A COMPUTATIONALLY EFFICIENT CALIBRATION ALGORITHM FOR THE LOFAR RADIO ASTRONOMICAL ARRAY

Yuntao Wu\textsuperscript{1,2}, Amir Leshem\textsuperscript{1} and Stefan J. Wijnholds\textsuperscript{3}

\textsuperscript{1} Faculty of Engineering, Bar-Ilan University, Ramat-Gan, 52900, Israel
\textsuperscript{2} School of Computer Science and Engineering, Wuhan Institute of Technology, Wuhan, 430205, China
\textsuperscript{3} ASTRON, Oude Hoogeveensedijk 4, 7991 PD Dwingeloo, The Netherlands
Email: leshem.amir2@gmail.com

ABSTRACT

In this paper, the problem of self-calibration for large astronomical arrays such as the Dutch Low Frequency Array (LOFAR) is considered. We assume direction dependent gain and phase errors which need to be estimated and calibrated out. Combining the subspace fitting and least square approaches, the signal subspace of the received single short-term interval (STI) sample data of the LOFAR is used to build a cost function whose minimizer is a statistically efficient estimator of the unknown parameters—gains and phases of the telescopes. Subsequently, an iterative algorithm for finding the minimum of the cost function is presented and the unknown calibration parameters of both the core stations and the external subarray are separated. As a result, the computational complexity of the proposed method is significantly reduced compared to the existing methods based on a direct covariance fitting. Finally, the performance of the proposed method is compared with the conventional peeling method in computer simulation. An example for calibrating the core of the LOFAR array on Cyg A is also provided.

Index Terms— Array Self-Calibration, Subspace Fitting, LOFAR, Radio Astronomy Arrays

1. INTRODUCTION

The low frequency array (LOFAR) is a low frequency radio astronomical array that is currently used in The Netherlands and neighboring countries (1, 2, 3) (see www.lofar.org). It is a synthesis radio astronomical array designed to obtain some of the faintest signals in the universe, such as the epoch of re-ionization of the universe [4], which is only observable at low frequencies (30 - 240 MHz) due to high redshift levels. Imaging with LOFAR requires to solve many hard signal processing tasks such as calibration, imaging with direction dependent beams, interference mitigation and very high dynamic range imaging as well as significant advances in efficient computation because of the huge amounts of data [5], [6], [7], [8], [9], [10], [11]. The self calibration for a LOFAR like array is an important research problem due to the fact that current radio astronomy self calibration algorithms cannot work in the LOFAR environment with the requirement of direction dependent calibration[5] [6]. The capability of jointly estimating independent complex gain terms for each combination of telescope sub-array and calibrating source is required. An early candidate for a self calibration algorithm named the peeling method was introduced in [5]. To overcome this problem, the imaging problem can be described as a parameter estimation problem where the pixels in the image are parametrized using their location (direction), power and polarization parameters as described in [12],[13],[14],[15]. Especially, the parameter estimation problem for the LOFAR calibration was considered in [9], and a data model as well relevant CRLB analysis were computed using a new general formulation. Furthermore, it was shown that the ambiguous problem of unconstrained direction dependent calibration can be solved by using either the physical constraints of the LOFAR array or multiple time interval snapshots. A demix peeling calibration algorithm corresponding to an improvement of the conventional peeling are also given for the case of multiple snapshots in [9]. However, the computational complexity of the nonlinear multidimensional minimization is high and the performance of the resultant peeling method is also seriously degraded due to the cross-source interference at the initialization of the peeling method.

Calibration methods for direction independent errors have been widely dealt in the signal processing literature, see e.g., [16], [17], [18] [19], [20]. However no generalization of these techniques to arrays with direction dependent errors have been considered. This is still a computationally important problem and this paper takes a step in this direction. In this paper, the inherent direction independent property of the central core stations in the LOFAR array is efficiently exploited to further simplify the calibration problem of full arrays. A cost function with respect to the unknown gain and phase parameters is proposed by using the estimated signal subspace of the received telescopes array data. An iterative algorithm for solving the minimization problem of the cost function is presented to give the estimation of all unknown calibration parameters.

