PROCEEDINGS OF SCIENCE

Simulations of radio detection of cosmic rays with SKA-Low

A. Corstanje,[a,b,*] **S. Buitink,**[a,b] **M. Desmet,**[a] **H. Falcke,**[b,d,e] **B. M. Hare,**[d] **J. R. Hörandel,**[a,b,e] **T. Huege,**[a,f] **V. B. Jhansi,**[c] **N. Karastathis,**[f] **G. K. Krampah,**[a] **P. Mitra,**[a] **K. Mulrey,**[b,e] **A. Nelles,**[g,h] **H. Pandya,**[a] **O. Scholten,**[i,j] **K. Terveer,**[h] **S. Thoudam,**[c] **G. Trinh**[k] **and S. ter Veen**[d]

[a] *Vrije Universiteit Brussel, Astrophysical Institute, Pleinlaan 2, 1050 Brussels, Belgium*
[b] *Department of Astrophysics/IMAPP, Radboud University Nijmegen, P.O. Box 9010, 6500 GL Nijmegen, The Netherlands*
[c] *Department of Physics, Khalifa University, P.O. Box 127788, Abu Dhabi, United Arab Emirates*
[d] *Netherlands Institute for Radio Astronomy (ASTRON), Postbus 2, 7990 AA Dwingeloo, The Netherlands*
[e] *Nikhef, Science Park Amsterdam, 1098 XG Amsterdam, The Netherlands*
[f] *Institut für Astroteilchenphysik, Karlsruhe Institute of Technology (KIT), P.O. Box 3640, 76021 Karlsruhe, Germany*
[g] *Deutsches Elektronen-Synchrotron DESY, Platanenallee 6, 15738 Zeuthen, Germany*
[h] *ECAP, Friedrich-Alexander-Universität Erlangen-Nürnberg, 91058 Erlangen, Germany*
[i] *Interuniversity Institute for High-Energy, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium*
[j] *University of Groningen, Kapteyn Astronomical Institute, Groningen, 9747 AD, Netherlands*
[k] *Department of Physics, School of Education, Can Tho University Campus II, 3/2 Street, Ninh Kieu District, Can Tho City, Vietnam*

E-mail: a.corstanje@astro.ru.nl

We present a simulation study on the cosmic-ray detection capabilities of the Square Kilometre Array, the low-frequency part of which is being built in Australia. With nearly 60,000 antennas in a 1 km diameter, its antenna density is two orders of magnitude higher than at LOFAR. The wider frequency band of 50 to 350 MHz allows to resolve the radio energy footprint into smaller detail, as well as providing more information via the signal frequency spectra. We discuss the improved resolution in depth of shower maximum X_{max} compared to LOFAR. Moreover, the next-level features open the possibility of measuring the longitudinal air shower profile in more detail than just its maximum X_{max}, for individual air showers. The benefits are twofold: it gives additional information on the mass composition (independent of X_{max}), and the main hadronic interaction models predict observable differences in longitudinal profiles and their distributions. Thus, it would help in constraining hadronic physics at energy levels beyond laboratory experiments.

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

*Speaker

© Copyright owned by the author(s) under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (CC BY-NC-ND 4.0).

PoS(ICRC2023)500

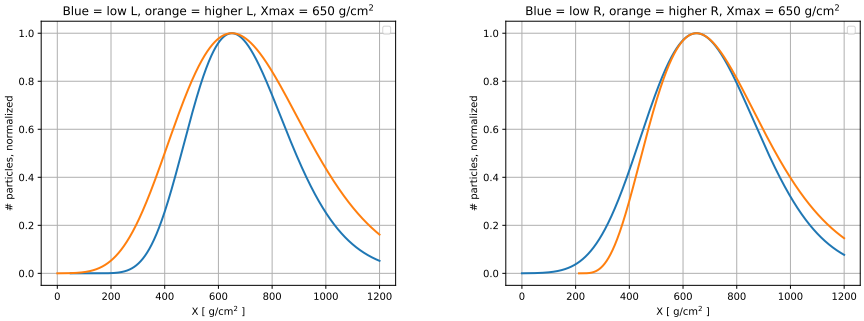

Figure 1: A longitudinal particle distribution example at $X_{\max} = 650\,\mathrm{g/cm^2}$, demonstrating the effect of the parameters L (left) and R (right). Variations have been exaggerated for demonstration purposes.

