



Robust Calibration of Radio Interferometers in Non-Gaussian Environment

Virginie Ollier, Mohammed Nabil El Korso, Rémy Boyer, Pascal Larzabal,
Marius Pesavento

► To cite this version:

Virginie Ollier, Mohammed Nabil El Korso, Rémy Boyer, Pascal Larzabal, Marius Pesavento. Robust Calibration of Radio Interferometers in Non-Gaussian Environment. IEEE Transactions on Signal Processing, Institute of Electrical and Electronics Engineers, 2017, 65 (21), pp.5649-5660. 10.1109/TSP.2017.2733496 . hal-01572142

HAL Id: hal-01572142

<https://hal-centralesupelec.archives-ouvertes.fr/hal-01572142>

Submitted on 4 Aug 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Robust Calibration of Radio Interferometers in Non-Gaussian Environment

Virginie Ollier, Mohammed Nabil El Korso, Rémy Boyer, *Senior Member, IEEE*, Pascal Larzabal, *Member, IEEE* and Marius Pesavento, *Member, IEEE*

Abstract—The development of new phased array systems in radio astronomy, as the low frequency array (LOFAR) and the square kilometre array (SKA), formed of a large number of small and flexible elementary antennas, has led to significant challenges. Among them, model calibration is a crucial step in order to provide accurate and thus meaningful images and requires the estimation of all the perturbation effects introduced along the signal propagation path, for a specific source direction and antenna position. Usually, it is common to perform model calibration using the a priori knowledge regarding a small number of known strong calibrator sources but under the assumption of Gaussianity of the noise. Nevertheless, observations in the context of radio astronomy are known to be affected by the presence of outliers which are due to several causes, e.g., weak non-calibrator sources or man made radio frequency interferences. Consequently, the classical Gaussian noise assumption is violated leading to severe degradation in performances. In order to take into account the outlier effects, we assume that the noise follows a spherically invariant random distribution. Based on this modeling, a robust calibration algorithm is presented in this paper. More precisely, this new scheme is based on the design of an iterative relaxed concentrated maximum likelihood estimation procedure which allows to obtain closed-form expressions for the unknown parameters with a reasonable computational cost. This is of importance as the number of estimated parameters depends on the number of antenna elements, which is large for the new generation of radio interferometers. Numerical simulations reveal that the proposed algorithm outperforms the state-of-the-art calibration techniques.

Index Terms—Calibration, robustness, spherically invariant random process, relaxed concentrated maximum likelihood, Jones matrices, radio astronomy

I. INTRODUCTION

Radio astronomy aims to study radio emissions from the sky, in order to detect, identify new objects and observe known structures at higher resolution, in a specific electromagnetic spectrum [1]. This fundamental thematic shines a new light on our universe, revealing more about its nature and history. In order to carry out particularly sensitive observations in a large

range of the spectrum, and to handle significant cosmological issues, largely distributed sensor arrays are currently being built or planned, such as the low frequency array (LOFAR) [2] and the square kilometre array (SKA) [3]. They will notably be composed of a large number of relatively low-cost small antennas with wide field of view, resulting in a large collecting area and high resolution imaging. Nevertheless, to meet the theoretical optimal performances of such next generation radio interferometers, a plethora of signal processing challenges must be overcome, among them, calibration, data reduction and image synthesis [4]–[7]. These aspects are intertwined and must be dealt with to take advantage of the new advanced radio interferometers. As an example, lack of calibration has dramatic effects in the image reconstruction by causing severe distortions. In this paper, we focus on calibration, which involves the estimation of all unknown perturbation effects and represents a cornerstone of the imaging step [8]–[10].

Array calibration aspects have been tackled for a few decades in the array processing community leading to a variety of calibration algorithms [11]–[13]. Such algorithms can be classified into two different approaches depending on the presence [14]–[16], or the absence [17]–[22], of one or more cooperative sources, named calibrator sources. In the radio astronomy context, calibration is commonly treated using the first approach as we have access to prior knowledge thanks to tables describing accurately the position and flux of the brightest sources [23].