2. DATA MODEL AND SELF-CALIBRATION FOR LOFAR

2.1. Problem Formulation

Assume that $Q$ known bright calibrating point sources are observed by $J$ subarrays. For the $k$-th subband centered at frequency $f_k$, the $J \times 1$ observed array sample vector is expressed as

$$x_k(n) = \sum_{q=1}^{Q} n_{k,q}(n)s_{k,q}(n) + z_k(n) \quad (1)$$

where $s_{k,q}(n)$ is the signal from the $q$-th calibrating source at time sample $n$ and frequency $f_k$, $n_{k,q}(n)$ is the array response vector for this source. The vector $z_k(n)$ is the noise sample vector. The work described in this paper was supported by the Israel Science Foundation grant: “Signal processing and imaging techniques for large radio telescopes 1248/2009” and the NWO TOP project.
with spatially and temporally white Gaussian process, and is statistically independent of the Q signal sources.

Let \( N \) be the number of time samples in a short-term integration interval. Assume that \( \mathbf{x}_k(n) \) is constant over such an interval, as a result, for the \( m \) th interval, \( \mathbf{x}_k(n) \) is wide sense stationary over \( (m-1)N \leq n \leq mN - 1 \). A single autocovariance of the observed array vector \( \mathbf{x}_k(n) \) is defined as

\[
\mathbf{R}_{k,m} = E\{(\mathbf{x}_k(n)\mathbf{x}_k^H(n))\} - \mathbf{A}_{k,m} \mathbf{\Sigma}_k \mathbf{A}_{k,m}^H + \sigma^2 \mathbf{I} \tag{2}
\]

where

\[
\mathbf{A}_{k,m} = \begin{bmatrix} \mathbf{a}_{k,1}(\lfloor (m-1)N \rfloor) & \cdots & \mathbf{a}_{k,Q}(\lfloor (m-1)N \rfloor) \end{bmatrix} \tag{3}
\]

\[
\mathbf{\Sigma}_k = \text{diag}\{\sigma_{k,1}^2, \ldots, \sigma_{k,Q}^2\} \tag{4}
\]

Where \( \sigma_{k,j}^2 \) is the power of the \( j \)th calibration source at frequency \( k \) and \( \sigma^2 \) is the noise power. The array response matrix \( \mathbf{A}_{k,m} \) can be factored into the product of a phase matrix \( \mathbf{K}_{k,m} \) due entirely to the propagation delays associated with the array and source geometry, and a complex calibration gain matrix \( \mathbf{G}_{k,m} \), which includes both source direction dependent ionospheric perturbations and electronic instrumentation gain errors[9]

\[
\mathbf{A}_{k,m} = \mathbf{G}_{k,m} \otimes \mathbf{K}_{k,m}
\]

\[
= \begin{bmatrix} g_{k,m}^1 \otimes k_{k,m}^1 \quad \cdots \quad g_{k,m}^Q \otimes k_{k,m}^Q \end{bmatrix} \tag{5}
\]

where \( g_{k,m}^q \) and \( k_{k,m}^q \) are the \( g \)-th column of gain matrix \( \mathbf{G}_{k,m} \) and phase matrix \( \mathbf{K}_{k,m} \) respectively, and \( \otimes \) denotes the Hadamard product. In the astronomical literature, \( \mathbf{k}_{k,m}^q (q = 1, \ldots, Q) \) are jointly determined using the position vector for the \( q \)-th array element and the unit length vector pointing in the direction of the \( q \)-th source during STI snapshot \( m \). In general, the position vector, the direction of source and the source power levels are all known with high accuracy for tabulated calibrating sources, therefore, \( \mathbf{\Sigma}_k \) and \( \mathbf{K}_{k,m} \) are regarded as known quantities in the radio astronomy array[see 9]).

The problem of self calibration in LOFAR-like radio astronomy arrays is to estimate \( \mathbf{G}_{k,m} \) given \( \mathbf{R}_{k,m} \) over a range of \( k \) and \( m \), i.e., all the \( JQ \) independent unknown complex gain parameters in the full matrix \( \mathbf{G}_{k,m} \) must be estimated to calibrate the array for imaging or beamforming. A 2JQ \( \times 1 \) parameter vector containing all unknown terms is defined as

\[
\mathbf{\Theta}_{k,m} = [|g_{k,m}^1|; \ldots; |g_{k,m}^Q|; \angle(g_{k,m}^1); \ldots; \angle(g_{k,m}^Q)]\tag{6}
\]

Where \( | \cdot | \) denotes the modulus of the complex gain coefficient, and \( \angle(\cdot) \) denotes the phase of the complex gain coefficient, respectively.