1. Introduction

The Square Kilometre Array, of which the low-frequency half is being built in Australia [1], will improve on current distributed radio telescopes such as LOFAR [2] by two orders of magnitude in terms of numbers of antennas. With nearly 60, 000 antennas in a diameter of 1 km and a frequency range up to 350 MHz, it allows for air shower reconstructions with a precision not seen before. Here we focus on the longitudinal distribution of particles along the air shower development, of which the maximum is defined as $X_{\max}$. A mass composition analysis based on $X_{\max}$ reconstructed from the radio signals has been done at LOFAR [3, 4] and at AERA [5]. The longitudinal evolution can be (in almost all cases) reliably described by two more parameters, akin to the second and third moment of the distribution. These parameters depend on both the mass composition and on the specific hadronic interaction model [6].

We describe a preliminary analysis to measure the longitudinal distribution, aiming to improve mass composition analysis and to provide an (indirect) test of the hadronic interaction models.

2. Methods

We parametrize the longitudinal distribution of particles using three parameters $X_{\max}$, L, and R [7]:

$$N(X) = \exp\left(-\frac{X - X_{\max}}{RL}\right)\left(1 - \frac{R}{L}\left(X - X_{\max}\right)\right)^{\frac{1}{R^2}}, \tag{1}$$

where N is the number of particles, X is the depth in the atmosphere (going down) in $\mathrm{g/cm^2}$, L is related to the variance, and R to the asymmetry (skewness) of the distribution. In Fig. 1 we show how L and R affect the shape of the distribution.

We have set up a Monte Carlo ensemble of 110 proton showers at 10^{17} eV primary energy, using Corsika and CoREAS 7.7410 [8, 9]. They were chosen to all have the same $X_{\max}$, that is, $645 \pm 0.5\,\mathrm{g/cm^2}$, for which 50, 000 CONEX [10] and 500 Corsika showers were used as pre-selection pool. This is useful to take a closer look at the effects of L and R on the footprints, unaffected by $X_{\max}$-variations.

We have simulated 208 antennas on a radial or 'star-shaped' grid, and interpolate the energy fluence and amplitude along the footprint using an interpolation scheme based on Fourier series we

PoS(ICRC2023)500

have developed [11, 12]. This allows for highly accurate reconstructions of the full pulse signals at arbitrary positions. Errors on amplitude and energy fluence are well below 1 % for reasonably strong signals, and the pulse shapes match to 99.8 % on average, using normalized cross-correlations to the 'true' simulation as a metric. Thus, simulating 208 antennas suffices for high-precision analysis of SKA-measured air showers; as simulating 60,000 antennas directly would be intractable, this overcomes an important practical obstacle.

In the next section we show a direct comparison of footprint shapes as a function of L and R; after this, we show a reconstruction of X_{max}, L, and R based on simulated measurements with realistic sky noise levels.

2.1 Comparing equal-X_{max} footprints as a function of L and R

A relatively straightforward way to compare footprints is to do a least-squares fit of all shower footprints in the ensemble to that of a chosen reference shower. This is similar to the chi-squared fit we do to real data, except that here we have no natural noise. So, we calculate the mean squared error or MSE which replaces χ^2:

$$\mathrm{MSE} = \sum_{\mathrm{ant}} \left(f_i(\vec{x}) - A f_j(\vec{x}) \right)^2 , \tag{2}$$

where A is energy scale factor A that is a free parameter in the fit, and f_i is the pulse energy fluence in shower i.

We have also looked at the radial dependence of the circular-symmetric and $\cos(\phi)$ modes of the energy fluence when taken around a circle in the footprint. Again we look at the effects from variations in L and R by taking one shower as reference and plotting the differences in the other showers.

