Calibration of Radio Interferometers Using a Sparse DoA Estimation Framework

M. BROSSARD[§], M. N. EL KORSO[*], M. PESAVENTO[†], R. BOYER[‡] and P. LARZABAL[‖]
[§‖]SATIE, UMR 8029, École Normale Supérieure de Cachan, Cachan, France
[*]Université Paris Ouest Nanterre La Défense, IUT de Ville d'Avray, LEME EA 4416, France
[§†]Communication Systems Group, Technische Universität, Darmstadt, Germany
[‡]L2S, UMR 8506, Université Paris Sud, Gif-sur-Yvette, France

***Abstract*—The calibration of modern radio interferometers is a significant challenge, specifically at low frequencies. In this perspective, we propose a novel iterative calibration algorithm, which employs the popular sparse representation framework, in the regime where the propagation conditions shift dissimilarly the directions of the sources. More precisely, our algorithm is designed to estimate the apparent directions of the calibration sources, their powers, the directional and undirectional complex gains of the array elements and their noise powers, with a reasonable computational complexity. Numerical simulations reveal that the proposed scheme is statistically efficient at low SNR and even with additional non-calibration sources at unknown directions.**

***Index Terms*—Calibration, radio astronomy, radio interferometer, sensor array, Direction-of-Arrival estimation**

I. INTRODUCTION

The calibration is a salient challenge for the new generation of radio interferometers [1], such as the LOw Frequency ARray (LOFAR) [2] or the Square Kilometer Array (SKA) [3]. These instruments consist in large sensor arrays for which calibration is essential to produce accurate images ($>10^8$ pixels) with high dynamic range ($60\,\mathrm{dB}$). Furthermore, the huge number of array elements imposes the need for designing computationally efficient algorithms.

For such radio interferometers, additional difficulties arise at low frequencies ($<300\,\mathrm{MHz}$), where the ionosphere causes phase delays which scale with the wavelength [4,5]. In this paper, we focus on the regime where all lines of sight toward a source in the sky cross the same ionospheric layer, where the thickness of the ionosphere can be direction dependent [6]. In this regime, the ionospheric phase delays are commonly modeled as a linear function of the distances between the so-called piercing points [7,8]. As a consequence, it modifies the geometric delays and introduces direction dependent angular-shifts for the source directions. By estimating the calibrator shifts (i.e., the difference between the true calibrator directions, known from tables [9], and their apparent directions, estimated from the observations of the radio interferometers), an interpolation method can be efficiently applied in order to obtain a phase screen model, that estimates the ionospheric delays over the entire Field-of-View [7]. In addition to the phase screen reconstruction step, calibration usually involves the estimation of the complex undirectional gains of the antennas, their directional gains toward each calibrator and their noise powers [10].

To solve this calibration problem, the a priori knowledge of some calibration sources is required, i.e., their true/nominal directions and powers without the effects of the ionosphere and antenna imperfections [8]. Based on this knowledge, state-of-the-art calibration algorithms to estimate the aforementioned parameters are mostly of iterative nature [1,8,10]. As an example, the (Weighted) Alternating Least Squares approach has been adapted for LOFAR calibration [10], in which closed form expressions have been obtained for the undirectional antenna gains, the source powers and the sensor noise powers. Nevertheless, regarding to the Direction-of-Arrival (DoA) estimation, no closed form expression can be obtained and classical subspace methods, such as MUSIC [11], have to be applied. A major drawback of these methods is that subspace techniques are not efficient in low Signal-to-Noise-Ratio (SNR) scenarios and require the exact number of sources in the scene.

As an alternative approach, recently, sparse reconstruction methods came into focus of DoA estimation for fully calibrated arrays [12] as well as for partially calibrated arrays [13]. They exhibit the super-resolution property, robustness and computational efficiency, without the aforementioned drawbacks of subspace-based methods [12]. Sparse representation methods have been successfully applied in radio astronomy for imaging [14], but, to the best of our knowledge, such methods have never been applied for this calibration problem.

In this paper, we focus on the calibration of a sensor array, involving its individual antennas and propagation disturbances. In addition, we assume that the sensor array has an arbitrary geometry, identical elements and is simultaneously excited by inaccurately known calibration sources and unknown non-calibration sources. We consider these non-calibration sources as outliers, i.e., as an additional noise term in the calibration step. From the calibration perspective, we propose a novel iterative scheme, which successively estimates the undirectional antenna gains along with the calibrator and noise parameters, to minimize a proper weighting cost function. The calibrator parameter estimation relies on the popular

This work was supported by the following projects: MAGELLAN (ANR-14-CE23-0004-01), MI-CNRS TITAN and ICode blanc.

978-0-9928-6265-7/16/$31.00 ©2016 IEEE

sparse representation framework and the sensor noise power estimation considers the presence of non-calibration sources (a.k.a. outliers in our calibration procedure), leading to a robust and computationally efficient algorithm in low SNR scenarios.

In the following, $\bar{(.)}, (.)^{\mathsf{T}}, (.)^{\mathsf{H}}, (.)^{\dagger}, (.)^{\odot\alpha}$ and $[.]_n$ denote, respectively, conjugation, transpose, Hermitian transpose, pseudo-inverse, element-wise raising to α and the n-th element of a vector. The expectation operator is $\mathcal{E}\{.\}, \circ$ denotes the Khatri-Rao product, $\exp(.)$ and $\odot$ represent the element-wise exponential function and multiplication (Hadamard product), respectively. The operator $\mathrm{diag}(.)$ converts a vector to a diagonal matrix with the vector aligned on the main diagonal, whereas $\mathrm{vecdiag}(.)$ produces a vector from the main diagonal of its entry and $\mathrm{vec}(.)$ converts a matrix to a vector by stacking the columns of its entry. The functions $\|.\|_0, \|.\|_2$ and $\|.\|_{\mathsf{F}}$ refer to l_0 norm, i.e., the number of non-zero elements of its entry, the l_2 and Frobenius norms, respectively. Finally, $\mathbf{x} \succeq 0$ means that each element in its vector $\mathbf{x}$ is non-negative.