For a single STI, the least squares calibration solution to the unknown parameters vector \( \mathbf{\Theta}_{k,m} \) in (2) is directly given from the following covariance fitting problem

\[
\mathbf{\Theta}_{k,m} = \arg \min_{\Theta} \| \mathbf{R}_{k,m} - ME_{k,m}(\Theta, \sigma^2) \|^2 \tag{7}
\]

where \( \| \cdot \|_F \) denotes the Frobenius matrix norm, and \( ME_{k,m}(\Theta, \sigma^2) \) is called the visibility measurement equation [5]

\[
ME_{k,m} = (\mathbf{g}_{k,m} \otimes \mathbf{k}_{k,m}) \mathbf{\Sigma}_k (\mathbf{g}_{k,m} \otimes \mathbf{k}_{k,m})^H + \sigma^2 \mathbf{I}
\]

Generally, directly solving the (7) is not computationally tractable, therefore, a computationally efficient method for estimating all the unknown calibration parameters \( \theta \) is highly desired. In addition, it is shown in [9] that the unknown calibration matrix \( \mathbf{G}_{k,m} \) is not identifiable using a single \( \mathbf{R}_{k,m} \) unless some additional constraints on the structure over \( k \) and \( m \), are added by exploiting the inherent physics of the source or across a range of time-frequency bins to solve the ambiguity in (7).

### 2.2. Calibration method using the Core Compact LOFAR Geometry

Due to the compact scene of the core subarray, the ionospheric phases are cancelled out when computing the correlations \( \mathbf{R}_{k,m} \) for the core subarray. Using the inherent physical property of the LOFAR geometry, the corresponding core gain matrix \( \mathbf{G}_c \) can be simplified as \( \mathbf{G}_c = \mathbf{g}_c \mathbf{I}^I \). For the rest \( \mathbf{J}_c = J - \mathbf{J}_c \) stations of the array, which are exterior to the core, both the field of view and all inter-element baselines are greater than the ionospheric irregularity scale. For these stations the corresponding gain matrix \( \mathbf{G}_c \) is direction dependent and best modeled as a full matrix (see [9]).

Under the above assumption, (5) can be written as

\[
\mathbf{A}_{k,m} = \mathbf{G}_c \otimes \mathbf{K}_c
\]

\[
= \begin{bmatrix} g_{c,1} \otimes k_{c,1} \quad \cdots \quad g_{c,Q} \otimes k_{c,Q} \end{bmatrix}
\]

where \( g_{c,q} \) and \( k_{c,q} \) are diagonal matrix containing the largest eigenvalues in decreasing order, and the associated eigenvectors are the columns of the matrix \( \mathbf{E}_c \), which spans the signal subspace. Based on the classical subspace property, it is easily derived to get the following equation [22]

\[
\mathbf{A} \Sigma_k^{1/2} = \mathbf{E}_c \Sigma^{1/2}_c \mathbf{P}
\]

where \( \mathbf{\Sigma}_c = \text{diag}\{\lambda_1, \ldots; \lambda_Q\} \) is a diagonal matrix containing the \( Q \) largest eigenvalues in decreasing order, and the associated eigenvectors are the columns of the matrix \( \mathbf{E}_c \), which spans the signal subspace. Based on the classical subspace property, it is easily derived to get the following equation [22]

\[
\mathbf{A} \Sigma_k^{1/2} = \mathbf{E}_c \Sigma^{1/2}_c \mathbf{P}
\]

where \( \mathbf{\Sigma}_c = \text{diag}\{\lambda_1, \ldots; \lambda_Q\} \) and \( \mathbf{P} \) is a unitary matrix. The noise power \( \sigma^2 \) is estimated by the mean of the \( J - Q \) smaller eigenvalues.

For a single STI the estimate of the covariance matrix \( \mathbf{R} \) in (9) is given by:

\[
\hat{\mathbf{R}}_{k,m} = \frac{1}{N} \sum_{m=0}^{m-1} \mathbf{x}_k(n)\mathbf{x}_k^H(n) \tag{11}
\]

The ML estimates \( \hat{\mathbf{E}}_c \) and \( \hat{\mathbf{\Sigma}}_c \) of \( \mathbf{E}_c \) and \( \mathbf{\Sigma}_c \) are given by the eigen-decomposition of \( \mathbf{R}_{k,m} \). Hence, a natural cost function whose minimum corresponds to estimates of the unknown gain parameters is defined as follows

\[
F(\Theta) = \| \mathbf{A}(\Theta) \hat{\mathbf{\Sigma}}_c^{1/2} - \hat{\mathbf{E}}_c \mathbf{\Sigma}_c^{1/2} \mathbf{P} \|_F^2
\]

\[
L = \hat{\mathbf{R}}_{k,m} \hat{\mathbf{\Sigma}}_c^{1/2} \tag{12}
\]