2.2 Monte Carlo setup for simulated air shower reconstruction at SKA

Taking a step towards a realistic detector and analysis simulation, we use the ensemble of simulations described above. Applying an antenna model, in this case SKALA2 [13], converts E-fields as they arrive at the antenna to measurable voltages. We add sky noise at approximately realistic levels as we have also used to calibrate LOFAR antennas [14], using spectra from LFMap [15]. Reconstruction of X_{max} is done as we have at LOFAR [3, 16] where we have used the same code where possible.

Adaptations are made to the pulse detection and measurement in a wider frequency window, taking a window width of 12 ns between 50 and 350 MHz, and 60 ns between 50 and 100 MHz. The low-frequency band is used for reconstructing L and R or a combination of them described in the next section. From 300 realisations of the noise, we get a trigger probability (taking a 5 sigma criterion for now) and an uncertainty on the measured energy fluence. Using the interpolation method in [11, 12] applied to the fluence and the uncertainty, we obtain a measured value and an uncertainty at all antennas in the SKA layout. In a future analysis we will use the full signal interpolation to have a more direct simulation of the measured quantities.

PoS(ICRC2023)500

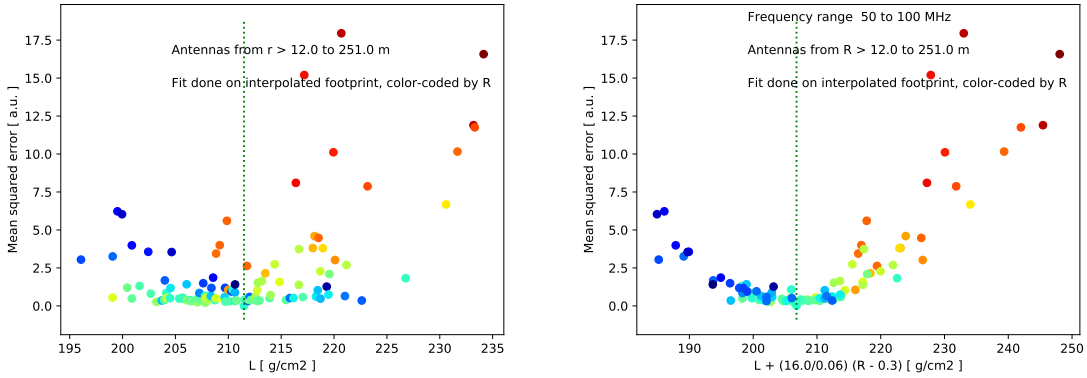

Figure 2: Left: mean squared error on fitting the footprints from the shower ensemble to a test shower at $L = 212\,\mathrm{g/cm^2}$, color-coded by R. Right: the same, but plotted against a linear combination of L and R defined in Eq. 3.

3. Footprint comparison results

We have plotted the mean-squared error in fitting shower footprints to a reference shower in the left panel of Fig. 2, as a function of L and color-coded by R. This is for a shower with L close to its median value, and the frequency range is 50 to 100 MHz as variations are more easily seen here.

This plot looks irregular, yet considering the color coding, the variations with R seem non-random. Therefore, in the right panel of Fig. 2 we plot the MSE versus a combination of L and R. A linear function of both is the simplest assumption, which we denote S (for sensitive parameter) as

$$S(L,R) = L + \frac{16\,\mathrm{g/cm^2}}{0.06}\,(R - 0.3)\,, \tag{3}$$

where the constant factor is used to have the same dimension of $\mathrm{g/cm^2}$ as R is dimensionless. The numbers correspond to about the median value of R (0.3) and half its range (0.06). The constant factor in general depends on the frequency range, which may later be used to disentangle L and R again. Its optimal value will likely depend on X_{max} and zenith angle as well.

We see now a clear, and to lowest order parabolic minimum around the 'true' position. So we conclude that energy fluence footprints are sensitive to S rather than to L and R separately.

3.1 The circular-symmetric and $\cos(\phi)$ Fourier modes as function of distance to shower core

The two most important Fourier modes are shown in Fig. 3. The zero level is defined by the shower at median L, and the differential curves of the other showers are shown, with a color coding by S as defined in Eq. 3. Especially the inner part of the footprint, i.e. inside the Cherenkov ring, is sensitive to S, with lower values showing more energy there. Also the angular variation is stronger for lower S-values.