II. Data model and problem statement

Let us consider an array of P elements, with known locations, each detonated by the Cartesian coordinates $\boldsymbol{\xi}_p = [x_p, y_p, z_p]^{\mathsf{T}}$ for $p = 1, \ldots, P$, that we stack in $\boldsymbol{\Xi} = [\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_P]^{\mathsf{T}} \in \mathbb{R}^{P\times 3}$. This array is exposed to Q known strong calibration sources and additional Q^{U} unknown weak non-calibration sources, with known true $\mathbf{D}^{\mathrm{K}} = \left[\mathbf{d}_1^{\mathrm{K}}, \ldots, \mathbf{d}_Q^{\mathrm{K}}\right] \in \mathbb{R}^{3\times Q}$ and unknown $\mathbf{D}^{\mathrm{U}} = \left[\mathbf{d}_1^{\mathrm{U}}, \ldots, \mathbf{d}_{Q^{\mathrm{U}}}^{\mathrm{U}}\right] \in \mathbb{R}^{3\times Q^{\mathrm{U}}}$ spatial coordinates, respectively, in which each direction $\mathbf{d} = [l, m, n]^{\mathsf{T}}$ can be uniquely described by a couple (l, m), since $n = \sqrt{1 - l^2 - m^2}$ [8]. The ionosphere introduces an unknown angular-shift for each source direction [1,7], and consequently, we distinguish between the unknown *apparent* directions for the calibrators, denoted by $\mathbf{D} = [\mathbf{d}_1, \ldots, \mathbf{d}_Q]$, and their *true* directions $\mathbf{D}^{\mathrm{K}}$, i.e., without the propagation disturbances.

Under the narrowband assumption, the steering vector $\mathbf{a}(\mathbf{d})$ toward the direction $\mathbf{d}$ is given by

$$\mathbf{a}(\mathbf{d}) = \mathbf{a}(l, m) = \frac{1}{\sqrt{P}} \exp\left(-\mathrm{j}\frac{2\pi}{\lambda}\boldsymbol{\Xi}\mathbf{d}\right) \in \mathbb{C}^P, \qquad (1)$$

where λ denotes the wavelength. For the calibration source signals, we consider the steering matrix

$$\mathbf{A} = \frac{1}{\sqrt{P}} \exp\left(-\mathrm{j}\frac{2\pi}{\lambda}\boldsymbol{\Xi}\mathbf{D}\right) \in \mathbb{C}^{P\times Q}, \qquad (2)$$

which contains the calibrator steering vectors. We define $\mathbf{A}^{\mathrm{U}}$, w.r.t. $\mathbf{D}^{\mathrm{U}}$, correspondingly for the non-calibration sources.

As in [10], we assume that all antennas have identical directional responses. Their directional gain responses (and propagation losses) can be modeled by two diagonal matrices $\boldsymbol{\Gamma} \in \mathbb{C}^{Q\times Q}$ and $\boldsymbol{\Gamma}^{\mathrm{U}} \in \mathbb{C}^{Q^{\mathrm{U}}\times Q^{\mathrm{U}}}$ toward the calibration and non-calibration sources, respectively.

The received signals from each antenna are stacked, for the n-th observation, into the vector

$$\mathbf{x}(n) = \mathbf{G}\left(\mathbf{A}\boldsymbol{\Gamma}\mathbf{s}(n) + \mathbf{A}^{\mathrm{U}}\boldsymbol{\Gamma}^{\mathrm{U}}\mathbf{s}^{\mathrm{U}}(n)\right) + \mathbf{n}(n), \qquad (3)$$

where $\mathbf{G} = \mathrm{diag}(\mathbf{g}) \in \mathbb{C}^{P\times P}$ models the undirectional antenna gains, $\mathbf{s}(n) \in \mathbb{C}^Q$ and $\mathbf{s}^{\mathrm{U}}(n) \in \mathbb{C}^{Q^{\mathrm{U}}}$ represent, respectively, the i.i.d. calibrator and non-calibrator signals and the vector $\mathbf{n}(n) \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma}^{\mathrm{n}})$ denotes the i.i.d. noise, in the n-th observation [8]. Consequently, the covariance matrix $\mathbf{R} = \mathcal{E}\left\{\mathbf{x}\mathbf{x}^{\mathsf{H}}\right\}$ of the observations is given by

$$\begin{aligned}\mathbf{R} = \mathbf{G}\mathbf{A}\boldsymbol{\Gamma}\boldsymbol{\Sigma}^{\mathrm{K}}\boldsymbol{\Gamma}^{\mathsf{H}}\mathbf{A}^{\mathsf{H}}\mathbf{G}^{\mathsf{H}} + \\ \mathbf{G}\mathbf{A}^{\mathrm{U}}\boldsymbol{\Gamma}^{\mathrm{U}}\boldsymbol{\Sigma}^{\mathrm{U}}\boldsymbol{\Gamma}^{\mathrm{U}\mathsf{H}}\mathbf{A}^{\mathrm{U}\mathsf{H}}\mathbf{G}^{\mathsf{H}} + \boldsymbol{\Sigma}^{\mathrm{n}},\end{aligned} \qquad (4)$$

where $\boldsymbol{\Sigma}^{\mathrm{K}} \in \mathbb{R}^{Q\times Q}, \boldsymbol{\Sigma}^{\mathrm{U}} \in \mathbb{R}^{Q^{\mathrm{U}}\times Q^{\mathrm{U}}}$ and $\boldsymbol{\Sigma}^{\mathrm{n}} = \mathrm{diag}\left(\boldsymbol{\sigma}^{\mathrm{n}}\right) \in \mathbb{R}^{P\times P}$ denote, respectively, the diagonal covariance matrix for the calibrators, non-calibration sources and sensor noises.

