

## Cosmic Ray Mass Measurements with LOFAR

Stijn Buitink<sup>1,\*</sup>, Antonio Bonardi<sup>2</sup>, Arthur Corstanje<sup>2</sup>, J. Emilio Enriquez<sup>2</sup>, Heino Falcke<sup>2,3,4</sup>, Jörg R. Hörandel<sup>2,4</sup>, Pragati Mitra<sup>1</sup>, Katie Mulrey<sup>1</sup>, Anna Nelles<sup>5</sup>, Jörg Paul Rachen<sup>2</sup>, Laura Rossetto<sup>2</sup>, Pim Schellart<sup>2</sup>, Olaf Scholten<sup>6,7</sup>, Satyendra Thoudam<sup>2</sup>, Gia Trinh<sup>6,7</sup>, Sander ter Veen<sup>3</sup>, and Tobias Winchen<sup>1</sup>

<sup>1</sup> Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels (Belgium)

<sup>2</sup> Department of Astrophysics/IMAPP, Radboud University, P.O. Box 9010, 6500 GL Nijmegen (the Netherlands)

<sup>3</sup> Astron, Oude Hoogeveensedijk 4, 7991 PD Dwingeloo (The Netherlands)

<sup>4</sup> Nikhef, Science Park 105, 1098 XG Amsterdam (The Netherlands)

<sup>5</sup> University of California Irvine, Irvine, CA 92697, (USA)

<sup>6</sup> KVI-CART, Zernikelaan 25, 9747 AA Groningen (The Netherlands)

<sup>7</sup> University of Groningen, Zernikelaan 25, 9747 AA Groningen (The Netherlands)

**Abstract.** In the dense core of LOFAR individual air showers are detected by hundreds of dipole antennas simultaneously. We reconstruct  $X_{\max}$  by using a hybrid technique that combines a two-dimensional fit of the radio profile to CoREAS simulations and a one-dimensional fit of the particle density distribution. For high-quality detections, the statistical uncertainty on  $X_{\max}$  is smaller than 20 g/cm<sup>2</sup>. We present results of cosmic-ray mass analysis in the energy regime of  $10^{17} - 10^{17.5}$  eV. This range is of particular interest as it may harbor the transition from a Galactic to an extragalactic origin of cosmic rays.

### 1 Introduction

Radio detection of extended air showers is a powerful and rapidly developing technique [1] that is suitable for measuring the atmospheric depth of the shower maximum  $X_{\max}$  [2, 3]. Arrays with a high density of radio antennas, like LOFAR [4] and the future SKA [5], are capable of accurately measuring the mass composition of cosmic rays below the ankle, where a transition of a Galactic to an extragalactic component is expected to occur [6, 7].

LOFAR is a digital radio telescope constructed in the North of the Netherlands with satellite stations across Europe. It consists of thousands of dipole antennas, sensitive in the frequency range of 10–240 MHz. It produces extremely detailed air shower radio data, by using the dense core region, or *superterp*, where 384 antennas are located within a circle of 320 m diameter. A particle array, LORA [8], has been installed in the core and is used for triggering and reconstruction. Each antenna contains a ring buffer which is read-out in case of a trigger. The full waveform is stored for offline analysis [9].

The vast amount of data per shower and the complicated radiation pattern require new reconstruction techniques. By fitting two-dimensional radiation profiles to the data, LOFAR has achieved a

---

\*e-mail: Stijn.Buitink@vub.ac.be

resolution on  $X_{\max}$  below  $20 \text{ g/cm}^2$  [3] which is comparable to the resolution of fluorescence detection. An analysis of the first set of high-quality LOFAR events has yielded mean  $X_{\max}$  values that are in agreement with existing data based on other techniques and is indicative of a strong contribution of light nuclei in the energy range of  $10^{17} - 10^{17.5} \text{ eV}$  [10].

