CONFUSION LIMITED ALL SKY IMAGING WITH LOFAR'S INITIAL TEST STATION APPLYING WIDE FIELD CALIBRATION TECHNIQUES

Stefan J. Wijnholds[1]

[1]*ASTRON, R&D, Oude Hoogeveensedijk 4, P.O. Box 2, 7990 AA Dwingeloo, The Netherlands*
Email: wijnholds@astron.nl

ABSTRACT

In this paper an all sky map at 30 MHz made with LOFAR's first full scale prototype station will be presented. This paper will provide an extensive description of the required data reduction techniques which include wide field imaging and calibration, spectral and spatial RFI detection and source subtraction. Due to observational conditions the SNR varies over the map. It will be demonstrated that in regions with sufficient SNR the map is confusion limited instead of noise limited. This indicates that the presented data model and reduction techniques produce proper results up to the point where the confusion limit is reached.

1. INTRODUCTION

The Initial Test Station (ITS) is the first full scale prototype of a LOFAR station. It consists of 60 sky noise limited inverse V-shaped dipoles with East-West orientation arranged in a five-armed spiral configuration offering an instantaneous synthesized aperture of almost 200 m diameter. Each antenna samples a superposition of signals coming from all directions on the sky in the same single polarization. The signal from each antenna is filtered by a 10-40 MHz band pass filter, digitized by a 12 bit 80 MHz A/D converter and stored in a 1 GB data buffer for further processing allowing snapshot integration times up to 6.7 s.

The ability to produce sky noise limited hemispheric snapshot images [1] poses a calibration challenge since there is not just one single dominant source within the field of view which can be used for calibration. Instead the complex electronic gains of the receiver chains need to be estimated based on an input signal which is a superposition of a multitude of sources. This makes the calibration more complicated compared to the calibration methods used for "classical" radio telescopes. On the other hand the stations are small enough to ensure that for each of the individual sources all dipoles are looking through the same piece of the ionosphere.

In this paper I will demonstrate that proper wide field imaging and calibration can be done by applying common techniques from signal processing. After presenting the required data reduction techniques, ITS's ability to do all sky imaging will be demonstrated. At the end the confusion limitedness of the final map will be discussed.

Notation: The transpose operator is denoted by ^{T}, the Hermitian transpose by ^{H}, an estimated value by $\hat{\cdot}$ and an expected value by $\mathcal{E}\{\cdot\}$. $\oslash$ is the elementwise matrix division. $\mathrm{diag}(\cdot)$ when applied to a vector converts the vector into a diagonal matrix with the vector placed on the main diagonal, when applied to a matrix it converts the main diagonal of the matrix to a column vector. $\|\cdot\|_F$ denotes the Frobenius norm of the argument, which is the square root of the sum of the squared absolute values of the individual entries of the argument. $\mathbf{I}$ denotes the identity matrix.

2. DATA REDUCTION TECHNIQUES

Following [2], consider a telescope or antenna array with p elements. Let the output signal of the i^{th} antenna be denoted by $x_i(t)$, and define the array signal vector $\mathbf{x}(t)$ by $\mathbf{x}(t) = [x_1(t), \cdots, x_p(t)]^T$. Assume that the narrow band condition holds, which implies that geometrical delay of a signal can be represented by a phase shift of the signal. Consider further q (astronomical) source signals $s_k(t)$ with $k = 1 \cdots q$. Denote the spatial signature vector of each of the sources k by $\mathbf{a}_k = [\mathrm{e}^{-\mathrm{j}\mathbf{k}_k\mathbf{r}_1}, \cdots \mathrm{e}^{-\mathrm{j}\mathbf{k}_k\mathbf{r}_p}]^T$, where $\mathbf{k}_k$ is the propagation vector of the incoming plane wave from source k and $\mathbf{r}_i$ is the position vector of element i, Furthermore, consider telescope noise signals $n_i(t)$ stacked in a $p \times 1$ vector $\mathbf{n}(t)$. Further let

the telescope dependent gains g_i be stacekd in a vector $\mathbf{g} = [g_1, \cdots, g_p]^T$, and in diagonal matrix form by $\mathbf{\Gamma} = \mathrm{diag}(\mathbf{g})$. Using these definitions, the array signal vector can be expressed as

$$\mathbf{x}(t) = \mathbf{\Gamma} \left(\sum_{k=1}^{q} \mathbf{a}_k s_k(t) \right) + \mathbf{n}(t) \tag{1}$$