As $\boldsymbol{\Gamma}$ and $\boldsymbol{\Gamma}^{\mathrm{U}}$ are diagonal matrices, we define in the sequel the diagonal matrix that contain the apparent calibrator powers as $\boldsymbol{\Sigma} = \boldsymbol{\Gamma}\boldsymbol{\Sigma}^{\mathrm{K}}\boldsymbol{\Gamma}^{\mathsf{H}} = \mathrm{diag}\left(\boldsymbol{\sigma}\right) \in \mathbb{R}^{Q\times Q}$. Since only $\boldsymbol{\Sigma}^{\mathrm{K}}$ is assumed to be known, $\boldsymbol{\Sigma}$ is generally unknown and $\boldsymbol{\Gamma}$ can be deduced from it. We further introduce the unknown covariance matrix for the non-calibration sources, $\mathbf{R}^{\mathrm{U}} = \mathbf{G}\mathbf{A}^{\mathrm{U}}\boldsymbol{\Gamma}^{\mathrm{U}}\boldsymbol{\Sigma}^{\mathrm{U}}\boldsymbol{\Gamma}^{\mathrm{U}\mathsf{H}}\mathbf{A}^{\mathrm{U}\mathsf{H}}\mathbf{G}^{\mathsf{H}}$, and rewrite the covariance matrix model [1] as

$$\mathbf{R} = \mathbf{G}\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^{\mathsf{H}}\mathbf{G}^{\mathsf{H}} + \mathbf{R}^{\mathrm{U}} + \boldsymbol{\Sigma}^{\mathrm{n}}. \qquad (5)$$

We then formulate the calibration problem as the estimation of the parameter vector $\mathbf{p} = \left[\mathbf{g}^{\mathsf{T}}, \mathbf{d}_1^{\mathsf{T}}, \ldots, \mathbf{d}_Q^{\mathsf{T}}, \boldsymbol{\sigma}^{\mathsf{T}}, \boldsymbol{\sigma}^{\mathrm{n}\mathsf{T}}\right]^{\mathsf{T}}$, from the sample covariance matrix $\hat{\mathbf{R}} = \frac{1}{N}\sum_{n=1}^{N}\mathbf{x}(n)\mathbf{x}^{\mathsf{H}}(n)$. Note that the estimation of the unknown matrix $\mathbf{R}^{\mathrm{U}}$ represents the imaging step [10] which is beyond the scope of the paper. The imaging step is done usually as a separate step after the calibration [14]. The main reason is that the calibration step is usually based on a source point model (unlike the imaging step) with a known number of calibrators, whereas, the effect of the weakest (non-calibration) sources, with an unknown source number, can be assumed absorbed by the noise component.

In order to overcome the scaling ambiguities in model (5), we consider the following commonly used assumptions in radio astronomy [8,10]: i) to solve the phase ambiguity of $\mathbf{g}$, we take its first element as the phase reference; ii) $\mathbf{g}$ and $\boldsymbol{\sigma}$ share a common scalar factor and when solving for the

Algorithm 1: Iterative and Sparsity Based Calibration Algorithm (ISBCA)

Input: sample covariance matrix $\hat{\mathbf{R}}$;
Init: set the iteration counter $i = 0$,
$\mathbf{g} = \mathbf{g}^{[0]}, \mathbf{D} = \mathbf{D}^{\mathrm{K}}, \boldsymbol{\sigma} = \mathrm{diag}\left(\boldsymbol{\Sigma}^{\mathrm{K}}\right), \boldsymbol{\Omega} = \mathbf{1}_{P\times P}$;
while $\left\|\mathbf{p}^{[i-1]} - \mathbf{p}^{[i]}\right\|_2 \geq \left\|\mathbf{p}^{[i]}\right\|_2 \epsilon_{\mathbf{p}}$ **do**
1 $i = i + 1$;
2 Estimate $\mathbf{g}^{[i]}$ with Algorithm 1.2;
3 Estimate $\mathbf{D}^{[i]}, \boldsymbol{\sigma}^{[i]}$ and $\boldsymbol{\sigma}^{\mathrm{n}[i]}$ with Algorithm 1.3;
4 Update $\boldsymbol{\Omega} = \left(\boldsymbol{\sigma}^{\mathrm{n}[i]}\boldsymbol{\sigma}^{\mathrm{n}[i]\mathsf{T}}\right)^{\odot-\frac{1}{2}}$;
Output: $\hat{\mathbf{p}} = \left[\mathbf{g}^{[i]\mathsf{T}}, \mathbf{d}_1^{[i]\mathsf{T}}, \ldots, \mathbf{d}_Q^{[i]\mathsf{T}}, \boldsymbol{\sigma}^{[i]\mathsf{T}}, \boldsymbol{\sigma}^{\mathrm{n}[i]\mathsf{T}}\right]^{\mathsf{T}}$;

calibrator directions, a single rotation of all steering vectors can be compensated by the undirectional gain phase solution. We therefore consider an additional known source to remove both ambiguities, by fixing its direction $\mathbf{d}_0$ and its apparant power σ_0.