From this we see that these variations must be measurable, at least with an ultra-dense antenna array. The only caveat is that they might be degenerate with variations in X_{max}, which we discuss further below from Fig. 6. However as this is only a first step towards a full exploitation of the information in the signals, in a limited frequency range, we expect the parameters to be separable.

PoS(ICRC2023)500

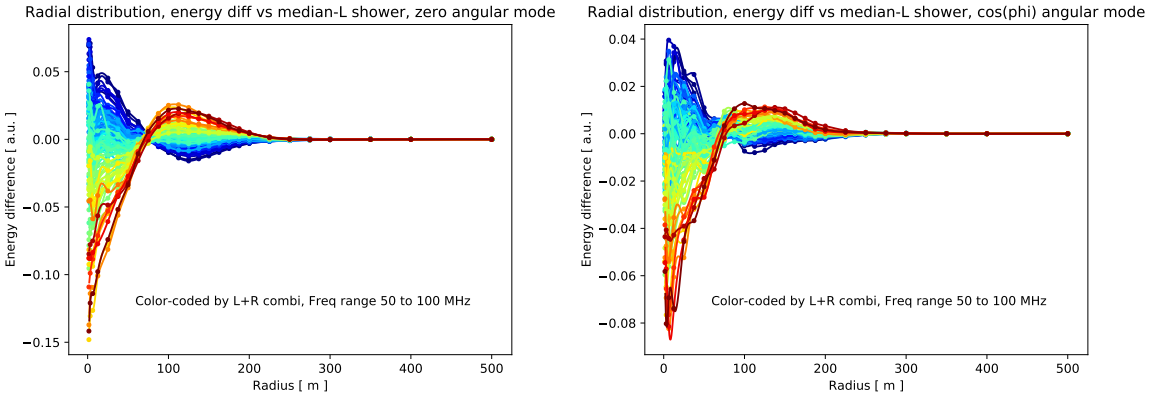

Figure 3: Left: Circular-symmetric mode of the energy fluence as a function of distance r to the shower core, color-coded by the parameter S defined in Eq. 3. Right: the same for the mode proportional to $\cos(\phi)$.

4. Result of simulated reconstructions

We have simulated the reconstruction process of X_{max} and of the parameter S which is a linear combination of the two longitudinal distribution parameters L and R. The reconstruction of X_{max} is done on showers in an X_{max}-range of 50 g/cm^2, whereas for the reconstruction of S we have used the constant-X_{max} ensemble ($\pm$0.5 g/cm^2).

4.1 Reconstruction of X_{max} in the 50 to 350 MHz frequency band

In Fig. 4 we show a reconstruction of X_{max}; no information on L and R has been used here. Typical precision numbers for X_{max} are 6 to 8 g/cm^2. This means that to a large degree, statistical errors have been removed using the very large number of antennas. Accuracy of X_{max} would then be mainly limited by systematic uncertainties, as from the atmosphere, the antenna (models), and from the simulations themselves.

4.2 Reconstruction of a sensitive parameter S in the 50 to 100 MHz band

For the reconstruction of S from Eq. 3 we use the ensemble at fixed X_{max}. An example is shown in Fig. 5. Although the differences in reduced χ^2 are quite small, the large number of antennas makes them significant, and the fit value is close to the simulation 'truth' at 204.6 g/cm^2.

A measurement with real data requires determining X_{max} and S (or L and R) together. For a preliminary exploration, we have simulated some additional showers for more X_{max} coverage. We have taken three showers using the 50 to 100 MHz band, added noise as described above, and fitted all simulated showers to them. Using the reduced χ^2 as color code, we have plotted the results in an S versus X_{max} plot, in Fig. 6. The range of χ^2 has been restricted to show only the part close to the optimum.

The fit quality landscape has a 'valley' along a line with a roughly constant slope, which means that to some extent X_{max} and S are interchangeable. Yet already in this first analysis, we can already distinguish low, intermediate, or high values of S, for example by using an X_{max} measurement as above, on the entire frequency range.