Consider an observation in which the signal is sampled with sample period T, and define the data sample matrix $\mathbf{X} = [\mathbf{x}(T), \mathbf{x}(2T), \cdots, \mathbf{x}(NT)]$. The array correlation matrix (ACM) $\widehat{\mathbf{R}}$ takes the form $\widehat{\mathbf{R}} = N^{-1}\mathbf{X}\mathbf{X}^H$. Define the source power by $\sigma_{s_k}^2 = \mathcal{E}\{|s_k(t)|^2\}$ and stack the source signals $s_k(t)$ in a $q \times 1$ vector $\mathbf{s}(t)$. Assuming the sources are mutually independent identically distributed Gaussian signals, the source signal covariance $\mathbf{\Sigma}_s = \mathcal{E}\{\mathbf{s}(t)\mathbf{s}(t)^H\}$ is diagonal: $\mathbf{\Sigma}_s = \mathrm{diag}(\boldsymbol{\sigma}_s)$, where $\boldsymbol{\sigma}_s = [\sigma_{s_1}^2, \cdots, \sigma_{s_q}^2]^T$. Assuming the noises $\mathbf{n}(t)$ are mutually independent identically distributed Gaussian signals as well results in a diagonal covariance matrix $\mathbf{\Sigma}_n = \mathrm{diag}(\boldsymbol{\sigma}_n)$, where $\boldsymbol{\sigma}_n = [\sigma_{n_1}^2, \cdots, \sigma_{n_p}^2]^T$ and $\sigma_{n_i}^2 = \mathcal{E}\{|n_i(t)|^2\}$. Assuming the spatial signature vectors are deterministic, and are stacked in a $p \times q$ matrix $\mathbf{A}$, the expected value $\mathbf{R} = \mathcal{E}\{\widehat{\mathbf{R}}\}$ has model

$$\mathbf{R} = \mathbf{\Gamma}\mathbf{A}\mathbf{\Sigma}_s\mathbf{A}^H\mathbf{\Gamma}^H + \mathbf{\Sigma}_n \tag{2}$$

Beam steering in phased array telescopes can be done by applying complex weights w_i to the individual elements such that signals from the selected direction are added coherently. These weights can be stacked in a $p \times 1$ beam steering vector $\mathbf{w} = [w_1, w_2, \cdots, w_p]^T$. In [3] it is shown that the expected power output of the tied array beam former is

$$\mathcal{E}\{|\mathbf{w}^H\mathbf{x}(t)|^2\} = \mathbf{w}^H\mathbf{R}\mathbf{w}. \tag{3}$$

In (2) the factor $\mathbf{A}\mathbf{\Sigma}_s\mathbf{A}^H$ contains the information on the spatial distribution and powers of the sources. According to the data model this factor is distorted by the telescope dependent gains g_i. This data model can be used to solve for these element specific variations by minimizing the difference between the model and the actual data with respect to $\mathbf{g}$ and $\boldsymbol{\sigma}_n$:

$$\{\hat{\mathbf{g}}, \hat{\boldsymbol{\sigma}}_n\} = \min_{\mathbf{g}, \boldsymbol{\sigma}_n} \parallel \mathbf{\Gamma}\mathbf{A}\mathbf{\Sigma}_s\mathbf{A}^{\mathrm{H}}\mathbf{\Gamma}^{\mathrm{H}} + \mathbf{\Sigma}_n - \widehat{\mathbf{R}} \parallel_F^2 \tag{4}$$

In [4] several approaches are presented to solve this minimization problem. These algorithms assume that the positions and powers of the sources in the field and therefore $\mathbf{A}$ and $\mathbf{\Sigma}_s$ are known. Unfortunately there is always some extended emission due to the galactic plane and the North Polar Spur which is hard to model with a discrete point source model. Therefore extended emission was filtered out by using only baselines longer than 4 wavelengths during calibration.

