PROCEEDINGS OF SCIENCE

Filtering and aliasing effects produced by time-discretization in the simulation of Extensive Air Shower radio signals

Pengfei Zhang ,[1,] Jiale Wang,[1] Xin Xu,[1] Fufu Yang,[1] Chao Zhang,[2,3] Sergio Cabana-Freire[4,*] and Matias Tueros[5,6]**

[1]*School of Electronic Engineering, Xidian Univ, Xi'an 710071, China*
[2]*Department of Astronomy, School of Physics, Peking University, Beijing 100871, China*
[3]*School of Astronomy and Space Science, Nanjing University, 210023 Nanjing, Jiangsu, China*
[4]*Instituto Galego de Física de Altas Enerxías (IGFAE), Universidade de Santiago de Compostela, 15782 Santiago de Compostela, Spain*
[5]*Instituto de Física La Plata, CONICET-UNLP, Diagonal 113 entre 63 y 64 , La Plata, Argentina*
[6]*Depto. de Física, Fac. de Cs. Ex., Universidad Nacional de La Plata, Casilla de Coreo 67 (1900), La Plata, Argentina*

E-mail: zhangpf@mail.xidian.edu.cn

Extensive Air Showers (EAS) triggered by high-energy particles can generate short radio pulses, allowing the indirect study of ultra-high energy (UHE) cosmic rays or neutrinos with sparse arrays of antennas. Accurate simulation of EAS and the radio pulses they generate is key in informing the design and detection capability of such arrays. These simulations are usually carried out with either the CoREAS or the ZHAireS simulation packages, which are considered to be the state-of-the-art in the field. In this work, it is pointed out for the first time that the time-discretization processes used in these packages lead to filtering and aliasing effects even though, at first glance, the Nyquist-Shannon sampling criterion seems to be respected. These effects are significant if the receiving antenna is located close to the EAS Cherenkov angle. We give an overview of the causes behind this effect, and suggestions on how to solve it or circumvent it.

10th International Workshop on Acoustic and Radio EeV Neutrino Detection Activities - ARENA2024
11-14 June 2024, Chicago, Illinois, United States

**Corresponding author
*Speaker

1. Introduction

Due to either their extremely low flux or their low cross-section, ultra-high energy (UHE) cosmic rays and neutrinos are almost impossible to detect directly. Instead, we observe the cascades of secondary particles they produce when they interact in the Earth's atmosphere, creating extensive air showers (EAS). Among the various methods used to detect these EAS, radio detection has emerged as a powerful technique due to its high sensitivity and cost-effectiveness [1, 2]

The radio detection of EAS exploits the fact that these showers generate brief electromagnetic pulses as they propagate through the Earth's atmosphere. The primary mechanisms behind these emissions are the *geomagnetic effect* [3], stemming from the deflection of electrons and positrons in the shower by the Earth's magnetic field, and the *Askar'yan effect* [4, 5], producing a coherent radio emission originated in the development of a net negative charge in the shower front. These radio emissions, ranging typically from a few ten of MHz to 1 GHz, can be detected by ground-based or balloon-borne radio antennas and provide valuable insights into the characteristics of the primary cosmic ray, including its energy, incoming direction and mass composition.

As we seek to refine our understanding of UHE cosmic rays and improve the sensitivity of our detection techniques, computer simulations, particularly those employing Monte Carlo methods, have become indispensable. These simulations allow us to model the shower development and the resulting radio emissions under varied atmospheric and geomagnetic conditions, and are crucial not only for interpreting the observed data but also for planning and optimizing the design of radio detection arrays and their components.

In this article, we will focus our attention on a limitation that emerges in simulations performed with CoREAS [6] and ZHAireS [7, 8] when trying to determine the electric field in the time domain. This limitation, usually overlooked, could lead to a misinterpretation of the simulation results.

2. Sampling of simulated radio signals: a filter in disguise

The simulation of radio signals from extensive air showers begins with the generation of a primary cosmic ray particle with predefined properties such as type, energy, and direction. The particle is propagated as it interacts with the atoms in the atmosphere, considering the production of secondaries and their interactions, decays and further particle production. Charged particles, in particular electrons and positrons, will produce electromagnetic fields. The calculation of these emissions and their propagation to a given detector requires solving Maxwell's equations under the constraints imposed by the atmosphere's varying density and the Earth's magnetic field. Two closely related formalisms, the ZHS (Zas, Halzen, and Stanev) algorithm [9] and the end-points algorithm [10] allow to compute the electric field produced by each particle in the cascade, as seen from an observing antenna. The details of these computations are outside of the scope of this article, but in essence each particle trajectory is modeled as a succession of finite impulsive straight-line translations (so-called "tracks"), each contributing with a step function to the vector potential **A**, or two Dirac delta functions to the electric field **E**, as sketched in figure 1. The end-points formalism computes the electric field directly, while the ZHS algorithm computes the vector potential, and once the simulation is finished the electric field is obtained as the negative time derivative of the total vector potential.