## 2 Simulations

At LOFAR, showers are reconstructed using dedicated sets of CORSIKA [11] simulations, i.e. for each measured shower we generate enough proton showers and iron showers to span the range of possible values for  $X_{\max}$ . Such simulations are computationally expensive, so we aim to minimise the number of simulated showers per events, while still ensuring that (a) the complete range of naturally occurring  $X_{\max}$  values is covered, and (b) the reconstruction is accurate enough.

We achieve this by using CONEX to quickly generate hundreds of showers [12]. For these showers only the first few interactions are simulated in CORSIKA, after which the rest of the shower development is based on parametrizations. Since the value of  $X_{\max}$  critically depends on these first interactions the set of CONEX showers contains a realistic distribution of  $X_{\max}$  values. Moreover, it is possible to select a subset of these showers for subsequent full CORSIKA simulation. We choose this subset by first selecting the two proton showers with the highest and lowest value for  $X_{\max}$ . Then, we select nine more proton showers with intermediate values, (nearly) equally spaced between the extreme values. Since the range of  $X_{\max}$  values for iron showers is much narrower, we select a total of five iron showers using the same algorithm.

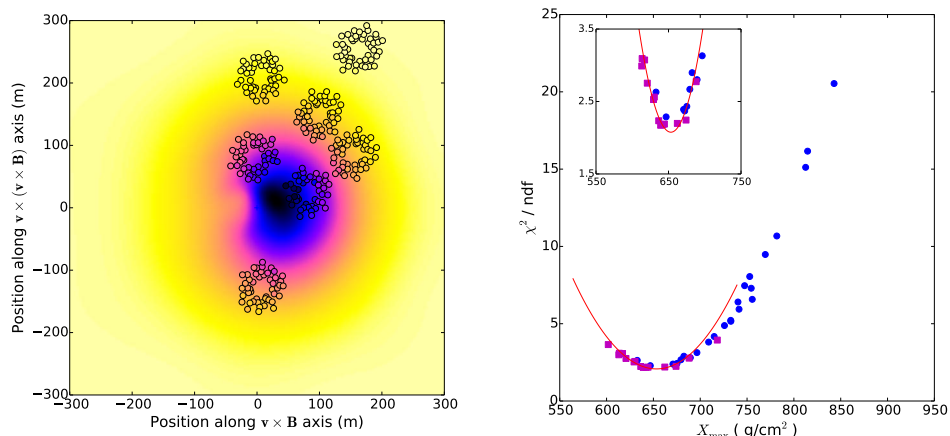
Finally we add additional shower simulations close to the estimated  $X_{\max}$  of the measured shower. This estimation is given by fitting a parametrization of the radio signal [13] to the data, that has an uncertainty of  $\sim 40 \text{ g/cm}^2$ . We select 11 additional showers around this expected value.

The subset acquired in this way is then simulated in CORSIKA including the CoREAS plugin for calculation of the radio emission [14]. The radio emission is simulated for antennas that form a star-shaped pattern in the shower plane. An antenna model is applied to the simulated electric fields to calculate the measured signals in the two dipoles of the LOFAR low band antennas. The measured pulse power is defined as the total power in a window of 55 ns centered around the pulse maximum, summed over the two polarisations. Finally, a two-dimensional map of the pulse power is created by interpolation of the star-shaped pattern. A detailed description of this procedure can be found in [3].

A GEANT4 simulation of the LORA scintillators is used to calculate the lateral distribution of the deposited energy. So, for each measured shower a dedicated set of simulations is produced, and from each simulation in a set a two-dimensional radio profile and a one-dimensional particle profile are produced.