Since the data model used for calibration is based on celestial sources, the calibration algorithm should be applied to RFI free channels. This can be achieved by using either a spectral or a spatial RFI detection algorithm. The spectral RFI detection algorithm calculated the median power level over the 51 channel frequency interval centered on the channel of interest. The standard deviation is assumed to be equal to this median value divided by the square root of the product of bandwidth and integration time. The algorithm uses a threshold three standard deviations above the median power. This procedure was repeated for each of the individual elements. If the power in the channel of interest was lying below the treshold for each of the individual elements, the channel was regarded RFI free.

The spatial RFI detection algorithm uses the Frobenius norm of the ACM $\widehat{\mathbf{R}}$ as measure for the amount of spatial structure in the map. To prevent gain variations over frequency from influencing the result, a normalized version of the ACM $\widetilde{\mathbf{R}} = \widehat{\mathbf{R}} \oslash \sqrt{\mathrm{diag}(\widehat{\mathbf{R}})\mathrm{diag}(\widehat{\mathbf{R}})^H}$ is used. The amount of spatial structure is thus quantified by

$$\parallel \widetilde{\mathbf{R}} \parallel_F^2 = \mathrm{tr}\left(\widetilde{\mathbf{R}}\widetilde{\mathbf{R}}^H\right). \tag{5}$$

To illustrate the operation of this measure, suppose that $\widehat{\mathbf{R}}$ is completely dominated by receiver noise. In that case $\mathcal{E}\{\widehat{\mathbf{R}}\} = \sigma_n^2\mathbf{I}$ assuming an array of p identical elements giving $\widetilde{\mathbf{R}} \approx \mathbf{I}$ and $\parallel \widetilde{\mathbf{R}} \parallel_F^2 \approx p$. If $\widehat{\mathbf{R}}$ is dominated by a single point source, $\mathcal{E}\{\widehat{\mathbf{R}}\} = \sigma_s^2\mathbf{a}\mathbf{a}^H$. In this case $\widetilde{\mathbf{R}} \approx \mathbf{a}\mathbf{a}^H$ giving $\parallel \widetilde{\mathbf{R}} \parallel_F^2 \approx p^2$.

Since the beam pattern varies only gradually with frequency, the amount of structure in the map should vary only gradually with frequency as well. By putting a threshold on the difference in the variance of the normalized ACMs of adjacent frequency channels, this measure can be used for spatial RFI detection.

After successful calibration the result produced by (3) depends on the distribution of sources in the field and the array beam pattern. The latter may cause contamination of weaker sources in the field by the sidelobes of strong sources. If the

