A new way of air shower detection: measuring the properties of cosmic rays with LOFAR

A. Nelles\textsuperscript{1}, S. Buitink\textsuperscript{1}, A. Corstanje\textsuperscript{1}, J.E. Enriquez\textsuperscript{1}, H. Falcke\textsuperscript{1,2,3,4}, J.R. Hörandel\textsuperscript{1,2}, J.P. Rachen\textsuperscript{1}, P. Schellart\textsuperscript{1}, O. Scholten\textsuperscript{5}, S. ter Veen\textsuperscript{1}, S. Thoudam\textsuperscript{1}, and T.N.G. Trinh\textsuperscript{5}

\textsuperscript{1} Department of Astrophysics/IMAPP, Radboud University, P.O. Box 9010, 6500 GL Nijmegen, The Netherlands
\textsuperscript{2} Nikhef, Science Park Amsterdam, 1098 XG Amsterdam, The Netherlands
\textsuperscript{3} Netherlands Institute for Radio Astronomy (ASTRON), Postbus 2, 7990 AA Dwingeloo, The Netherlands
\textsuperscript{4} Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany
\textsuperscript{5} KVI-CART, University of Groningen, P.O. Box 72, 9700 AB Groningen, The Netherlands
E-mail: a.nelles@astro.ru.nl

Abstract. High-energy cosmic rays impinging onto the atmosphere of the Earth initiate cascades of secondary particles: extensive air showers. Many of the particles in a shower are electrons and positrons. During the development of the air shower and by interacting with the geomagnetic field, the electromagnetic cascade creates radiation, which we detect at frequencies of tens of MHz with the LOFAR radio telescope in the Netherlands. After many years of struggling to understand the emission mechanisms, the radio community has achieved the breakthrough. We are now able to determine direction, energy, and type of the shower-inducing primary particle from the radio measurements. The large number of antennas at LOFAR allows us to have a high precision and very detailed measurements. We will elaborate on the shower reconstruction, a precise description of the intensity of the radio signal at ground level (at frequencies from 10 to 240 MHz), a precise measurement of the shape of the radio wavefront, and on the reconstruction of the shower energy.

1. Introduction
After three years of measuring the radio emission of air showers with LOFAR (the LOw-Frequency ARay), a number of key findings have been obtained. The vast number of antennas available at the core of LOFAR has allowed us to measure the signal distribution, polarization pattern and times of arrival with unprecedented quality.

These measurements have strengthened the case that detecting the radio emission of air showers has matured as a technique. The emission mechanisms are understood as a combination of the interaction of the shower with the magnetic field of the Earth [1, 2, 3] and the accumulation of a negative charge in the shower front [4, 5, 6]. Both effects have to be considered in a medium with changing index of refraction [7, 8], which also determines the regions and frequencies in which the signals are added up coherently.
Figure 1. Left: Low-band antennas at the core of LOFAR. The amplifier is located in the black top, above the stabilizing pole. In the background, the black tiles of the high-band antennas are visible. Right: Schematic view of the superterp. Open squares are the high-band antennas, crossed depict the low-band antennas and the full squares are the LORA detectors. Dashed lines indicate the grouping of LORA detectors for electronic and triggering purposes.

2. LOFAR – The LOw-Frequency ARray
LOFAR is currently the largest distributed radio interferometer for the lowest radio frequencies [9]. Two types of antennas are employed at LOFAR. The low-band antennas (LBA, $10^{-90}$ MHz) are slanted dipole antennas and used as standard tool for cosmic-ray operation. The high-band antennas (HBA, $110^{-240}$ MHz) are fat dipole antennas, optimized for astronomical interferometric observations.