## 3 Reconstruction

The radio and particle profiles are fitted to the data simultaneously. The fit has four free parameters: two for the shower core position and scaling parameters for both the radio and the particle profiles. Inclusion of the particle data in the fit does not have much effect on the fitted core position in practice. However, it serves two important purposes: as an additional consistency check and to determine the energy scale. The analysis presented here does not yet include the absolute calibration of the antennas [15] so the scaling parameter for the radio power has an arbitrary value. In effect, the ratio of two scaling parameters also has an arbitrary value. It should however be the same constant for



**Figure 1.** Left: event reconstruction. The small circles indicate LOFAR antenna positions (grouped in larger circles) projected into the shower plane. Their color reflects the measured power of the radio pulse. The background is a two-dimensional radio map from interpolated CoREAS simulations. The fit quality is high when the colors inside the circles blend into the background. Right: reconstruction of  $X_{\max}$ . Every data point corresponds to a single simulated shower. Together they form the dedicated simulation set produced for this particular shower. The reduced  $\chi^2$  of the fit is plotted against  $X_{\max}$  and a parabola is fitted to find the reconstructed  $X_{\max}$ . The inlay is a zoomed in version of the minimum of the curve demonstrating the sharpness of the minimum. Purple squares correspond to iron shower and blue circles to proton shower.

all measured showers. Therefore, the width of the distribution of this parameter is a measure for the energy resolution of the reconstruction<sup>1</sup>.

The reduced  $\chi^2$  of the fit is very sensitive to  $X_{\max}$  as is seen in Fig. 1. The minimum of the  $\chi^2$ -curve is found by fitting a parabola to the points near the lowest value. The curve is not smooth because shower-to-shower fluctuations are not just variations in  $X_{\max}$ . Other parameters that will vary include the width of the longitudinal profile, the shape of the tail of the shower, and the lateral distribution of the shower. These variations cause the jitter in the data points. If an infinite amount of simulations could be produced, the data points are expected to form a band rather than a line.

The reconstructed value for  $X_{\max}$  found in this method is the atmospheric depth corresponding to the atmosphere used in the simulation. For this analysis we used the standard US atmosphere which is also a good average for northern Europe. For each measurement we construct a more accurate atmosphere using the GDAS database [16]. A first order correction is made by assuming that the reconstructed  $X_{\max}$  corresponds to the correct altitude  $h$  above sea level. The corrected value for  $X_{\max}$  is then found by evaluating the atmospheric depth of the GDAS profile at altitude  $h$ . However, it was realised that the change in refractive index (which also depends on the humidity profile) also affects the reconstructed  $X_{\max}$ . This effect is included as a systematic effect in the present analysis, although it can be corrected for in the future [17].

<sup>1</sup>In fact this is a conservative choice since it includes the energy resolution of both the radio and particle fit, instead of relying on the most accurate of the two.

## 4 Event selection

For this analysis we have selected showers that were detected by at least four LOFAR antenna groups (each containing 48 active antennas). At trigger level, we require that 16 out of the 20 scintillator detectors find a signal. Both conditions introduce a possible bias towards light or heavy primaries in the sample. Heavy nuclei interact higher in the atmosphere, so the particle density at the ground becomes smaller, especially for more inclined showers. This leads to reduced trigger probability. On the other hand, protons and light nuclei penetrate deeper into the atmosphere and produce showers closer to the ground. Therefore, the size of the radio footprint is reduced and the probability of having a detectable signal in at least four antenna groups is reduced.

To ensure that our final event sample is bias-free, we require that *all* of simulated showers in the dedicated simulation set of a particular event (a) produce a trigger in the LORA array and (b) have a detectable radio signal in at least four antenna groups. This requirement removes only a few events above  $10^{17}$  eV, but below that energy the amount of removed events rapidly increases, mainly because of the trigger efficiency. Composition measurements below  $10^{17}$  eV become possible by changing the triggering strategy in the future, for example by triggering on radio-only [18] or a combination of radio and particles.

A quality cut is imposed in order to only include showers with relatively small uncertainty in the reconstruction on the core, the energy, and  $X_{\max}$ . In fact, those three uncertainties are strongly correlated and a single cut on the precision of the core suffices. The uncertainties are calculated using the dedicated set of simulations. They therefore depend on the arrival direction, core position and energy of the measured shower, but not on its value for  $X_{\max}$ . Thus, no additional bias on composition is introduced by this cut.