PoS(ICRC2023)500

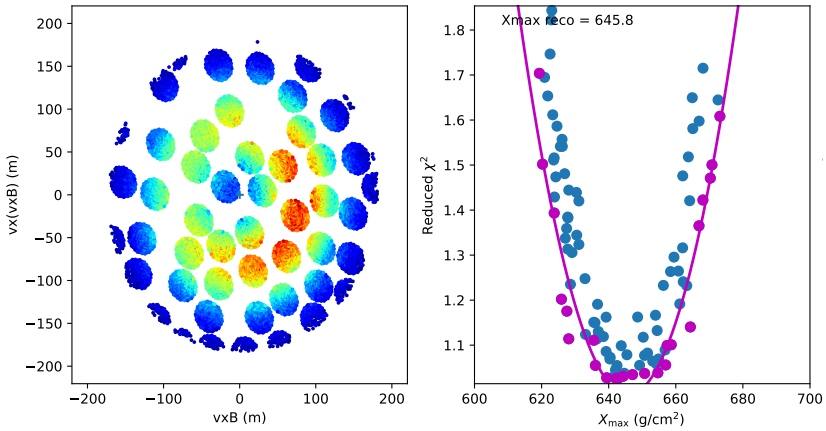

Figure 4: Left: pulse energy fluence in antenna positions projected onto the shower plane, from simulated measurements with noise. Right: reduced χ^2 obtained from fitting the showers in the ensemble to the measured shower. A lower envelope of data points (shown in magenta) has been selected to perform the parabola fit on.

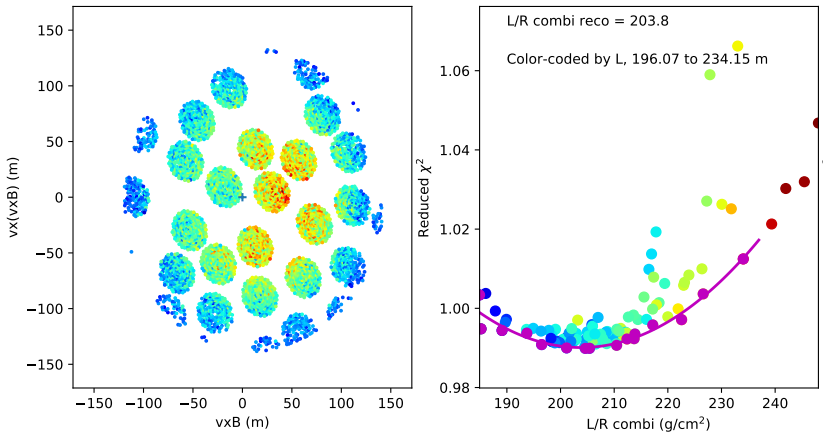

Figure 5: Same as in Fig. 4, now analyzed in a 50 to 100 MHz frequency window; color-coding in the right panel corresponds to the value of L

We expect a more extensive analysis to improve considerably on this result. For instance, by making use of the entire frequency band, and of the frequency spectra and polarizations of the pulses which we have not yet considered.

5. Summary

The presented analysis demonstrates that SKA will be able to deliver measurements of air showers to a level of detail that is beyond current radio observatories. Parametrizing the longitudinal distribution of air showers by $X_{\max}$, L, and R, observable sensitivity is found to $X_{\max}$ as well as

PoS(ICRC2023)500

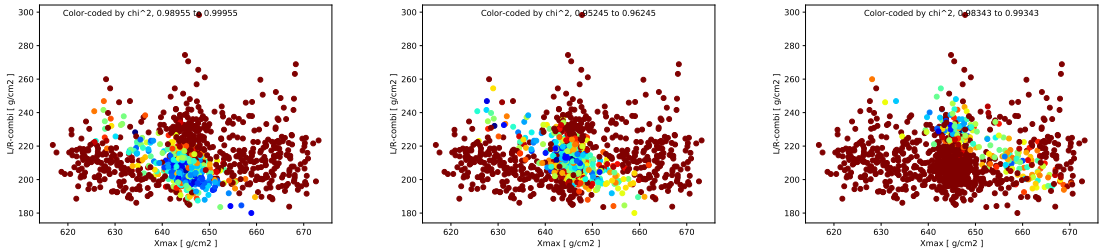

Figure 6: Fit quality (reduced χ^2, color-coded) as a function of X_{max} and the parameter S from Eq. 3, for a low, medium, and high value of S, respectively. Along a slanted line, there is a region of near-optimal fits.

to a parameter S which is a linear combination of L and R. This particular function also depends on frequency range, and probably on zenith angle and X_{max}. We have conducted a simulated measurement by making use of a sky noise model, and applying the SKALA2 antenna model to pulses and noise.