The layout of LOFAR is based on stations. About one-hundred antennas of each type are grouped on an irregular grid into a cluster, called station. Stations are positioned all over Europe with increasing density towards the core of LOFAR in the north of the Netherlands. In the center, six stations are grouped on a circular elevated area, called the superterp. This area has also been instrumented with particle detectors [10], providing triggers for measurements of air showers. The LOFAR Radboud Air Shower Array (LORA) consists of 20 detector units, (a unit contains two 0.45$^2$m scintillators, type: NE 114), which are distributed between the antennas on an irregular grid. Whenever operation conditions allow and an air shower is detected in more than 13 detector units, a read-out of the transient buffer boards (TBBs) of the antennas is initialized. The antenna set and observation mode are always set by the primary astronomical observation, which also determines the data-stream that is buffered by the TBBs. Due to the multi-purpose capabilities of LOFAR, cosmic ray measurements are possible in the background of the primary observation. The TBBs store up to 5 s of data of a selected antenna type (LBA, HBA, inner or outer ring per station). Triggers can be received from inside and outside of LOFAR.

3. Data reconstruction
For every triggered data-set, the data from all selected LOFAR antennas, as well as the particle signal are available. Currently, 2.1 ms of radio data are stored following a trigger from the particle detectors. The data are processed automatically every night, delivering relatively calibrated pulse powers per antennas for air showers detected above the noises threshold [11]. An absolute calibration is currently underway [12].

The raw voltage traces from all antennas per station are first cleaned from narrow-band interference. Furthermore, relative differences between antennas are corrected for, by
normalizing the average received power to a value corresponding to the expected galactic power as function of local sidereal time [11]. Using the arrival time and arrival direction reconstructed from the particle signal, the cleaned antenna signals are beamformed in order to identify the timing of a detectable air shower signal. This allows us to narrow the search window and to determine the air shower signal in absolute timing in an additional pulse-finding step. Using this time, the signals from the two dipoles are combined and corrected for the antenna response, while refitting the arrival direction iteratively. Once the final arrival direction is reached, the pulse quantities such as power, polarization, timing and corresponding uncertainties are calculated. If a pulse is successfully identified in more than one station, the results are combined and further analysis steps like a non-planar wavefront-fit and fit of the signal pattern are performed.

All results are accessible from a PostgreSQL database. Additional quality cuts, in order to exclude periods with thunderstorms or contamination with broad-band interference are applied at a later stage, allowing us to also review these contaminated measurements. The details of the reconstruction framework can be found in [11].

4. Polarization and emission mechanisms

The two main emission mechanisms can be identified by their polarization signature. The geomagnetic effect, the interaction of the shower with the magnetic field of the Earth, will induce electric fields that are polarized in the $\vec{v} \times \vec{B}$-direction, where $\vec{v}$ is the direction of propagation of the shower and $\vec{B}$ the direction of the magnetic field at the detector. This direction of polarization will be independent of the location of the observing antenna around the shower axis. The charge excess, however, being an effect that is symmetric around the shower axis, will induce an electric field that is dependent on the position of the observer. It will therefore be changing from parallel to anti-parallel to the geomagnetically induced electric field, when moving around the shower axis. LOFAR is highly suited to observe such changes on an event-by-event level or as function of distance to the shower axis [13].

Especially, the calculation of the fraction of charge excess $a$ has been proven instructive. The air showers measured with LOFAR contain signals that show a degree of nearly 99% of polarization. As the geomagnetic and the charge-excess emission are linearly polarized in different directions, their relative contributions can be determined. Following [6] the expected electric field at any given time $t$ can be written as

$$\vec{E}(t) = \vec{E}_G(t) + \vec{E}_C(t)$$

$$= (|\vec{E}_G(t)| + |\vec{E}_C(t)| \cos \phi')\hat{e}_{\hat{v} \times \hat{B}} +$$

$$((|\vec{E}_C(t)| \sin \phi')\hat{e}_{\hat{v} \times \hat{v} \times \hat{B}}.$$  

Here $\vec{E}_G(t)$ is the electric field produced by the geomagnetic contribution that is purely polarized along the direction of the Lorentz force, and $\vec{E}_C(t)$ the radial electric field, produced by charge-excess. The charge-excess fraction is then defined as

$$a \equiv \sin \alpha \frac{|E_C|}{|E_G|},$$

where $|E_C|$ is the amplitude of the electric field produced by charge-excess when the total electric field vector amplitude reaches its maximum value.