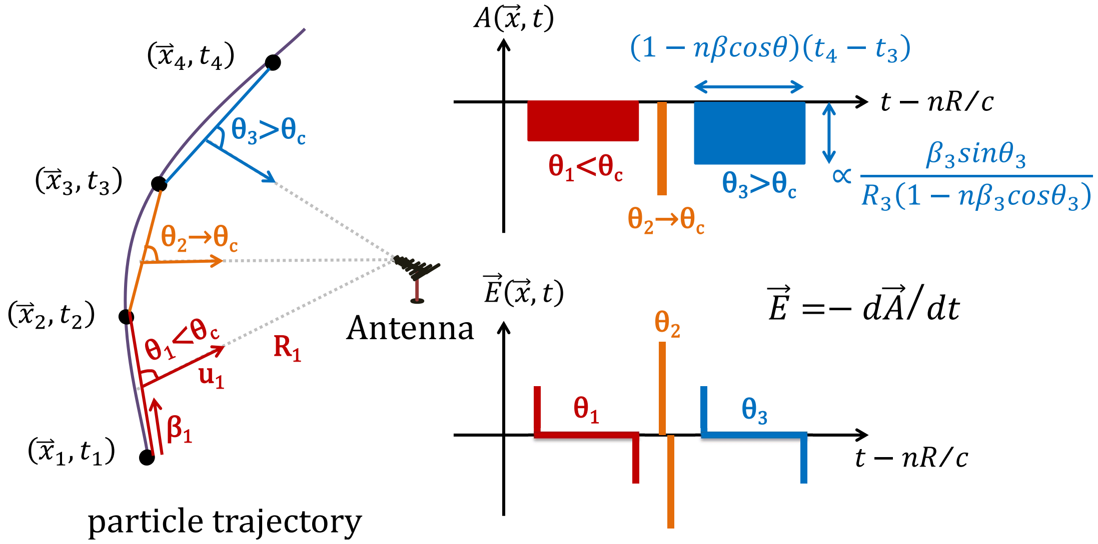

Figure 1: Contribution of a single particle's sub-trajectory to the magnetic vector potential and electric field. Adapted from [11]

An EAS can involve hundred of millions of charged particles, each one potentially contributing to the electromagnetic emissions observed. In practical terms, it becomes computationally prohibitive to store the electric field information for every single track within a shower. However, the desired output from the simulation is the total electric field received at a given antenna. Since real antennas will be connected to an analog-to-digital converter taking snapshots of the electric field value at discrete time intervals, we only require to know the electric field at those precise sampling moments, within a limited bandwidth.

In order to get the field received by the antenna as a function of time, both CoREAS (on the end-points formalism) and ZHAireS (on the ZHS formalism) follow a time-discretized approach based on averaging the electric field contributions (or electromagnetic vector potential contributions in the case of the ZHS formalism) over discrete time bins of width ΔT. The contribution to the field from each track ($\mathbf{A}(t)$ or $\mathbf{E}(t)$) at time t_n is:

$$A(t_n) = \frac{1}{\Delta T}\int_{t_n-\Delta T/2}^{t_n+\Delta T/2} dt\, A(t) \tag{1}$$

This time-discretized approach allows for an efficient computation of the superposition of the electric fields from numerous tracks while the shower is being simulated, on a finite amount of computer memory.