With the partition of \( \mathbf{A}(\Theta) \) in equation (8), one can write the above equation as

\[
F(\Theta) = \| \mathbf{A}(\Theta) \hat{\mathbf{\Sigma}}_c^{1/2} - \mathbf{L} \mathbf{P} \|_F^2
\]

\[
= \| \{\mathbf{g}_c \otimes \mathbf{k}_c\} \hat{\mathbf{\Sigma}}_c^{1/2} - \mathbf{L} \mathbf{P} \|_F^2
\]

\[
+ \| \{\mathbf{g}_c \otimes \mathbf{k}_c\} \hat{\mathbf{\Sigma}}_c^{1/2} - \mathbf{L} \mathbf{P} \|_F^2 \tag{13}
\]
where the matrix \( L_e \) consists of the first \( J_e \) rows of \( L \) and \( L_c \) consists of the last \( J_e \) rows of matrix \( L \). The minimization of (13) should be performed using the known structure of the matrices \( G_e(\Theta_e) \) and \( G_c(\Theta_c) \). In the following, a sequence of estimation of unknown parameters is given by minimizing the cost function iteratively and changing one unknown matrix at a time.

a. Minimization with respect to \( P \): Given the matrices \( A \) and \( L \), the orthonormal matrix \( P \) is given by minimizing the cost function \( F(\Theta) \) results in:

\[
\mathbf{P} = \mathbf{U} \mathbf{V}^H
\]  

(14)

where the columns of \( \mathbf{U} \) and \( \mathbf{V} \) are the left singular and right singular vectors of the singular value decomposition (SVD) of \( L^H A \Sigma_k^{1/2} \). This is a direct generalization of a similar result for real matrices proved in [22, 23].

b. Minimization with respect to \( G_c(\Theta_c) \): Given the matrices \( G_c(\Theta_c) \) and \( P \), the minimization of \( F(\Theta_c) \) with respect to \( \{g_1^c, \ldots, g_{J_e}^c\} \), is obtained by noting that

\[
F(\Theta_c) = \sum_{q=1}^{J_e} ||g_q^c - \bar{L}_q||_F^2 + ||K_c (G_c(\Theta_c) \ominus \bar{K}_c) \Sigma_k^{1/2} - \bar{L}_c P||_F^2
\]

(15)

where \( \bar{L}_q = k_q \ominus h \) and \( k_q = \text{diag}(\Sigma_k^{1/2}) \) and \( \bar{L}_c \) denotes the \( q \)-th row of \( L_c \). \( P \). Hence, using the diagonal property of \( G_c(\Theta_c) \), the minimization of \( F(\Theta_c) \) with respect to the \( \{g_1^c, \ldots, g_{J_e}^c\} \) can be performed by a separate fine search for each \( g_q^c \). However, noting that the second term of (15) does not depend on \( \Theta_e \), it may be dropped under minimization. Then, using a least squares formulation, it is easily shown that the estimates of \( g_q^c \), \( q = 1, \ldots, J_e \), can be obtained by (note that these vectors are row vectors):

\[
g_q^c = \frac{k_q^H a_q^H}{||k_q a_q||_F^2}, \quad q = 1, \ldots, J_e.
\]

(16)

c. Minimization with respect to \( G_e(\Theta_e) \): Given the matrices \( P \) and \( G_e(\Theta_e) \), the estimates of each column of the matrix \( G_e(\Theta_e) \) are given by minimizing the cost function \( F(\Theta_e) \)

\[
F(\Theta_e) = ||(G_e(\Theta_e) \ominus \bar{K}_e) \Sigma_k^{1/2} - \bar{L}_e P||_F^2 + \sum_{q=1}^{Q} ||(g_q^e \ominus k_q^e) \Sigma_k^{1/2} - d_q||_F^2
\]

(17)

where \( d_q \) is the \( q \)-th column of matrix \( L_e P \). Again, under minimization we drop the first term in (17) which does not depend on \( \Theta_e \) and we can obtain the estimates of each column \( g_q^e \) in \( G_e(\Theta_e) \) as follows:

\[
g_q^e = h_q^{-1/2} d_q \ominus (K_q^{\ominus -1}), \quad q = 1, \ldots, Q
\]

(18)

where \( h_q \) is the \( q \)-th element of \( h \), and \( (\cdot)^{\ominus -1} \) stands for the inverse operator of each element in \( (\cdot) \).