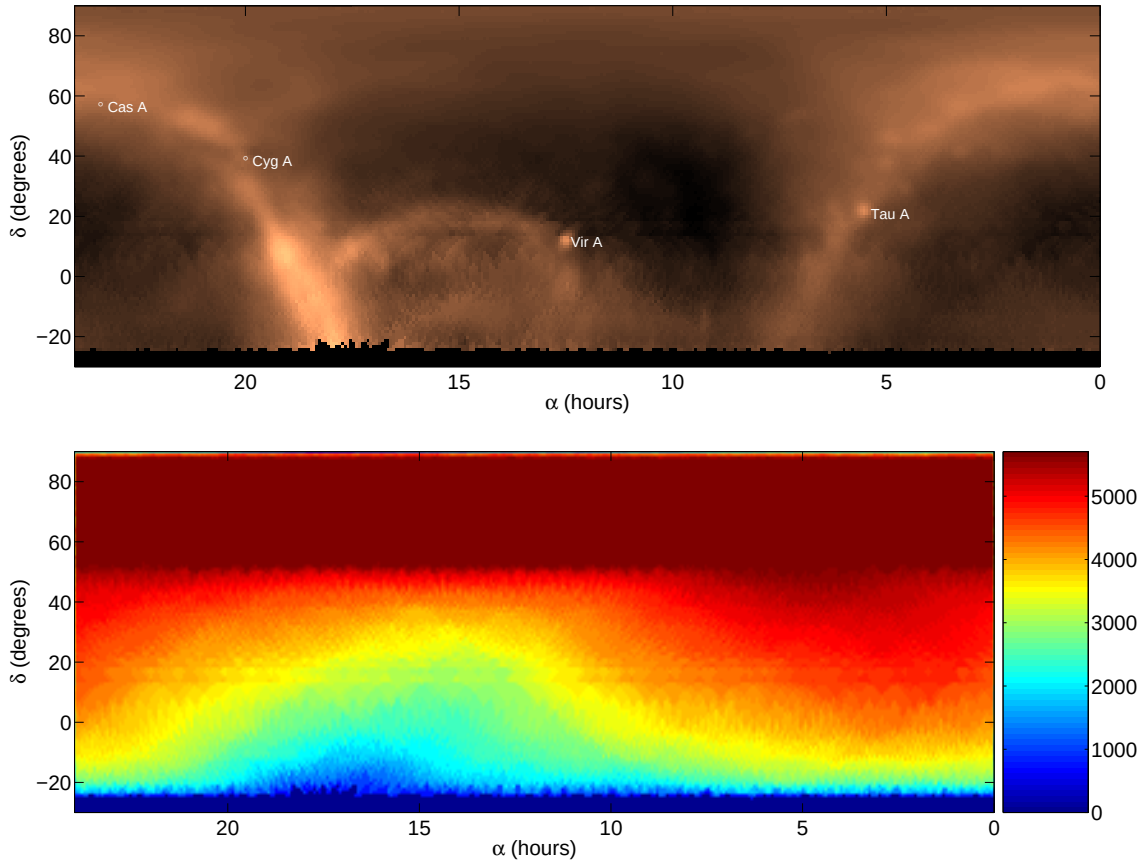

Fig. 1. The upper image shows an all sky map based on 56 6.7 s hemispheric snapshot integrations using all RFI free frequency channels between 29.5 and 30.5 MHz. Cas A and Cyg A have been removed to avoid contamination of weak source by their sidelobes. The image in the lower panel shows the distribution of $\sqrt{B\tau}$ over this map.

spatial signature vector $\mathbf{a}$ and power σ_s^2 of a strong source are known, it may be subtracted from the measured ACM [5] producing the reduced ACM

$$\widehat{\mathbf{R}}_{-} = \widehat{\mathbf{R}} - \sigma_s^2 \mathbf{a}\mathbf{a}^H. \tag{6}$$

3. MEASUREMENT RESULTS

Between September 24 and October 9, 2004 a total of 86 hemispheric 6.7 s snapshot observations were done with ITS. An 8192 point FFT with Hanning Window was applied to the data to produce spectra with 9.77 kHz frequency bins. To reduce the computational load only the channels within the pass band of the analog filter (channels 1050 through 3900) were correlated. From this data all RFI free channels between 29.5 MHz and 30.5 MHz were selected for further processing. Snapshots having less than 50 RFI free channels in this frequency range were rejected from further processing leaving 56 snapshots.

After calibration using (4) on the individual frequency channels for every snapshot, Cas A and Cyg A were removed by applying (6) to avoid contamination of weaker sources in the map by their sidelobes. Imaging using (3) was done separately for every frequency channel before adding all images from a single snapshot observation. The resulting snapshot images were projected on an (α, δ) grid with 1° resolution and averaged weighted by the square root of band width and integration time of the snapshot observation. The resulting map is shown in the upper panel of Fig. 1.