Taking the sum-average of all contributions on discrete time bins is akin to sampling the moving average of the total field, with a sampling rate equal to the inverse of the bin size. The moving average can be represented as a convolution with a normalized boxcar function of width ΔT, whose Fourier transform is a *sinc* function

$$B(f) = \text{sinc}\,(\pi\,\Delta T\, f) \tag{2}$$

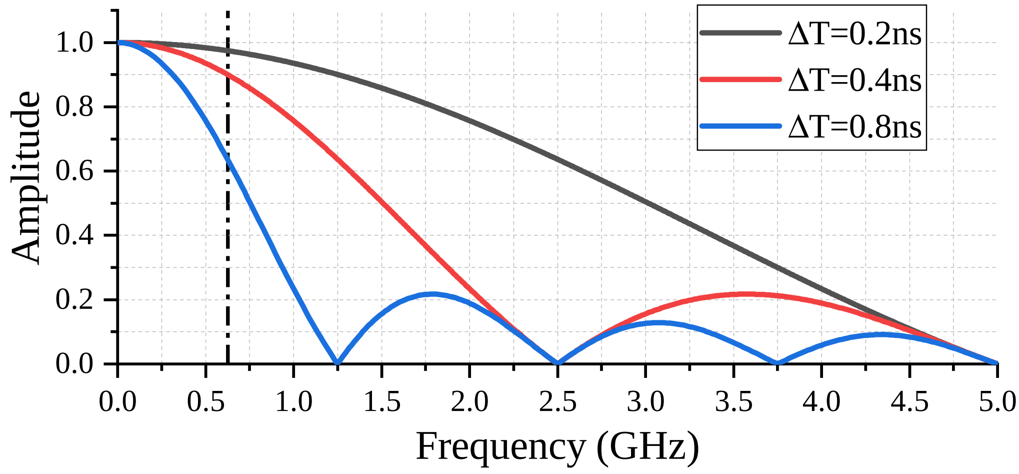

Figure 2: Sinc Function for different time bin width settings, namely 0.8 ns, 0.4 ns and 0.2 ns corresponding to 1, 2 and 4 times the Nyquist sampling rate needed to reconstruct accurately a 625 MHz signal (dashed line).

As a convolution in time domain is equivalent to a product in frequency space, the sampling algorithm described in the previous paragraphs distorts the *real* spectrum of the field. The appearance of a filter due to the moving average is usually overlooked.

As an example, let us suppose that we are interested in working with signals up to a frequency of 625 MHz. When working with sampled signals, the Nyquist-Shannon theorem [12] states that *if a continuous-time signal contains no frequencies higher than B, it can be completely and accurately reconstructed from samples taken at a rate greater than* $2B$ *samples per unit of time*. This is known as the Nyquist rate. In our example, the Nyquist rate would be 1.25 GHz, or a time bin of 0.8 ns. The moving average filter associated with this time bin can bee seen in figure 2.

The *sinc* filter is characterized for having a very slow roll-off and meager stop-band attenuation. The immediate consequence of this is that the moving average filter removes frequency content from the field, and changes the slope of the spectrum. If we naively were to set the sampling rate f_S following the Nyquist criterion, the amplitude at the Nyquist frequency $f_N = 1/2\Delta T$ after sampling becomes $A_{\text{sampled}} = A_{\text{true}} \times \text{sinc}\left(\frac{\pi}{2}\right) \sim 0.64 \times A_{\text{true}}$.

In order to minimize this effect we can reduce the time bin size so that the filter cut-off is several times above the maximum frequency of interest. Using a sampling rate 4 times larger (in our example 5 GHz or 0.2 ns) results in only a 3% loss at half the original sampling rate, at the expense of proportionally higher utilization of CPU time and output disk space. This is how the community has sometimes overcome this issue so far, albeit more for pragmatism (i.e. spectra look better, results are more stable) than for awareness of the underlying phenomena [1].

[1]The discrete time derivative needed in ZHS to get **E** from **A** introduces an additional but unrelated moving average filter that we won't discuss in this article, a more detailed full-length article is in preparation.