The polarization signal of one air shower measured with LOFAR is shown in figure 2. The dominant polarization in the $\hat{v} \times \hat{B}$-direction is clearly visible. The additional charge excess fraction needed is small, however, measurable. The average contribution is shown on the right of figure 2. The charge excess fraction is a function of distance to the shower axis and zenith angle varying between close to zero and 20%. This indicates that both emission mechanisms
Figure 2. Left: Direction of the measured electric field vector for an air shower as detected with LOFAR. The vectors are shown in the shower plane together with their (small) uncertainties. The main direction is defined by the cross-product of propagation direction of the shower $\vec{v}$ and the geo-magnetic field $\vec{B}$. Right: The parameter $a$, as fractional deviation from a purely geo-magnetic polarization, as function of distance to the shower axis. The relation is shown for different zenith angle bins [13].

show a different behavior with the zenith angle, possibly the shower development and the lateral extent of the shower. Details of this analysis can be found in [13].

5. Wavefront and arrival directions

The arriving front of radio signals can be approximated to reasonable resolution with a planar wavefront. At a more closer look, it however becomes apparent that there is significant curvature in the wavefront [14, 15]. With a timing resolution of better than nanoseconds, and several hundreds of antennas, LOFAR is the ideal tool to study the details of the wavefront.

From basic assumptions, three models of a wavefront can be motivated. A sphere, a cone and a hyperbola. In order to avoid being dominated by changing conditions between air showers, these three shapes are compared on an event-by-event basis.

The inspection of a single air shower is already intuitive. Figures 3 and 4 show the three wavefront shapes in their best-fitting version for one air shower. The core position was used as a free parameter in order to allow for a fair comparison. This explains the different distribution of measurements. It is clearly visible that using a spherical fit results in too much curvature, which causes a problem at the antennas of larger distances to the shower axis. The conical fit, however, cannot be used to explain all curvature and therefore residual curvature can be seen in the bottom half of the figure. Only with the hyperbola the right balance of curvature can be obtained.

Since one needs one additional free parameter for the hyperbola, all three models cannot be compared based directly on their $\chi^2$. Furthermore, the hyperbola can turn into a cone or a sphere for the appropriate choice of parameters. Therefore, a likelihood ratio test is performed, given the fact that both cone and sphere are a subset of the hyperbola. This results in a probability of effectively zero of either cone or sphere matching the dataset better than the hyperbola. More details of this analysis can be found in [14].

The right image in figure 4 shows the advantage of using a curved wavefront in order to fit the arrival direction. For the reconstructed arrival direction, it is not very important which form you chose. All parameterization result in the same arrival direction with a resolution of...
Figure 3. Two best fitting solutions of a conical (left) and spherical (right) wavefront. Shown are the arrival times in the shower plane as function of distance to the axis. Also shown are the residuals of the fit and data in units of the uncertainties [14].

Figure 4. Left: Best fitting solutions of a hyperbolic wavefront to the arrival times of an air shower measured with LOFAR. Right: The difference in reconstruction of the arrival direction between the fit of an hyperbola and a cone. The comparison to a spherical fit looks comparable [14].

better than 0.1°. When comparing this arrival direction, however, to the results of a planar wavefront, the spread increases to about 1°. The same holds true for the comparison to the arrival directions as obtained with the particle detectors. As those already have an independent resolution of about 1°, the resolution of the curved wavefront needs to be significantly better, which is essential once arrival directions, anisotropies and possible sources are studied.

6. Energy and the depth of the shower maximum
If radio is to be used as standard detection technique, it will need to provide an independent measure of the energy of the cosmic ray and of the depth of the shower maximum, $X_{\text{max}}$. The data taken with LOFAR have been used to show that this is indeed possible with an excellent resolution. Two approaches have been tested. A direct comparison to simulations, as will be
explained in section 6.1 and a parametric approach that will be discussed in section 6.2.

They both utilize the following quantities: The integrated pulse power as measured in the low-band antennas, the particle densities as measured with the LORA scintillators and derived quantities.