III. Calibration scheme

A. Overview of the Proposed Algorithm

A statistically efficient estimator of the model parameters can be obtained via the Maximum-Likehood formulation, but it appears intractable in practice. However, with a large number of samples, statistically efficient estimations can be reached using the Weighting Least Squares approach. Consequently, we define the following cost function to minimize: $\kappa(\mathbf{p}) = \left\| \mathbf{W}^{-\frac{1}{2}} \left(\mathbf{R}(\mathbf{p}) - \hat{\mathbf{R}} \right) \mathbf{W}^{-\frac{1}{2}} \right\|_{\mathsf{F}}^2$, with $\mathbf{R}(\mathbf{p}) = \mathbf{G}\mathbf{A}\mathbf{\Sigma}\mathbf{A}^{\mathsf{H}}\mathbf{G}^{\mathsf{H}} + \mathbf{\Sigma}^{\mathrm{n}}$ denoting the covariance matrix in the absence of the non-calibration sources, and $\mathbf{W}$ is the weighting matrix. The optimal weighting matrix for Gaussian noise is $\mathbf{W} = \mathbf{R}$ [10], which is generally unknown. In radio astronomy, sources are typically much weaker than the noise [15], so the covariance matrix can be approximated by $\mathbf{R} \approx \mathbf{\Sigma}^{\mathrm{n}}$. Since the array consists of identical antennas and mutual coupling is negligible, it is commonly assumed that $\mathbf{\Sigma}^{\mathrm{n}} = \mathrm{diag}\left(\boldsymbol{\sigma}^{\mathrm{n}}\right) \approx \sigma^{\mathrm{n}}\mathbf{I}$. Consequently, we consider $\mathbf{W} = \mathbf{I}$ as an initial step and refine it with $\mathbf{W} = \mathbf{\Sigma}^{\mathrm{n}}$ once we obtain an estimate of $\mathbf{\Sigma}^{\mathrm{n}}$. Since $\mathbf{\Sigma}^{\mathrm{n}}$ is diagonal, we rewrite the cost function as

$$\kappa(\mathbf{p}) = \left\| \left(\mathbf{R}(\mathbf{p}) - \hat{\mathbf{R}} \right) \odot \mathbf{\Omega} \right\|_{\mathsf{F}}^2, \tag{6}$$

with $\mathbf{\Omega} = \left(\boldsymbol{\sigma}^{\mathrm{n}} \boldsymbol{\sigma}^{\mathrm{n}\mathsf{T}} \right)^{\odot -\frac{1}{2}}$.

We aim at minimizing the cost function $\kappa(\mathbf{p})$ in an iterative manner. We first minimize $\kappa(\mathbf{p})$ w.r.t. $\mathbf{g}$, with the remaining parameters in $\mathbf{p}$ fixed as described in the Subsection III-B. In a second step, we minimize (6) w.r.t. the variables $\boldsymbol{\sigma}, \mathbf{d}_1, \ldots, \mathbf{d}_Q$ and $\boldsymbol{\sigma}^{\mathrm{n}}$ for fixed $\mathbf{g}$, by using a sparse representation approach as described in the Subsection III-C. The overall procedure, referred to as the Iterative and Sparsity Based Calibration Algorithm (ISBCA), is presented in Algorithm 1. The algorithm is initialized with the true/nominal calibrator parameters and an initial guess for the undirectional gains or the unit sensor gain. In the following, we detail the two major alternating optimization steps of the ISBCA.

B. Undirectional Antenna Gain Estimation (Algorithm 1.2)

In this subsection, we describe Algorithm 1.2 of the ISBCA. We follow the same iterative approach as discussed in [15], that we adapt for our cost function (6) that we optimize w.r.t. $\mathbf{g}$ for the remaining parameter in $\mathbf{p}$ fixed.

Toward this aim, we consider $\mathbf{g}$ and $\bar{\mathbf{g}}$ as two independent variables. We first regard $\bar{\mathbf{g}}$ as fixed and minimize (6) w.r.t. $\mathbf{g}$, only, and without considering the diagonal elements in the cost function (6) that contain the unknown noise variances $\boldsymbol{\sigma}^{\mathrm{n}}$. In this case, the cost function becomes separable w.r.t. the elements of $\mathbf{g}$, hence,

$$\kappa(\mathbf{g}) = \sum_{p=1}^{P} \kappa^p([\mathbf{g}]_p), \tag{7}$$

where $\kappa^p([\mathbf{g}]_p)$ corresponds to the cost function for the p-th row of $\mathbf{R}$, which depends only on $[\mathbf{g}]_p$ since the remaining parameters are considered as fixed in this step. Let us define the operator $\mathcal{S}_p(.)$, that converts to a vector the p-th row of a matrix and removes the p-th element of this selected vector. Further, define the vector $\hat{\mathbf{r}}_p = \mathcal{S}_p\left(\hat{\mathbf{R}}\right)$ and the weighting vector $\boldsymbol{\omega} = \mathcal{S}_p\left(\mathbf{\Omega}\right)$. We can thus write $\kappa^p([\mathbf{g}]_p)$ in (7) as

$$\kappa^p([\mathbf{g}]_p) = \left\| (\hat{\mathbf{r}}_p - \mathbf{z}[\mathbf{g}]_p) \odot \boldsymbol{\omega} \right\|_2^2, \tag{8}$$