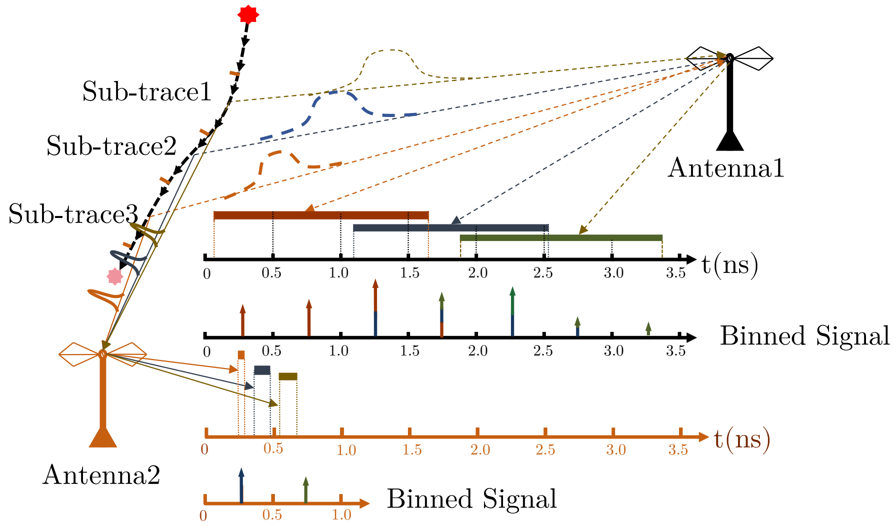

Figure 3: Schematic diagram of time binning. Antennas looking at the shower from the front see the contributions from each track compressed in time, many with a duration smaller than the time bin, while antennas looking from the side see longer contributions, spanning several time bins. The sum of the contributions averaged over the time bin is the sampled signal.

3. ZHS and end-points formalisms are not band limited

Even if the filtering introduced by the bin-averaging can be made negligible by a sufficiently high sampling rate, there is a conceptually deeper issue that we have overlooked so far. Both the end-points and ZHS formalisms build up the electric/vector potential fields from contributions that are infinite in frequency content (Dirac's delta or Heaviside's functions). The superposition of a finite number of these contributions is necessarily also infinite in frequency content. This means that the Nyquist-Shannon criterion can never be formally met, and that the sampled results from the simulations will always be affected by aliasing. Luckily, the sinc filter applied by the moving average works towards reducing the high-frequency content, but it's slow fall-off and wide side-lobes still keep frequency components above the Nyquist limit, that will fold into the spectrum we reconstruct from the samples. Due to the relativistic nature of the particles in the cascade, the amplitude and duration (and thus the frequency content) of each contribution depends on the relative orientation between the particle velocity and the line of sight to the antenna position, as depicted in fig 3. When the relative orientation is close to that of the Cherenkov angle the duration of the contribution tends to 0, significantly increasing its high frequency content. In this region, the effects of aliasing will become more evident.

To illustrate the effects that the moving average and the aliasing can have on simulations, we show in this section the electric fields simulated with ZHAireS for a proton cosmic ray with an energy of 2.154 PeV and incident zenith angle $\theta = 60°$. Similar results have been obtained using CoREAS. Antennas are located at an altitude of 1100m above sea level, in a line perpendicular to the shower axis, with 3m spacing between the antennas. The simulation is located in Dunghuang, China (longitude 94.66°, latitude 40.14°).

In figure 4 (top), we show the electric field maximum amplitude in the E-W polarization, as

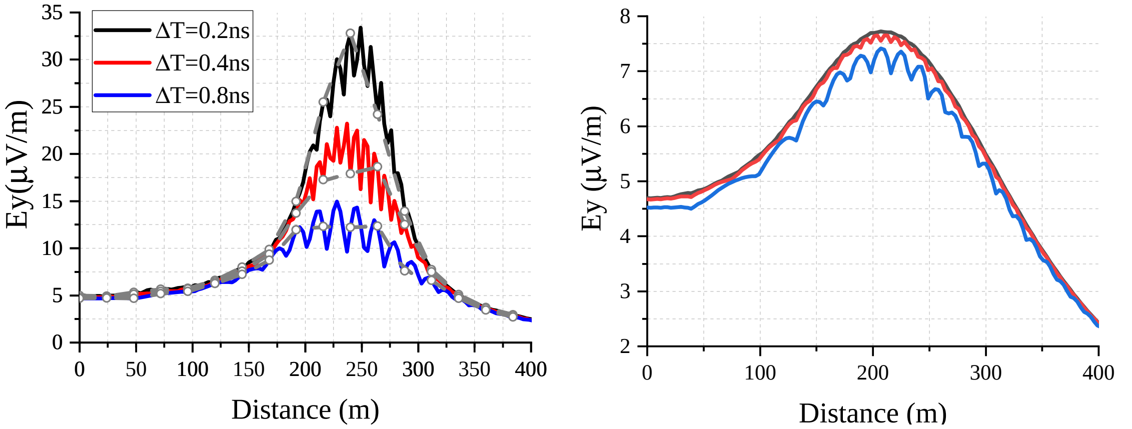

Figure 4: Maximum signal amplitude as a function of distance to the shower core for different time bin values, "full-bandwidth" (left panel) and after filtering in the 0 – 250 MHz frequency range (right panel). Lines in gray show what the result would be if we had simulated only one antenna every 24m.