6.1. Direct comparison to air shower simulations

Air shower simulations predict a strong dependence of the shape of the signal distribution on the distance to the shower maximum. Given the large number of antennas at LOFAR, it is possible to test this prediction without a parameterization step in-between.

CORSIKA [16] and its radio extension CoREAS [17] are used to generate emission maps for several depths of the shower maximum, X\text{max}, for a given air shower using its arrival direction and energy estimate from the particle detectors as input. The maps are generated using 160 ground positions at which the radio emission is evaluated. These ground positions are located on a star-pattern with eight arms, which is symmetric in the shower plane and centered around the shower axis. Furthermore, two arms are aligned with the \( \vec{v} \times \vec{B} \)-direction and two are aligned with the \( \vec{v} \times (\vec{v} \times \vec{B}) \)-direction. This star-shaped pattern allows us to sample the shower at the relevant positions. In the \( \vec{v} \times \vec{B} \)-direction the largest signal is expected due to the positive interference of geomagnetic emission and charge-excess. Starting from these individual simulations, full patterns can be extrapolated as shown in figure 5. Using a robust interpolation routine delivers the total power at every given position with an uncertainty of about 2.5\% with respect to dedicated simulations for a given position [18].

For every measured shower, about 40 of these patterns have to be generated in order to sample the ranges of X\text{max} sufficiently. It can now be tested, which pattern best describes the data by letting the shower position vary. CORSIKA also delivers the particles generated by the air shower. The particle content can therefore also be matched to what was measured with the LORA detectors. The following fit procedure is used:

\[
\chi^2 = \sum_{\text{antennas}} \left( \frac{P_{\text{ant}} - f^2 P_{\text{sim}}(x_{\text{ant}} - x_0, y_{\text{ant}} - y_0)}{\sigma_{\text{ant}}} \right)^2 + \sum_{\text{particle detectors}} \left( \frac{d_{\text{det}} - f d_{\text{sim}}(x_{\text{det}} - x_0, y_{\text{det}} - y_0)}{\sigma_{\text{det}}} \right)^2, \tag{5}
\]

Figure 5. Two simulated signal pattern of an air shower. The left pattern was derived for a shower that is observed at a large distance to X\text{max}, the right pattern for a small distance [19].
Figure 6. Left: The goodness of fit parameter $\chi^2$/ndf for 40 air shower simulations (squares iron, circles proton primaries), all differing by the depth of shower maximum, $X_{\text{max}}$, of the air shower. The inset shows a zoom of the minimum of the distribution. Right: Distribution of uncertainties on $X_{\text{max}}$ as obtained for the first 50 high-quality air showers measured with LOFAR [18].

where $P_{\text{ant}}$ is the measured power integrated over a 55 ns window at an antenna at location $(x_{\text{ant}}, y_{\text{ant}})$ with noise level $\sigma_{\text{ant}}$. $P_{\text{sim}}$ is the simulated power, $d_{\text{det}}$ is the deposited energy as measured by a LORA detector at location $(x_{\text{det}}, y_{\text{det}})$ with noise $\sigma_{\text{det}}$, and $d_{\text{sim}}$ is the simulated deposited energy. The scaling factors $f_r$ and $f_p$ are needed due to the missing absolute calibration and potential uncertainties in the particle reconstruction.

As shown in figure 6 the distribution of $\chi^2$-values shows a clear preference for a range of $X_{\text{max}}$ for a given shower, which allows us to determine the $X_{\text{max}}$ of every shower with a very accurate estimate. The obtainable uncertainty stems from the method itself, the atmospheric correction and the uncertainty in the arrival direction [18]. The choice of hadronic interaction model seems not to be very relevant for this kind of analysis. The largest offset observed between, for example, QGSJETII-04 and EPOS-LHC was 4.3 g/cm$^2$. The same holds true for the choice of simulation code, for example when using ZHAires [20] instead of CoREAS. For the first fifty air showers measured at LOFAR, the distribution of uncertainties is shown on the right in figure 6. With a mean uncertainty of 17 g/cm$^2$ the results are highly competitive with respect to for example fluorescence detectors and encouraging to develop the radio technique further.