Following this methodology, the majority of proposed calibration schemes in radio interferometry are least squares-based approaches. The state-of-the-art consists in the so-called alternating least squares approach [24]–[27], which leads to statistically efficient algorithm under a Gaussian model, since the least squares estimator is equivalent to the maximum likelihood (ML) estimator in this case. On the other hand, expectation maximization (EM) [28]–[30] and EM-based algorithms, such as the space alternating generalized expectation maximization algorithm [31], have been proposed in order to enhance the convergence rate of the least squares-based calibration algorithms [32]. Nevertheless, the major drawback of these schemes is the Gaussianity assumption which is not realistic in the radio astronomy context. Specifically, the presence of outliers has multiple causes, among which i) the radio frequency interferers, which corrupt the observations and are not always perfectly filtered in practice [33,34], ii) the presence of unknown weak sources in the background [35], iii) the presence of some punctual events such as interference due to the Sun or due to strong sources in the sidelobes which

Virginie Ollier and Pascal Larzabal are with SATIE, UMR 8029, École Normale Supérieure de Cachan, Université Paris-Saclay, Cachan, France (e-mail: virginie.ollier@satie.ens-cachan.fr, pascal.larzabal@satie.ens-cachan.fr). Virginie Ollier and Rémy Boyer are with L2S, UMR 8506, Université Paris-Sud, Université Paris-Saclay, Gif-sur-Yvette, France (e-mail: remy.boyer@l2s.centralesupelec.fr). Mohammed Nabil El Korso is with LEME, EA 4416, Université Paris-Ouest, Ville d’Avray, France (e-mail: m.elkorso@parisnanterre.fr). Marius Pesavento is with Communication Systems Group, Technische Universität Darmstadt, Darmstadt, Germany (e-mail: mpesa@nt.tu-darmstadt.de).

This work was supported by MAGELLAN (ANR-14-CE23-0004-01), ON FIRE project (Jeunes Chercheurs GDR-ISIS) and ANR ASTRID project MARGARITA.

can also randomly create outliers [36]. To the best of our knowledge, the proposed scheme in [35], represents the only alternative to the existing calibration algorithms based on a Gaussian noise model.

In [35], theoretical and experimental analyses have been conducted in order to demonstrate that the effect of outliers in the radio astronomy context can indeed be modeled by a non-Gaussian heavy-tailed distributed noise process. Nevertheless, the algorithm presented in [35] has its own limits, since the noise is specifically modeled as a Student's t with independent identically distributed entries. To improve the robustness of the calibration, we propose, in this paper, a new scheme based on a broader class of distributions gathered under the so-called spherically invariant random noise modeling [37,38], which includes the Student's t distribution. A spherically invariant random vector (SIRV) is described as the product of a positive random variable, named texture, and the so-called speckle component which is Gaussian, resulting in a two-scale compound Gaussian distribution [39]. The flexibility of the SIRV modeling allows to consider non-Gaussian heavy-tailed distributed noise in the presence of outliers, but also to adaptively consider Gaussian noise in the extreme case when there are no outliers. Under the SIRV model, we estimate the unknown parameters iteratively based on a relaxed ML estimator, leading to closed-form expressions for the noise parameters while a block coordinate descent (BCD) algorithm [40,41] is designed to obtain the estimates of parameters of interest efficiently and at a low cost.

Finally, it is worth mentioning that the parametric model used in this paper to describe the perturbation effects is based on the so-called Jones matrices [42,43]. Such formalism describes in a flexible way the conversion of the incident electric field into voltages. Indeed, along its propagation path, the signal is affected by various effects and transformations which correspond to matrix multiplications in the mathematical Jones framework. Multiple distortion effects caused by the environment and/or the instruments can be easily incorporated into the model using an adequate parametrization of the Jones matrices. Such effects can represent, for example, the ionospheric phase delay resulting in angular shifts, the atmospheric distortions, the typical phase delay due to geometric pathlength difference, the voltage primary beam, the cross-leakage or also the electronic gains [44,45]. For the above reasons and due to its flexibility [1,32,42]–[44], we adopt this parametric model. We make a distinction between the non-structured and the structured cases: in the first one, one total Jones matrix stands for all the effects along the full signal path while in the second case, we regard each physical effect separately thanks to individual Jones terms in a cumulative product. Thus, different corruptions are described by different kinds of Jones matrices. We emphasize that the proposed algorithm, entitled relaxed concentrated ML estimator, is a generic algorithm as it is based on a non-structured Jones matrices formulation as a first step. However, it can be adapted to various regimes describing distinct calibration scenarios in which an array can operate [46]. In this paper, we consider the specific example of the direction dependent distortion regime with a compact set of antennas, which we refer to as the 3DC regime. The

array is therefore considered as a closely packed group of antennas but the array elements have a wide field of view. This is particularly well-adapted for calibration of compact arrays, typically a LOFAR station.