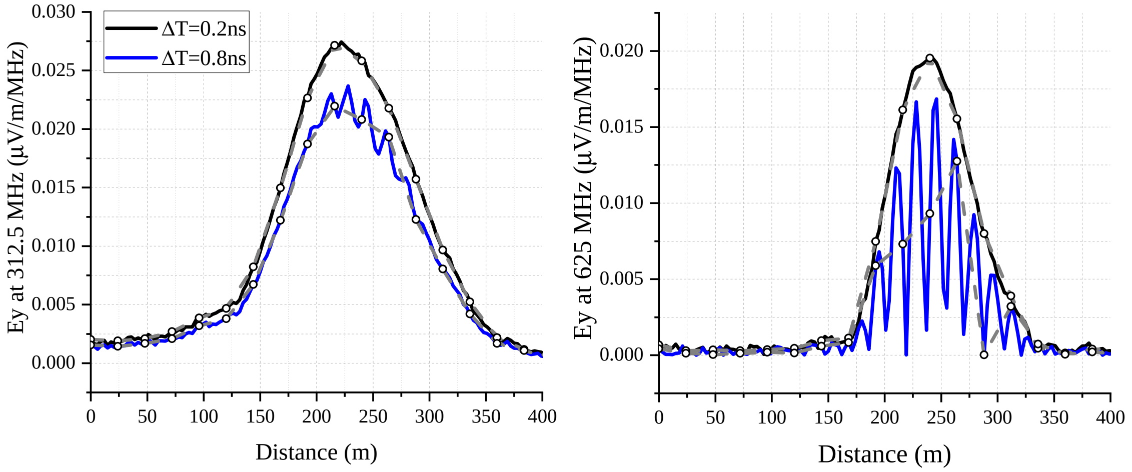

Figure 5: Spectral amplitude as a function of distance to the shower core for different time bin values, evaluated at 312.5 MHz (left panel) and 625 MHz (right panel). Lines in gray show what the result would be if we had simulated only one antenna every 24m.

a function of the distance to the shower core. Each line corresponds to a time bin width of 0.2, 0.4 and 0.8 ns, which would correspond to a moving average filter with its first zero-crossing at 5 GHz, 2.5 GHz and 1.25 GHz respectively (see figure 2). It can be seen that as we get closer to the Cherenkov angle, located in this case close to the 225m mark, the peak value gets more and more diluted the bigger the time bin, as the signals start having higher frequency content that is removed by the narrower moving average filter. The maximum amplitude of the sampled field depends strongly on the phase of the signal with respect to the averaging bins. The averaged amplitude changes significantly if the signal is mostly contained in one time bin or is split across 2 contiguous time bins.

Aliasing effects also depend on the phase of the pulse relative to the sampling bins. Nearby antennas that would receive approximately the same field at slightly different times end up having different spectra, as evidenced by the fluctuations in the Fourier components seen in fig 5. In this case, 0.8 ns (0.2 ns) time bin represents 2 (8) times the Nyquist rate for the 312.5 MHz component and 1 (4) times the Nyquist rate for the 625 MHz component.

Filtering the signal after simulation to remove the high frequency components can mitigate the problem with the signal amplitude (see figure 4 right), but does not remove it completely, as the output of the simulation is already biased before filtering.

It is important to note that these effects would not be apparent (but would still be present) if we had simulated the antennas more evenly spaced[2] which is probably the reason why this issue has not been brought to attention until now. Dotted lines in fig 4 and 5 show what the amplitudes would look like if, instead of using one antenna every 3m we had chosen to set one every 24m.

Further insight is given by figure 6 (animated on modern .pdf viewers), showing the electric field received by an antenna on the Cherenkov cone, sampled at 100 GHz to show its high frequency content. Averaging it over 1 ns time bins (a 1 GHz sampling rate) greatly reduces its amplitude. Depending on the position of the signal relative to the time bins, the maximum value of the field changes significantly as does the slope of the spectrum. The change in the spectrum slope could be specially worrisome for methods using this observable as a handle to reconstruct the energy [13].