in which $\mathbf{z} = \mathcal{S}_p\left(\mathbf{R}^{\kappa}\bar{\mathbf{G}}\right)$ and where $\mathbf{R}^{\kappa} = \mathbf{A}\mathbf{\Sigma}\mathbf{A}^{\mathsf{H}}$ represents the calibrator sky model. Since $\kappa^p([\mathbf{g}]_p)$ is a Least Squares function of $[\mathbf{g}]_p$, by using standard inversion techniques and introducing $\mathbf{z}_{\mathrm{w}} = \mathbf{z} \odot \mathcal{S}_p\left(\mathbf{\Omega}\right)$, its minimizer is given by $[\hat{\mathbf{g}}]_p = \frac{\mathbf{z}_{\mathrm{w}}^{\mathsf{H}}\hat{\mathbf{r}}_p}{\mathbf{z}_{\mathrm{w}}^{\mathsf{H}}\mathbf{z}}$. After this, we directly update $[\bar{\mathbf{g}}]_p$ and proceed in the same manner with the remaining parameters in $\mathbf{g}$. This procedure, summarized in Algorithm 1.2, is repeated until convergence.

C. Calibrator Parameter and Noise Power Estimation (Algorithm 1.3)

In this subsection, we describe Algorithm 1.3 of the ISBCA for optimizing (6) w.r.t. the calibrator parameters and noise powers, that is based mainly on the popular sparse representation framework.

Assuming that the calibration sources are well separated, which is common in radio astronomy [8,10], we consider in the sequel that: i) each apparent calibration source lies in an sector of displacements around its nominal location; ii) the displacement sectors of different calibration sources are not overlapping; iii) each following dictionary shall represent the displacement set corresponding to its source.

Algorithm 1.2: undirectional antenna gain estimation

Input: sample covariance matrix $\hat{\mathbf{R}}$;
Init: set the iteration counter $k = 0$,
$\mathbf{g} = \mathbf{g}^{[i-1]}, \mathbf{R}^{\kappa} = \mathbf{A}^{[i-1]}\mathbf{\Sigma}^{[i-1]}\mathbf{A}^{[i-1]\mathsf{H}}$;
while $\left\|\mathbf{g}^{[k-1]} - \mathbf{g}^{[k]}\right\|_2 \geq \left\|\mathbf{g}^{[k]}\right\|_2 \epsilon_{\mathbf{g}}$ **do**
1 $k = k + 1$;
for $p = 1, \ldots, P$ **do**
2 $\hat{\mathbf{r}}_p = \mathcal{S}_p\left(\hat{\mathbf{R}}\right)$;
3 $\mathbf{z} = \mathcal{S}_p\left(\mathbf{R}^{\kappa}\bar{\mathbf{G}}\right)$;
4 $\mathbf{z}_{\mathrm{w}} = \mathbf{z} \odot \mathcal{S}_p\left(\mathbf{\Omega}\right)$;
5 $[\mathbf{g}]_p^{[k]} = \frac{\mathbf{z}_{\mathrm{w}}^{\mathsf{H}}\hat{\mathbf{r}}_p}{\mathbf{z}_{\mathrm{w}}^{\mathsf{H}}\mathbf{z}}$;
Output: $\hat{\mathbf{g}} = \mathbf{g}^{[k]}$;

Let us define Q dictionaries of steering vectors $\tilde{\mathbf{A}}_q$, for $q = 1, \ldots, Q$, as

$$\tilde{\mathbf{A}}_q = \Big[\mathbf{a}(l_{q,1}, m_{q,1}), \ldots, \mathbf{a}(l_{q,1}, m_{q,N_q^m}), \mathbf{a}(l_{q,2}, m_{q,1}), \ldots, \mathbf{a}(l_{q,N_q^l}, m_{q,N_q^m})\Big] \in \mathbb{C}^{P \times N_q^l N_q^m}, \tag{9}$$

which contain $N_q^l N_q^m$ steering vectors, centered around the true/nominal direction of the q-th calibrator, namely $\mathbf{d}_q^{\text{K}}$, with resolution $(\Delta l_q, \Delta m_q)$ and $N_q^l \gg 1, N_q^m \gg 1$. These Q dictionary steering matrices are gathered in

$$\tilde{\mathbf{A}} = \Big[\tilde{\mathbf{A}}_1, \ldots, \tilde{\mathbf{A}}_Q\Big] \in \mathbb{C}^{P \times N_g}, \tag{10}$$

with $N_g = \sum_{q=1}^{Q} N_q^l N_q^m$ denoting the total number of directions on the grid. We define the sparse calibrator power vector as

$$\tilde{\boldsymbol{\sigma}} = \left[\tilde{\boldsymbol{\sigma}}_1^{\mathsf{T}}, \ldots, \tilde{\boldsymbol{\sigma}}_Q^{\mathsf{T}}\right]^{\mathsf{T}} \in \mathbb{R}^{N_g}, \tag{11}$$

which contains the powers of all calibrators. Due to the previous assumption of non-overlapping displacement sectors, we postulate that $\tilde{\boldsymbol{\sigma}}_q$ is exactly 1-sparse, i.e., $\|\tilde{\boldsymbol{\sigma}}_q\|_0 = 1$.