The rest of the paper is organized as follows: in Section II, we present the data model in the context of radio astronomy, first with non-structured Jones matrices and thereafter, we study an example of structured Jones matrices for the 3DC calibration regime. In Section III, we give an overview of the proposed robust ML estimator, based on spherically invariant random process (SIRP) noise modeling. An efficient estimation procedure of the distortions introduced on each signal propagation path is derived in Section IV. Then, the algorithm is adapted to the case of structured Jones matrices in Section V for the 3DC calibration regime. Finally, we provide numerical simulations in Section VI to assess the robustness of the approach and draw our conclusions in Section VII.

In this paper, we use the following notation: symbols $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^H$ denote, respectively, the transpose, the complex conjugate and the Hermitian transpose. The Kronecker product is represented by \otimes , $E\{\cdot\}$ denotes the expectation operator, $\text{bdiag}\{\cdot\}$ is the block-diagonal operator, whereas $\text{diag}\{\cdot\}$ converts a vector into a diagonal matrix. The trace and determinant operators are, respectively, referred by $\text{tr}\{\cdot\}$ and $|\cdot|$. The symbol \mathbf{I}_B represents the $B \times B$ identity matrix, $\text{vec}(\cdot)$ stacks the columns of a matrix on top of one another, $\|\cdot\|_F$ is the Frobenius norm, while $\|\cdot\|_2$ denotes the l_2 norm. Finally, $\Re\{\cdot\}$ represents the real part and we note j the complex number whose square equals -1 .

II. DATA MODEL

A. Case of non-structured Jones matrices

Let us consider M antennas with known locations that receive D signals emitted by calibrator sources. Each antenna is dual polarized and composed of two receptors, in order to provide sensitivity to the two orthogonal polarization directions (x, y) of the incident electromagnetic plane wave. Consequently, the relation between the i -th source emission and the measured voltage at the p -th antenna is given by [42,43,47]

$$\mathbf{v}_{i,p}(\boldsymbol{\theta}) = \mathbf{J}_{i,p}(\boldsymbol{\theta})\mathbf{s}_i \quad (1)$$

where $\mathbf{s}_i = [s_{i_x}, s_{i_y}]^T$ is the incoming signal, $\mathbf{v}_{i,p}(\boldsymbol{\theta}) = [v_{i,p_x}(\boldsymbol{\theta}), v_{i,p_y}(\boldsymbol{\theta})]^T$ is the generated voltage with one output for each polarization direction and $\mathbf{J}_{i,p}(\boldsymbol{\theta})$ denotes the so-called 2×2 Jones matrix, parametrized by the unknown vector of interest $\boldsymbol{\theta}$. The Jones matrix models the array response and all the perturbations introduced along the path from the i -th source to the p -th sensor. Since each propagation path is particular, we can associate a different Jones matrix with each source-antenna pair (i, p) , leading to a total number of DM Jones matrices. In this section, we consider the non-structured case where no specific perturbation model is used to describe the physical mechanism behind each perturbation effect and the unknown elements correspond to the entries of all Jones matrices [32,48] (a structured example is given for 3DC calibration regime, in Section II-B).

For each antenna pair, we compute the correlation of the output signals, resulting in the typical observations recorded by a radio interferometer. The correlation between voltages is given, in the case of noise free measurements, for the (p, q) antenna pair, by

$$\begin{aligned} \mathbf{V}_{pq}(\boldsymbol{\theta}) &= \mathbb{E} \left\{ \left(\sum_{i=1}^D \mathbf{v}_{i,p}(\boldsymbol{\theta}) \right) \left(\sum_{i=1}^D \mathbf{v}_{i,q}^H(\boldsymbol{\theta}) \right) \right\} \\ &= \sum_{i=1}^D \mathbf{J}_{i,p}(\boldsymbol{\theta}) \mathbf{C}_i \mathbf{J}_{i,q}^H(\boldsymbol{\theta}) \quad \text{for } p < q, \quad p, q \in \{1, \dots, M\}, \end{aligned} \quad (2)$$

where the signals emitted by the sources are assumed uncorrelated and the 2×2 matrix $\mathbf{C}_i = \mathbb{E}\{\mathbf{s}_i \mathbf{s}_i^H\}$ is known from prior knowledge. Let us remark that autocorrelations are not considered as shown by the condition $p < q$ in (2) (this is a typical situation in the radio astronomy context where radio interferometric systems automatically flag the autocorrelations [2]).