4. Possible solutions to be implemented

As we have mentioned in section 2, one way of mitigating these effects is to use a shorter time bin length, resulting in a sampling rate significantly higher than the maximum frequency of interest, and then filtering the simulations to the band of interest. This will significantly reduce aliasing, at the expense of increased CPU time and disk space usage. For simulations with multiple antennas, an adaptive bin size depending on the antenna position relative to the Cherenkov cone, as measured from the shower maximum, would be more efficient than using the same time bin for all antennas. However, this will still suffer from some degree of aliasing very close the Cherenkov cone, and will be prohibitively expensive on very high frequency applications. A more definitive solution, currently being considered for its implementation in ZHAireS, would be to filter each individual contribution to the field *before sampling it*, in order to guarantee that the Nyquist sampling criteria is always respected, eliminating all possible aliasing problems.

5. Acknowledgments

We would like to thank the technical support and fruitful discussions with the GRAND and ZHAireS working group, and the computer resources from the National Key Laboratory of Radar Detection and Sensing Xian, 710071, China employed to perform the simulations leading to this study.

[2]or if the antenna time windows had more or less the same starting phase for all antennas, as is done in ZHAireS when the expected position of shower maximum is given in the input

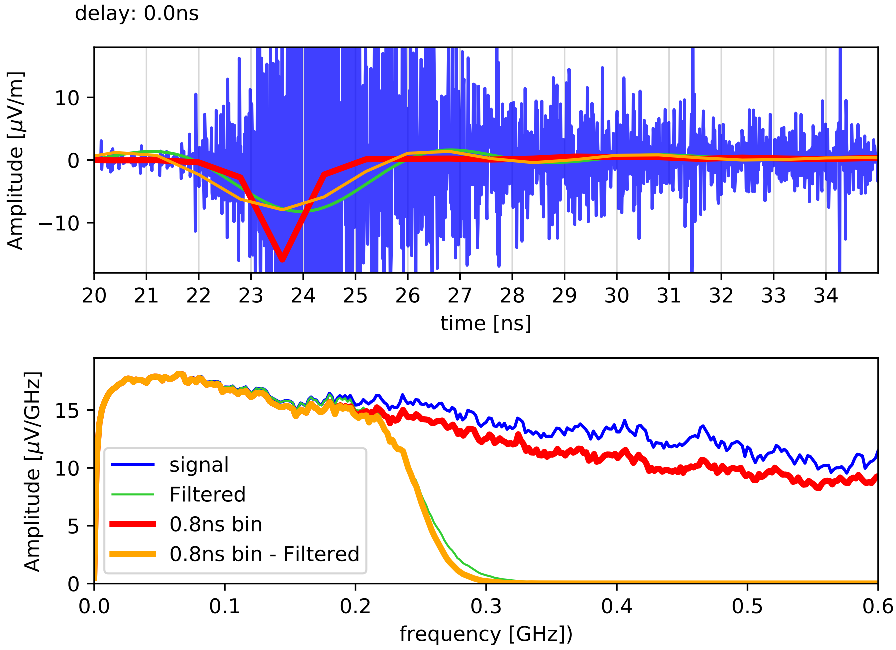

Figure 6: Electric field signal computed at 0.01ns bin size (blue) and at 1 ns bin size (red), showing the effect of averaging (animated on modern .pdf readers). When a delay is added to the signal (akin to moving the antenna position), the amplitude of the averaged signal and the slope of its spectrum change. Filtering the simulation results (in this case with a 250 MHz low-pass filter) diminishes the differences.

References

[1] T. Huege, Phys. Rep. 620, 1 (2016).

[2] F. G. Schröder, Prog. Part. Nucl. Phys. 93, 1 (2017).

[3] O. Scholten, K. Werner, and F. Rusydi, Astropart. Physics 29, 94 (2008)

[4] G.A. Askar'yan, Sov.Phys.JETP 14 (1962) 441

[5] Pierre Auger Collaboration, Phys. Rev. D 89 (2014) 052002

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

[7] J. Alvarez-Muñiz, W.R. Carvalho Jr., and E. Zas, Astropart. Phys. 35 (2012) 325

[8] J. Alvarez-Muñiz et al., Phys.Rev.D 86 (2012) 123007,[arXiv:1208.0951]

[9] E. Zas, F. Halzen, T. Stanev, Phys. Rev. D 45 (1992) 362

[10] C. James, H. Falcke, T. Huege, and M. Ludwig, Phys. Rev. E 84 (2011) 056602

[11] J. Alvarez-Muñiz; E. Zas.,Universe 2022, 8, 297

[12] C.E. Shannon, Proc. Institute of Radio Engineers (1949), vol. 37, no.1, pp. 1021.

[13] C. Welling et al JCAP10(2019)075