Finally, the proposed calibration method can be summarized as follows:

1. Initialization. Set the iteration counter to zero \( i = 0 \) and the initial value of \( P \) is set as \( P = I \).

2. Compute the \( G_c(\Theta_c) \) using (16) and \( G_e(\Theta_e) \) with (18).

3. Compute the cost function \( \varepsilon_i = F(\Theta^{(i)}) \) using (13).

4. Based on the estimated \( G_e(\Theta_e) \) and \( G_c(\Theta_c) \), compute the new value of \( P \) using (14).

5. Increment the iteration counter \( i = i + 1 \).

6. Using the updated \( G_e(\Theta_e) \), \( G_c(\Theta_c) \) and \( P \), compute the cost function value \( \varepsilon_{i+1} = F(\Theta^{(i+1)}) \).

7. Compute the difference \( \Delta \varepsilon = \varepsilon_i - \varepsilon_{i+1} \). If \( \Delta \varepsilon \leq \delta \) stop, otherwise go to step 2.

It is noteworthy that the higher dimension nonlinear searching computation in (7) is avoided by separating the unknown parameters in the proposed method, which the main computational complexity of the proposed method with \( I \) iterations is about \( o(J^3) + I(J^2Q + o(Q^3)) \), and the computational complexity is largely reduced compared with the peeling algorithm [5, 9].

3. SIMULATION RESULTS AND REAL MEASUREMENT DATA TEST

In this section we evaluated the performance of the proposed method and compared it to the peeling algorithm in computer simulation. To reduce the computational load of the peeling algorithm, a 32 element array with \( J_e = 8 \) core stations and only 26 radio sources were included in the simulation, and only the single short term integration interval (STI) with 500 time samples is used. The parameter set above, all 112(8*2 + 21*2 + 21*2 = 112) unknown gain and phase parameters will be estimated in the self calibration algorithm, while 64 unknown parameters for each source are required for the peeling algorithm [5]. The calibration parameters were randomly generated by Gaussian gain magnitudes with a mean of 1.0 and standard deviation of 0.3 as well phases uniformly distributed in the range \([-\pi, \pi]\). The direction parameter (direction cosine) of the two sky point sources is set as \((1, m, 0) = (0.2962, 0.1710), (I_2, m_2) = (0.1485, 0.5546)\), and the SNRs for two sources are [-10dB -15dB]. The additive noise is assumed to be white Gaussian process with variance 1, and all the estimated results for calibration parameters in the proposed method are averaged by 200 independent runs, and \( I = 5 \) passes in the peeling algorithm [5] are done. The mean square error (MSE) of the estimated gain and phase parameters of the proposed method is compared with that of the conventional peeling algorithm. The parameter index number is a function of the telescope station index and is ordered as defined in (6). Figs. 1-2 show the performance of the proposed method as compared to that of the peeling algorithm.

Finally, real LOFAR observations of Cyg A are used to show the performance of the proposed method. Due to technical issues involved in polarization calibration of the external stations we only consider the core stations and demonstrate the subspace approach for these stations. The LOFAR test station data were recorded using 21 frequency subbands of 61 kHz for the 13 core stations. Hence only the core gain matrix \( G_e \) being estimated in this case. A single target located at the center of the field view is observed, i.e. \((l_0, m_0, w_0) = (0, 0, 0)\). The core gain of 13 antennas is firstly estimated and then the received station data is calibrated by the estimated gain before imaging. The two images using the uncalibrated data and calibrated data based on the MVDR beamforming are given in Figure 3 and Figure 4. It is seen from Figs. 3-4 that the image after calibration is largely improved compared to the image before calibration.
4. CONCLUSION

The problem of self-calibration of the LOFAR astronomy array is discussed, and an iterative algorithm for estimating all unknown gain parameters is proposed by combining signal subspace fitting and a least square approach. Since the inherent physical property of the compact core LOFAR geometry is exploited, and the unknown parameters in both core subarray and external subarray are separated, the proposed method is computationally more efficient than that of the direct covariance fitting. Simulation results show that the performance of the proposed method is comparable to that of the available peeling algorithm but with lower computational complexity. Finally, a real data test of LOFAR is also included to demonstrate the significant improvement of the beamforming image after calibration. Combining these results with the approach in [24] is expected to yield very high dynamic range.

![The beamforming image before calibration.](image)

**Fig. 3.** The MVDR beamforming image before calibration.

![The beamforming image after calibration.](image)

**Fig. 4.** The MVDR beamforming image after calibration.

5. REFERENCES