Using the property $\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\text{vec}(\mathbf{B})$, we rewrite (2) as a 4×1 vector

$$\tilde{\mathbf{v}}_{pq}(\boldsymbol{\theta}) = \text{vec}(\mathbf{V}_{pq}(\boldsymbol{\theta})) = \sum_{i=1}^D \mathbf{u}_{i,pq}(\boldsymbol{\theta}) \quad (3)$$

in which $\mathbf{u}_{i,pq}(\boldsymbol{\theta}) = (\mathbf{J}_{i,q}^*(\boldsymbol{\theta}) \otimes \mathbf{J}_{i,p}(\boldsymbol{\theta})) \mathbf{c}_i$ with $\mathbf{c}_i = \text{vec}(\mathbf{C}_i)$. We stack all the noisy measurements within a full vector $\mathbf{x} = [\mathbf{v}_{12}^T, \mathbf{v}_{13}^T, \dots, \mathbf{v}_{(M-1)M}^T]^T \in \mathbb{C}^{4B \times 1}$, where $B = \frac{M(M-1)}{2}$ denotes the total number of antenna pairs and $\mathbf{v}_{pq} = \tilde{\mathbf{v}}_{pq}(\boldsymbol{\theta}) + \mathbf{n}_{pq}$ with \mathbf{n}_{pq} the noise sample at a specific antenna pair. Specifically, \mathbf{x} reads

$$\mathbf{x} = \sum_{i=1}^D \mathbf{u}_i(\boldsymbol{\theta}) + \mathbf{n} \quad (4)$$

in which $\mathbf{u}_i(\boldsymbol{\theta}) = [\mathbf{u}_{i,12}^T(\boldsymbol{\theta}), \mathbf{u}_{i,13}^T(\boldsymbol{\theta}), \dots, \mathbf{u}_{i,(M-1)M}^T(\boldsymbol{\theta})]^T$ and $\mathbf{n} = [\mathbf{n}_{12}^T, \mathbf{n}_{13}^T, \dots, \mathbf{n}_{(M-1)M}^T]^T$ is the full noise vector which accounts for Gaussian noise, but also the presence of outliers in our data. Therefore, the noise can no longer be considered Gaussian and a robust calibration method is required. To investigate non-Gaussian noise modeling and encompass a broad range of noise distributions, we propose to adopt the SIRP noise model [37,38]. Specifically, the noise at each antenna pair is assumed to be generated as

$$\mathbf{n}_{pq} = \sqrt{\tau_{pq}} \mathbf{g}_{pq}, \quad (5)$$

where the positive real random variable τ_{pq} is referred to as texture, whereas the complex speckle component \mathbf{g}_{pq} follows a zero-mean Gaussian distribution¹, i.e.,

$$\mathbf{g}_{pq} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Omega}). \quad (6)$$

In order to remove scaling ambiguities, we impose $\text{tr}\{\boldsymbol{\Omega}\} = 1$. Note that the choice of this constraint is arbitrary and does not affect the estimates of interest as argued in [49].

¹Let us note that it is possible to consider a different covariance matrix $\boldsymbol{\Omega}_{pq}$ for each speckle component \mathbf{g}_{pq} in (6). In this case, the proposed algorithm requires a few modifications and the corresponding expressions are presented in Appendix A.