6.2. Parameterization of the signal distribution

The only drawback of the method described in section 6.1 is the large computational effort. In the current set-up, several weeks of CPU time are needed to generate a sufficient number of patterns for an air shower. The effort could be limited, if the energy was precisely known for every shower (the particle detectors are small and subject to large systematic uncertainties due to the missing sensitivity to shower-to-shower fluctuations) and if the range of $X_{\text{max}}$-values could be limited.

In [19] an analytic parameterization was suggested that can be used for a fast reconstruction of the shower axis. Furthermore, two fit parameters of the function are predicted to show a strong correlation with the energy of the shower and the distance to the shower maximum.
Based on [19] the following function has been tested on LOFAR data:

\[
P(x', y') = A_+ \cdot \exp \left( \frac{-\left[(x' - X_c)^2 + (y' - Y_c)^2\right]}{\sigma^2} \right) - C_0 \cdot A_+ \cdot \exp \left( \frac{-\left[(x' - (X_c + x_-))^2 + (y' - Y_c)^2\right]}{(C_3 \cdot e^{C_1 + C_2 \sigma^2})^2} \right)
\]  

The function is given in coordinates \((x', y')\) in the shower plane, where the coordinates are aligned with the \(\vec{v} \times \vec{B}\)- and the \(\vec{v} \times (\vec{v} \times \vec{B})\)-axis. The parameters \(C_1, C_2\) and \(C_3\) are fixed from studies of simulations. All other parameters are used for the fit.

More than 85% of all showers recorded in three or more LOFAR stations can be reconstructed with the parameterization with good quality of fit. Showers that cannot be reconstructed are mostly showers with low signal-to-noise ratios and those where the shower axis is far away from the closest radio antennas. The shower axis that is reconstructed with the parameterization is in very good agreement with the axis obtained with the method of the direct comparison, as evidenced in figure 7. Here, the shape of the signal distribution as function of the distance to the shower axis would change immediately, if the axis was shifted. For this air shower, the axis position differs by less than three meters. The average spread between the two methods is less than 15 m. The small differences in signals and uncertainties are due to the fact that one method is applied on the instrumental signal (no correction for the antenna model on data, but application of antenna model on the simulation) and the other on the data (including its noise) corrected for the influence of the antenna model.

On data, it is also confirmed that the parameter \(A_+\) shows a strong correlation with the energy of the primary particle. It scales quadratically with the energy, as it is a power quantity and the amplitude scales linearly with the energy due to the coherence criterion. Compared to the energy estimate given by the LORA particle detectors, the \(A_+\)-parameter is usable for more showers. First of all, the radio antennas cover a larger portion of the surface with provides a larger fiducial area. Also, the radio signal pattern can still be reconstructed if the shower axis was outside the instrumented area, which is due to the asymmetry in the signal. For the whole data-set only about half of the air showers detected in radio can reliably be reconstructed with the particle signal. Another advantage of the radio detection is that it almost delivers a calorimetric
signal. The geometric component of the shower is represented in the $\sigma_+\,-$parameter. Therefore, the $A_+\,-$parameter is expected to have small systematic uncertainties ($<10\%$) with respect to the primary composition. In order to provide a good energy estimate for the simulation input, this missing uncertainty will help to significantly reduce the simulation effort, while the larger data-set allows for the retaining of more showers, which improves composition studies. Details on the application of the parameterization to the LOFAR data can be found in [21].

7. Signal distribution in different frequency ranges

Most air showers are measured at LOFAR with the low-band antennas covering the frequency range of $10-90\,\text{MHz}$. There are two reasons for this choice. First of all, the emission footprint is expected to be stronger and spread wider at lower frequencies and the first observations have also been conducted in this range. The second reason is more technical and LOFAR-specific. The high-band antennas $110-240\,\text{MHz}$ have been optimized for astronomical observations. In this frequency range, the antennas are system noise dominated, which calls for a direct enhancement of the astronomical signals to ensure a good sensitivity. Thus, the signals from 16 antennas are combined directly in hardware before being forwarded to the TDBs. In this combination step a time delay is introduced making the combined set of antennas more sensitive to a fraction of the sky that corresponds to the induced delays. Signals from other directions will be distorted. This consequently means that air showers can only be measured with best sensitivity, if their arrival direction happens to coincide with the $10^{\circ}$ field of view of the observation. This essentially limits the usefulness of the high-band antennas for studies of air showers.