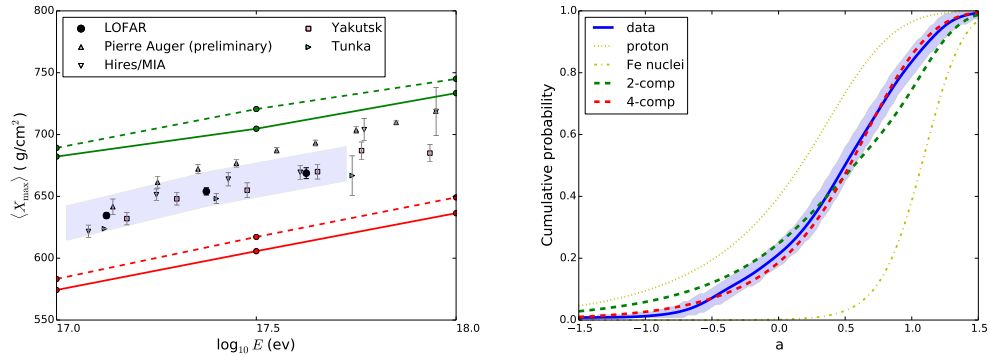
The final event sample counts 118 showers recorded before January 2015. The mean energy resolution of the sample is 32% and the resolution on  $X_{\max}$  is 17 g/cm<sup>2</sup>. The reduced  $\chi^2$  values of the fits range from 0.9 to 2.9.

Several systematic effects have been investigated. As mentioned above, humidity in the atmosphere increases the index of refraction and shifts the reconstructed  $X_{\max}$ . Since the simulations were performed for dry air this introduces an asymmetric systematic uncertainty of 10 g/cm<sup>2</sup>. The magnitude of the effect is expected to depend on zenith angle. Indeed, separating the event sample in two groups above or below 32 degrees zenith angle, yields a  $\pm 8$  g/cm<sup>2</sup> difference in mean  $X_{\max}$ . The choice of hadronic interaction model also introduces another uncertainty of 5 g/cm<sup>2</sup> in the reconstruction method<sup>2</sup>. The total systematic uncertainty on  $X_{\max}$  is +14/-10 g/cm<sup>2</sup> and on the energy is 27%. More details are given in [3] and [10].

## 5 Results

The mean value of  $X_{\max}$  as a function of energy is shown in the left panel of Fig. 2. The error bars indicate  $1\sigma$  statistical uncertainties and the shaded region shows the systematic uncertainty. The values are in agreement with other experiments using different techniques (fluorescence detection and non-imaging cherenkov). There is some tension with the data from the low-energy extension of the Pierre Auger observatory [19] towards  $10^{17.5}$  eV although the results still agree within systematic uncertainties.

<sup>2</sup>It is important to note that this is the uncertainty when *reconstructing*  $X_{\max}$ . When *interpreting*  $X_{\max}$  in terms of primary masses there is an additional difference between models of order 10 g/cm<sup>2</sup>.



**Figure 2.** Left: mean  $X_{\max}$  as a function of energy. The LOFAR results are in agreement with results based on different techniques within systematic uncertainties. The green lines are for pure proton composition and the red lines for pure iron (dashed: EPOS-LHC, solid: QGSJETII). Right: cumulative distribution of  $a$  for LOFAR data (blue; with  $1\sigma$  uncertainty band) and different compositions. A four-component model is required to reproduce the data.

Further information on the composition can be extracted by studying the shape of the distribution of  $X_{\max}$ . For each shower, we calculate:

$$a = \frac{\langle X_{\text{proton}} \rangle - X_{\text{shower}}}{\langle X_{\text{proton}} \rangle - \langle X_{\text{iron}} \rangle} \quad (1)$$

where  $X_{\text{shower}}$  is the reconstructed  $X_{\max}$ , and  $\langle X_{\text{proton}} \rangle$  and  $\langle X_{\text{iron}} \rangle$  are mean values predicted by the hadronic interaction code QGSJETII.04.