Extended emission from the galactic plane and the North Polar Spur is clearly visible in the image due to the considerable number of short baselines in the ITS configuration. A closer examination of the region around Tau A reveals a number of blobs which have been identified as sources from the 3C catalog (e.g. 3C157, 3C123, 3C134 and Per A). Due to the projection

effect the resolution of the snapshot images is varying with elevation. Since not all sources reach the same elevation the resolution is varying over the map. Most sources however will be measured with an average resolution of about 4°.

The SNR is varying over the map as well due to the difference in the product of bandwidth and integration time that has been achieved on different regions of the sky. The quality of the map in the upper panel of Fig. 1 can therefore only be judged by the image in the lower panel which shows the variation in the square root of the product of bandwidth B and integration time τ for the all sky map.

At the time of observation the sources at hour angles between 10 h and 20 h were measured during day time. Since we detect more RFI at day time than at night time, the usable number of frequency channels at day time is smaller which reduces the SNR achievable in day time measurements. The gradient with declination can be explained by the fact that when the declination is lower, the maximum elevation reached by the source is lower causing it to be visible for a shorter period of time. This restricts the achievable integration time on sources at lower declination.

4. CONFUSION LIMIT

At a spatial resolution (half power beam width) of 4° about 260 independent beams per steradian can be defined. To detect sources in a map separately one requires on average 10 beams per source. Therefore ITS is able to detect individual sources as long as their fluxes are above its detection limit and as long as their density on the sky is less than 26 sources per steradian.

The detection limit of ITS was determined by taking the difference of a map based on only the even channels and a map based on only the odd channels of the individual snapshots. Since the SNR is not constant over the sky a $20^\circ \times 20^\circ$ region centered on Tau A was selected. The standard deviation of the image values in this region was taken as estimate of the 1σ noise level in this region of the map.

Using a contour plot of the region around Tau A with contours every 10σ, Tau A was found to be a 480σ source. Since the flux of Tau A is 1420 Jy at 178 MHz, 1σ would correspond to 3 Jy at 178 MHz. The complete 3C catalog [6] with its limiting flux of 9 Jy should therefore be detectable at 3σ level. The total number of sources in the 3C catalog with $\delta \geq 0^\circ$ is 303, which gives a source density of 48.2 sources per steradian. We may therefore conclude that most regions of the map presented in the previous section are limited by source confusion, not by SNR.

5. CONCLUSIONS

In this paper a confusion limited all sky map at 30 MHz made with LOFAR's first full scale prototype station was presented. The successful result of the data reduction pipeline indicates that the data model and reduction procedures presented in this paper produce proper results up to the point where the confusion limit is reached.

6. REFERENCES

[1] S.J. Wijnholds, J.D. Bregman, and A.J. Boonstra, "Sky Noise Limited Snapshot Imaging in the Presence of RFI with LOFAR's Initial Test Station," *Exp. Astr.*, 2005, in press.

[2] A. Leshem, A.J. van der Veen, and A.J. Boonstra, "Multichannel Interference Mitigation Techniques in Radio Astronomy," *The Astrophysical Journal Supplement Series*, vol. 131, no. 1, pp. 355–373, November 2000.

[3] Dimitris G. Manolakis, Vinay K. Ingle, and Stephen M. Kogon, *Statistical and Adaptive Signal Processing*, McGraw-Hill, 2000.

[4] A.J. Boonstra, S.J. Wijnholds, S. van der Tol, and B. Jeffs, "Calibration, Sensitivity and RFI Mitigation Requirements for LOFAR," in *2005 IEEE International Conference on Acoustics, Speech, and Signal Processing*, March 2005.

[5] A.J. Boontra and S. van der Tol, "Spatial Filtering of Interfering Signals at the Initial LOFAR Phased Array Test Station," *Radio Science*, in press.

[6] A.S. Bennett, "The Revised 3C Catalog of Radio Sources," *Memoirs of the Royal Astronomical Society*, vol. 68, pp. 163–172, 1962.