Algorithm 1.3: calibrator parameter and noise power estimation

Input: sample covariance matrix $\hat{\mathbf{R}}$;
Init: set the iteration counter $k = 0$,
$\mathbf{g} = \mathbf{g}^{[i-1]}, \boldsymbol{\sigma} = \boldsymbol{\sigma}^{[i-1]}, \mathbf{D} = \mathbf{D}^{[i-1]}$;
while $\left\|\tilde{\boldsymbol{\sigma}}^{[k-1]} - \tilde{\boldsymbol{\sigma}}^{[k]}\right\|_2 \geq \left\|\tilde{\boldsymbol{\sigma}}^{[k]}\right\|_2 \epsilon_{\tilde{\boldsymbol{\sigma}}}$ **do**
1 $k = k + 1$;
for $q = 1, \ldots, Q$ **do**
2 Calculate the residual vector for the q-th source: $\check{\tilde{\mathbf{r}}}_q = \check{\tilde{\mathbf{r}}} - \sum_{q'=1}^{q-1} \check{\mathbf{C}}_{q'} \tilde{\boldsymbol{\sigma}}_{q'}^{[k]} - \sum_{q'=q+1}^{Q} \check{\mathbf{C}}_{q'} \tilde{\boldsymbol{\sigma}}_{q'}^{[k-1]}$;
3 Calculate the stepsize μ_q^k as in [16];
4 $\tilde{\boldsymbol{\sigma}}_q^{[k]} = \mathcal{H}_1\left(\tilde{\boldsymbol{\sigma}}_q^{[k-1]} + \mu_q^k \check{\mathbf{C}}_q^{\dagger}\left(\check{\tilde{\mathbf{r}}}_q - \check{\mathbf{C}}_q \tilde{\boldsymbol{\sigma}}_q^{[k-1]}\right)\right)$;
5 Estimate $\boldsymbol{\sigma}^{\text{n}}$ with (20);
Output: $\hat{\tilde{\boldsymbol{\sigma}}} = \tilde{\boldsymbol{\sigma}}^{[k]}$, that leads to $\hat{\boldsymbol{\sigma}}, \hat{\mathbf{d}}_1, \ldots, \hat{\mathbf{d}}_Q$ and $\hat{\boldsymbol{\sigma}}^{\text{n}}$;

Using (5), the covariance model can be rewritten as

$$\mathbf{R} = \mathbf{G}\tilde{\mathbf{A}}\tilde{\boldsymbol{\Sigma}}\tilde{\mathbf{A}}^{\mathsf{H}}\mathbf{G}^{\mathsf{H}} + \mathbf{R}^{\text{U}} + \boldsymbol{\Sigma}^{\text{n}}, \tag{12}$$

in which $\tilde{\boldsymbol{\Sigma}} = \text{diag}(\tilde{\boldsymbol{\sigma}})$. Let us then define

$$\mathbf{C}_q = \left(\boldsymbol{\Sigma}^{\text{n}-\frac{1}{2}}\bar{\mathbf{G}}\bar{\tilde{\mathbf{A}}}_q\right) \circ \left(\boldsymbol{\Sigma}^{\text{n}-\frac{1}{2}}\mathbf{G}\tilde{\mathbf{A}}_q\right), q = 1, \ldots, Q \tag{13}$$

$$\mathbf{C} = [\mathbf{C}_1, \ldots, \mathbf{C}_Q] = \left(\boldsymbol{\Sigma}^{\text{n}-\frac{1}{2}}\bar{\mathbf{G}}\bar{\tilde{\mathbf{A}}}\right) \circ \left(\boldsymbol{\Sigma}^{\text{n}-\frac{1}{2}}\mathbf{G}\tilde{\mathbf{A}}\right) \tag{14}$$

$$\mathbf{N} = \boldsymbol{\Sigma}^{\text{n}-\frac{1}{2}} \circ \boldsymbol{\Sigma}^{\text{n}-\frac{1}{2}}, \hat{\mathbf{r}} = \text{vec}(\hat{\mathbf{R}} \odot \boldsymbol{\Omega}). \tag{15}$$

Thus, we formulate the minimization problem of (6) as

$$\hat{\tilde{\boldsymbol{\sigma}}}, \hat{\boldsymbol{\sigma}}^{\text{n}} = \underset{\tilde{\boldsymbol{\sigma}}, \boldsymbol{\sigma}^{\text{n}}}{\arg\min} \|\hat{\mathbf{r}} - \mathbf{C}\tilde{\boldsymbol{\sigma}} - \mathbf{N}\boldsymbol{\sigma}^{\text{n}}\|_2^2 \quad \text{subject to } \tilde{\boldsymbol{\sigma}} \succeq 0, \boldsymbol{\sigma}^{\text{n}} \succeq 0, \|\tilde{\boldsymbol{\sigma}}_q\|_0 = 1, q = 1, \ldots, Q, \tag{16}$$

in which we make use of the fact that the Q displacement sectors are not overlapping.

To consider the l_0 constraints in (16), which are generally difficult, we choose the Iterative Hard Thresholding scheme of [17]. This greedy algorithm is based on a projected gradient descend direction algorithm and offers strong theoretical guarantees that have already been employed in the DoA estimation context [18]. Particularly, when the grid is fine and the columns of $\tilde{\mathbf{A}}_q$ are strongly similar, we can guarantee that each $\tilde{\boldsymbol{\sigma}}_q$ obtained from (16) is exactly 1-sparse. Thus, using the Coordinate Descent algorithm [19] to minimize (16), we obtain an analytic solution for each sub-problem and the sparsity of the desired minimizer $\tilde{\boldsymbol{\sigma}}$ reduces the computational complexity [19]. Each step involves the hard thresholding operator $\mathcal{H}_s(.)$, that keeps the s-largest components of a vector and sets the remaining entries equal to zero, thus, it automatically satisfies both constraints of sparsity and positivity. We allow a step size factor μ_q^k that depends on $\tilde{\boldsymbol{\sigma}}_q$ and the k-th iteration, hence considering the Normalized Iterative Hard Thresholding procedure of [16], where the choice of μ_q^k assures convergence toward a local minimum.