The cumulative probability density function (CDF) is shown in the right panel of Fig. 2. The yellow curves at the extremes represent pure proton (centered at  $a=0$ , wide distribution) and pure iron (centered at  $a=1$ , narrow distribution) composition.

We fit theoretical CDFs based on composition with two or four mass components to the data. The test statistic in the fit is the maximum deviation between the data and the model CDFs. The p-value is given by the probability of observing this deviation, or a larger one, assuming the fitted composition model. The two-component model consists of proton and iron nuclei, and has only one free parameter: the mixing ratio. The best fit is found for a proton fraction of 62%, but it describes the data rather poorly with a p-value of  $1.1 \times 10^{-6}$ .

The fit improves when using a four-component model (p+He+N+Fe), yielding a p-value of 0.17. While the best fit is found for a Helium fraction of 80%, the fit quality deteriorates only slowly when replacing helium by protons. Although there is a large uncertainty in the contribution of individual elements, the total fraction of light element (p+He) is well constrained. It has a best fit value of 0.8 and lies within the range [0.38,0.98] at a 99% confidence level based on QGSJETII.

## 6 Conclusion

It has been shown that the radio technique can be used to measure mass composition of cosmic rays and is competitive with other experiments. The first composition study with LOFAR yields a relatively high fraction of light particles just above  $10^{17}$  eV. The energy range can still be extended to lower and higher energies by expanding the triggering array and improving trigger strategies. Accurate composition measurements in the  $10^{16}$ - $10^{18}$  eV range are of crucial importance for understanding the transition of galactic to extragalactic origin of cosmic rays [7]. In the future, SKA will offer the opportunity of exploiting this technique with increased area, antenna density and frequency bandwidth, leading to extremely accurate measurements [5].

## References

- [1] T. Huege, *Physics Reports* **620**, 1-52 (2016).
- [2] W. Apel *et al.* [LOPES collaboration] *Phys. Rev. D* **90**, 062001 (2014).
- [3] S. Buitink *et al.* [LOFAR collaboration], *Phys. Rev. D* **90**, 082003 (2014).
- [4] M. P. van Haarlem *et al.* [LOFAR collaboration], *Astron. & Astrophys.* **556**, A2 (2013).
- [5] T. Huege *et al.*, these proceedings (2016); A. Zilles *et al.*, these proceedings (2016).
- [6] K.-H. Kampert and M. Unger, M., *Astropart. Phys.* **35**, 660 (2012).
- [7] S. Thoudam *et al.*, arXiv: 1605.03111 (2016).
- [8] S. Thoudam *et al.*, [LOFAR collaboration], *Nucl. Instr. and Meth. in Phys. Res. A* **767**, 339 (2014).
- [9] P. Schellart *et al.* [LOFAR collaboration], *Astron. & Astrophys.* **560**, A98 (2013).
- [10] S. Buitink *et al.* [LOFAR collaboration], *Nature* **531**, 70-73 (2016).
- [11] D. Heck, Report FZKA **6019** (1998).
- [12] T. Bergmann *et al.*, *Astropart. Phys.* **26**, 420-432 (2007).
- [13] A. Nelles *et al.*, *J. Cosm. Astropart. Phys.* **5**, 18 (2015).
- [14] T. Huege, M. Ludwig, and C. James, *AIP Conference Proceedings* **1535**, 128-132 (2012).
- [15] A. Nelles *et al.* [LOFAR collaboration], *JINST* **10**, P11005 (2015).
- [16] NOAA Air Resources Laboratory (ARL), Global Data Assimilation System (GDAS1) Archive Information, Tech. rep. (2004).
- [17] A. Corstanje *et al.* [LOFAR collaboration], these proceedings (2016).
- [18] A. Bonardi *et al.* [LOFAR collaboration], these proceedings (2016).
- [19] A. Porcelli *et al.* [Pierre Auger Collaboration], *PoS (ICRC2015)* 420; arXiv:1509.03732 (2015).