We find an X_{max} precision (statistical uncertainty only) of about 8 g/cm^2 using the full frequency range, from a reconstruction method copied from LOFAR. Further improvements to the analysis might even lower this number further. Of course, the complete X_{max}-uncertainty would then be dominated by systematic effects, from uncertainties in the atmospheric profiles, the antenna models, and the simulations themselves.

We also find a reliable reconstruction of S in the range 50 to 100 MHz in an ensemble at constant X_{max}. Using a large ensemble and combining these results, we can already separate low, medium, and high values of S.

As a first exploration, the presented results are only a starting point. Making more effective use of the information in the signals, i.e. by using the full bandwidth, the frequency spectra and/or the polarization of the pulses, we can expect the reconstruction of S, or of L and R separately, to still improve considerably. Analysis methods have to be extended beyond what was done at LOFAR, to make use of the capabilities of a next-level observatory like SKA. Currently we have to simulate large numbers of showers per measured event, owing to the three dimensions X_{max}, L, and R to cover, to make accurate reconstructions. Efforts to limit this practical obstacle are under way, for instance by using the template synthesis method presented in [17]. We have already found that simulating all 60,000 antennas at SKA will not be necessary as the full signal traces can be accurately interpolated from around 200 simulated antennas [12].

The longitudinal parameters beyond X_{max} are at present not readily measured for individual air showers at any cosmic-ray observatory. Therefore, being able to do this at SKA would be a significant advancement of the radio detection method for detailed reconstructions of air showers.

PoS(ICRC2023)500

Acknowledgements

BMH is supported by ERC Grant agreement No. 101041097; AN and KT acknowledge the Verbundforschung of the German Ministry for Education and Research (BMBF). NK acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Projektnummer 445154105. MD is supported by the Flemish Foundation for Scientific Research (grant number G0D2621N). ST acknowledges funding from the Abu Dhabi Award for Research Excellence (AARE19-224).

References

[1] G. Tan *et al.*, "The square kilometre array baseline design v2.0," in *2015 1st URSI Atlantic Radio Science Conference (URSI AT-RASC)*, pp. 1–1. 2015.

[2] M. P. van Haarlem *et al. Astronomy and Astrophysics* **556** (2013) A2.

[3] A. Corstanje *et al. Phys. Rev. D* **103** (May, 2021) 102006.

[4] S. Buitink *et al. Nature* (Mar, 2016) 70 .

[5] B. Pont for the Pierre Auger Observatory *PoS* **ICRC2021** (2021) 387.

[6] S. Buitink for the SKA High Energy Particle Science Working Group *PoS(ICRC2023)503* .

[7] S. Andringa, R. Conceição, and M. Pimenta *Astroparticle Physics* **34** no. 6, (2011) 360–367.

[8] D. Heck *et al.*, *CORSIKA: a Monte Carlo code to simulate extensive air showers.* Feb., 1998.

[9] T. Huege, M. Ludwig, and C. W. James *ARENA 2012, AIP Conf. Proc. 1535* (2013) 128–132.

[10] T. Bergmann *et al. Astroparticle Physics* **26** no. 6, (2007) 420 – 432.

[11] A. Corstanje *et al. Submitted to JINST* (2023) .

[12] A. Corstanje *et al. PoS(ICRC2023)501* (2023) .

[13] E. de Lera Acedo *et al. Experimental Astronomy* **39** (2015) 567 – 594.

[14] K. Mulrey *et al. Astroparticle Physics* **111** (2019) 1–11.

[15] E. Polisensky *http://www.faculty.ece.vt.edu/swe/lwa/memo/lwa0111.pdf* (2007) .

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

[17] M. Desmet *et al. PoS(ICRC2023)216* .

PoS(ICRC2023)500