Since the p-th element of $\boldsymbol{\sigma}^{\text{n}}$, $[\boldsymbol{\sigma}^{\text{n}}]_p$, is only present in the p-th diagonal term of $\mathbf{R}$, ignoring this term does not affect the estimation of $\tilde{\boldsymbol{\sigma}}$. Thus, $\boldsymbol{\sigma}^{\text{n}}$ is estimated after the estimation of $\tilde{\boldsymbol{\sigma}}$, for which $\boldsymbol{\sigma}^{\text{n}} \succeq 0$ holds in low SNR scenarios.

Let us denote $\check{\tilde{\mathbf{r}}}$ and $\check{\mathbf{C}}$, that refer, respectively, to $\hat{\mathbf{r}}$ and $\mathbf{C}$ without their elements corresponding to the diagonal of $\mathbf{R}$. The solution of (16) w.r.t. $\tilde{\boldsymbol{\sigma}}$ can then be obtained as

$$\hat{\tilde{\boldsymbol{\sigma}}} = \underset{\tilde{\boldsymbol{\sigma}}}{\arg\min} \left\|\check{\tilde{\mathbf{r}}} - \check{\mathbf{C}}\tilde{\boldsymbol{\sigma}}\right\|_2^2 \quad \text{subject to } \tilde{\boldsymbol{\sigma}} \succeq 0, \|\tilde{\boldsymbol{\sigma}}_q\|_0 = 1, q = 1, \ldots, Q, \tag{17}$$

which is used in Algorithm 1.3.

Afterward, the minimizer of (16) w.r.t. $\boldsymbol{\sigma}^{\text{n}}$ is given by

$$\check{\boldsymbol{\sigma}}^{\text{n}} = \text{vecdiag}\left(\hat{\mathbf{R}} - \hat{\mathbf{G}}\hat{\mathbf{R}}^{\text{K}}\hat{\mathbf{G}}^{\mathsf{H}}\right). \tag{18}$$

Let us remove the bias introduced by the non-calibration sources as follows: we calculate the power

$$\sigma_r = \frac{\mathbf{a}(\mathbf{d}_r)^{\mathsf{H}}\left(\hat{\mathbf{R}} - \hat{\mathbf{G}}\hat{\mathbf{R}}^{\text{K}}\hat{\mathbf{G}}^{\mathsf{H}}\right)\mathbf{a}(\mathbf{d}_r)}{\|\mathbf{a}(\mathbf{d}_r)\|_2^2} \tag{19}$$

of the residual sample covariance matrix for a random direction $\mathbf{d}_r$, where no source is supposed to be present. We then approximate $\mathbf{a}(\mathbf{d}_r)^{\mathsf{H}}\mathbf{a}(\mathbf{d}_q) \approx 0$ for any $\mathbf{d}_r \neq \mathbf{d}_q$, which yields σ_r as the sum of the sensor noise powers. By imposing $\sum_{p=1}^{P}[\boldsymbol{\sigma}^{\text{n}}]_p = \sigma_r$, the unbiased solution is given by

$$\hat{\boldsymbol{\sigma}}^{\text{n}} = \check{\boldsymbol{\sigma}}^{\text{n}} + \frac{1}{P}\left(\sigma_r - \mathbf{1}_{P\times 1}^{\mathsf{T}}\check{\boldsymbol{\sigma}}^{\text{n}}\right)\mathbf{1}_{P\times 1}, \tag{20}$$

that concludes Algorithm 1.3.

IV. Simulations

The proposed method is tested using Monte-Carlo simulations and compared with the deterministic Cramér–Rao bound (CRB), that expresses a lower bound on the variance of any unbiased estimator. We test the algorithm in standard situations, with similar sensor locations, sky and parameters as in [1,8,10].

The antenna locations correspond to the LOFAR's Initial Test Station [20], with $P = 60$ antennas disposed in a five-armed spiral. We assume a sky model at 30 MHz ($\lambda = 10\,\mathrm{m}$) consisting of $Q = 2$ strong calibration sources and $Q^{\mathrm{U}} = 8$ weak non-calibration sources, provided from the ten strongest sources in the table of [9]. The total power of these sources is assumed to be 1% of the total antenna noise power, a typical scenario for radio interferometers [10]. Data are generated via the signal model given in (3), in order to obtain the sample covariance matrix. We choose initially a coarse grid, with the same resolution for each coordinate and each calibrator. To avoid off-grid mismatch, we apply grid refinements [12] until we achieve the theoretical limits given by the CRB.

To investigate the algorithm performances, we perform 500 Monte-Carlo runs for each sample size. The variances of the errors on the complex undirectional gains, the calibrator directions and the powers, are plotted in Fig. 1 and Fig. 2, as a function of the number of samples N and compared to the corresponding CRB. As expected, the method appears to be asymptotically statistically efficient in a low SNR scenario, even with the presence of non-calibration sources. The (Relative) Mean Square Error may be slightly lower than the CRB for some values of N as shown in Fig. 1 and Fig. 2. This is mainly due to the fact that our algorithm takes into account the true/nominal direction of the calibrators (see (9)), whereas the classical CRB does not include this prior knowledge.

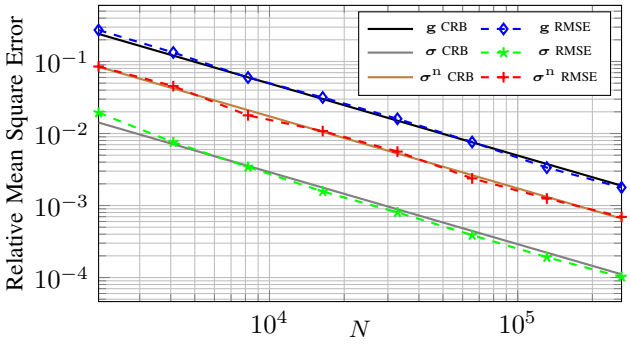

Fig. 1. Relative Mean Square Error on the undirectional antenna gains, the powers of the calibrators and the antenna noise powers.