However, as the primary observation determines the antenna set, there is air shower data available in the higher frequency band. Also, a small number of signals have been recorded without distortion and can be compared to simulations with the same approach as described above.

Figure 8 shows an air shower as measured in the frequency range of $110-190\,\text{MHz}$. In this frequency range, a clear ring-like structure is visible. This can be explained by the increasing importance of the relativistic time-compression for higher frequencies. At higher frequencies,
the coherence is no longer given at all observer positions around the shower, which decreases the signal strength measured. At the Cherenkov angle, however, an enhancement is visible that translates into a ring of strong emission. LOFAR is the first experiment to measure this ring on an event-by-event basis and to confirm the theoretical predictions.

This ring size has been postulated to correlate with the depth of the shower maximum [23]. Indications for such a correlation are indeed found via the direct comparison to simulations. However, a much larger set is needed to confirm this relation. Furthermore, the ring size would have to be estimated with a high precision to obtain depths of the shower maximum that are well enough constrained for composition studies. Even with the large number of antennas at LOFAR, this precision is difficult to obtain. Details of this analysis can be found in [22].

8. Conclusion
The measurement and the reconstruction of the radio emission of air showers have made significant progress over the last years. The radio-telescope LOFAR with its high antenna density has been highly suitable to measure the subtle features of the emission and confirm theoretical predictions on a single-event basis. The emission shows a predominant polarization in the direction defined by the cross-product of shower axis and magnetic field, which can be attributed to the dominant emission process. A second emission process is needed to explain the secondary radial component in the signal polarization. This charge excess fraction is not constant, but is a function of the distance to the shower axis and the zenith angle of the shower. Also, we have shown that the wavefront of the radio signal is best represented by a hyperbola. Conical or spherical shapes introduce too little or too much curvature and cannot reproduce the data convincingly. A milestone in the understanding of the emission was the direct comparison of air shower simulations with the LOFAR data. Not only were all asymmetric shapes very well reproduced, but the strong dependence on the depth of the shower maximum was proven. With LOFAR data the shower maximum can be resolved with an average uncertainty of 17 g/cm², which makes the radio detection of air showers a strong competitor for composition studies. Based on these simulations an analytical parameterization was developed that can provide a fast reconstruction of the air showers. This parameterization will help to speed up the detailed energy and composition studies. In addition, LOFAR data also shows that the emission mechanisms are well understood also beyond the standard frequency range of below 100 MHz. At higher frequencies the emission is clearly dominated by Cherenkov-like effects and ring structures of enhanced emission appear.

Acknowledgments
The LOFAR cosmic ray key science project very much acknowledges the scientific and technical support from ASTRON. Furthermore, we acknowledge financial support from the Netherlands Research School for Astronomy (NOVA), the Samenwerkingsverband Noord-Nederland (SNN), the Foundation for Fundamental Research on Matter (FOM) and the Netherlands Organization for Scientific Research (NWO), VENI grant 639-041-130. We acknowledge funding from an Advanced Grant of the European Research Council under the European Union’s Seventh Framework Program (FP/2007-2013) / ERC Grant Agreement n. 227610.

LOFAR, the Low Frequency Array designed and constructed by ASTRON, has facilities in several countries, that are owned by various parties (each with their own funding sources), and that are collectively operated by the International LOFAR Telescope (ILT) foundation under a joint scientific policy.
References


[21] A. Nelles et al., The radio emission pattern of air showers as measured with LOFAR - a tool for the reconstruction of the energy and the shower maximum, Submitted to JCAP (2014) [arXiv:1411.7868].
