)^2 \,(t_2 - t_1) \,\Delta t \,\sigma_e^4 \,. \tag{5.19}$$