V. Conclusion

In this paper, we presented a novel iterative scheme for the calibration of radio interferometers, where different shifts affect the apparent directions and powers of the calibration sources. The proposed algorithm, named Iterative and Sparsity Based Calibration Algorithm (ISBCA), iteratively minimizes a weighting Least Squares function in order to estimate the complex undirectional antenna gains and their noise powers, whereas, it jointly estimates the directions and powers of the calibrators using the Normalized Iterative Hard Thresholding procedure. This leads to a statistically efficient, computationally efficient and robust scheme as shown by numerical simulations.

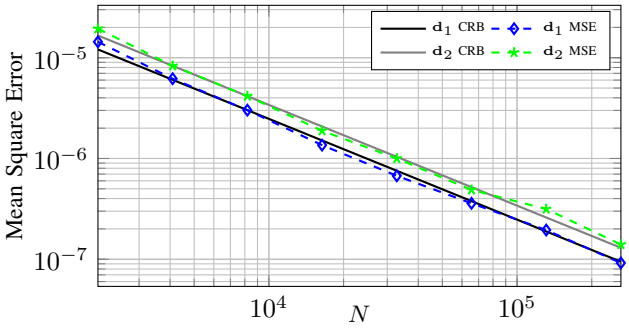

Fig. 2. Mean Square Error on the directions of the calibrators.

References

[1] S. Wijnholds, S. van der Tol, R. Nijboer, and A.-J. van der Veen, "Calibration challenges for future radio telescopes," *IEEE Signal Processing Magazine*, vol. 27, pp. 30–42, 2010.

[2] M. Van Haarlem, M. Wise, A. Gunst, et. al., "LOFAR: The LOw-Frequency ARray," *Astronomy and Astrophysics*, vol. 556, 2013.

[3] P. Dewdney, P. Hall, R. Schilizzi, and T. Lazio, "The Square Kilometre Array," *Proceedings of the IEEE*, vol. 97, no. 8, pp. 1482–1496, 2009.

[4] S. van der Tol, *Bayesian Estimation for Ionospheric Calibration in Radio Astronomy*. PhD thesis, Delft University of Technology, 2009.

[5] H. Intema, S. van der Tol, W. Cotton, et. al., "Ionospheric calibration of low frequency radio interferometric observations using the peeling scheme. I. Method description and first results," *Astronomy and Astrophysics*, vol. 501, pp. 1185–1205, 2009.

[6] C. Lonsdale, "Calibration approaches." 2004.

[7] W. Cotton, J. Condon, R. Perley, et. al., "Beyond the isoplanatic patch in the VLA Low-frequency Sky Survey," in *Ground-based Telescopes*, vol. 5489 of *Proc. SPIE*, pp. 180–189, 2004.

[8] A.-J. van der Veen and S. Wijnholds, "Signal processing tools for radio astronomy.," in *Handbook of Signal Processing Systems*, pp. 421–463, Springer, 2013.

[9] A. Bennett, "The revised 3C catalog of radio sources," *Memoirs of the Royal Astronomical Society*, vol. 68, pp. 163–172, 1962.

[10] S. Wijnholds, *Fish-eye Observing with Phased Array Radio Telescopes*. PhD thesis, Delft University of Technology, 2010.

[11] R. Schmidt, "Multiple emitter location and signal parameter estimation," *Antennas and Propagation, IEEE Transactions on*, vol. 34, no. 3, pp. 276–280, 1986.

[12] D. Malioutov, M. Cetin, and A. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *Signal Processing, IEEE Transactions on*, vol. 53, no. 8, pp. 3010–3022, 2005.

[13] C. Steffens, P. Parvazi, and M. Pesavento, "Direction finding and array calibration based on sparse reconstruction in partly calibrated arrays," in *Proceedings of the 8th IEEE Sensor Array and Multichannel Signal Processing Workshop*, pp. 21–24, 2014.

[14] Y. Wiaux, L. Jacques, G. Puy, et. al., "Compressed sensing imaging techniques for radio interferometry," *Monthly Notices of the Royal Astronomical Society*, vol. 395, pp. 1733–1742, 2009.

[15] S. Salvini and S. Wijnholds, "Fast gain calibration in radio astronomy using alternating direction implicit methods: Analysis and applications," *Astronomy and Astrophysics*, vol. 571, p. A97, 2014.

[16] T. Blumensath and M. Davies, "Normalized iterative hard thresholding: Guaranteed stability and performance," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 4, no. 2, pp. 298–309, 2010.

[17] T. Blumensath and M. Davies, "Iterative hard thresholding for compressed sensing," *Applied and Computational Harmonic Analysis*, vol. 27, no. 3, pp. 265–274, 2009.

[18] E. Ollila, K. Hyon-Jung, and V. Koivunen, "Robust iterative hard thresholding for compressed sensing," in *International Symposium on Communications, Control and Signal Processing*, pp. 226–229, 2014.

[19] J. Friedman, T. Hastie, H. Höfling, and R. Tibshirani, "Pathwise coordinate optimization," *Ann. Appl. Stat.*, vol. 1, no. 2, pp. 302–332, 2007.

[20] S. Wijnholds, J. Bregman, and A. Boonstra, "Sky noise limited snapshot imaging in the presence of RFI with LOFAR's Initial Test Station," *Experimental Astronomy*, vol. 17, no. 1, pp. 35–42, 2004.