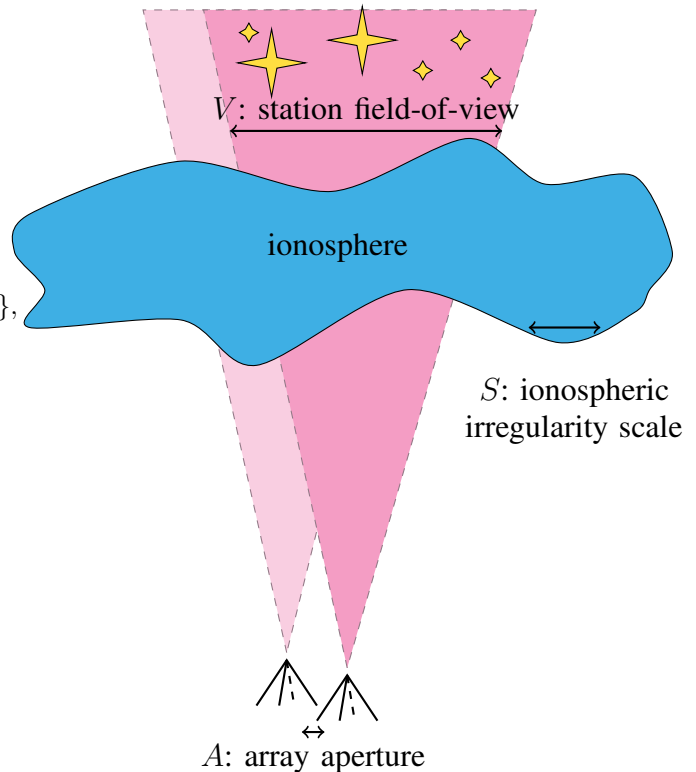


Fig. 1. 3DC calibration regime, for which $V \gg S$ and $S \gg A$. All receiving elements in the station see the same ionosphere part but, due to their wide field of view, a multitude of sources are visible and perturbations are highly direction dependent.

In this section, we adopted the non-structured Jones matrices formulation which is relevant in the radio astronomical context [32,48]. In this case, there is no need to specify the full propagation path, avoiding misspecification in the model. Besides, it is highly flexible and can be adapted to different scenarios [46]. In the following, we present the direction dependent distortion regime with a compact set of antennas, named 3DC regime.

B. Specific case of the 3DC calibration regime

For a specific propagation path, from the i -th source to the p -th antenna, the global Jones matrix $\mathbf{J}_{i,p}$ accounts for multiple effects which can be described explicitly. Indeed, each global matrix can be decomposed into individual Jones terms which stand for specific physical effects [43,44]. This way, instead of estimating entries of all Jones matrices as done in the non-structured case, we will estimate physically meaningful parameters, thus reducing the total number of parameters to estimate. Introducing structured Jones matrices can be done in the context of calibration scenarios [10,46]. In what follows, we target one specific commonly used calibration regime that we call 3DC calibration regime, described in Fig. 1, which is well adapted for calibration of LOFAR on station level [26], and also for stations of the future SKA radio interferometer [10]. In this scenario, direction dependent distortions play a significant role since individual elements in the array have

a wide field of view. Indeed, this implies different propagation conditions towards distinct sources in the field of view. However, the array being relatively compact, made of similar elements, some effects might be the same for all antennas. In the following, we introduce a particular sequence of Jones matrices with specific parametrizations, in the context of 3DC calibration regime [50]

$$\mathbf{J}_{i,p}(\boldsymbol{\theta}_{i,p}^{3\text{DC}}) = \mathbf{G}_p(\mathbf{g}_p)\mathbf{H}_{i,p}\mathbf{Z}_{i,p}(\boldsymbol{\alpha}_i)\mathbf{F}_i(\vartheta_i) \quad (7)$$

for $i \in \{1, \dots, D\}$, $p \in \{1, \dots, M\}$ and $\boldsymbol{\theta}_{i,p}^{3\text{DC}} = [\vartheta_i, \mathbf{g}_p^T, \boldsymbol{\alpha}_i^T]^T$. We note $\mathbf{H}_{i,p}$ the only assumed known matrix thanks to electromagnetic simulations in terms of antenna response and a priori knowledge given by calibrator source and antenna positions [43,44,50,51], whereas the remaining matrices are explained in the following items.

• *Ionospheric effects* :

Propagation through the ionosphere, the outer layer of the earth's atmosphere, introduces propagation delay on the signal which is affected by spatially variable fluctuations. If these perturbation effects are not corrected for, the sources may appear shifted from their intrinsic positions [10,52]. In the case of a compact array, the ionospheric delay matrix is in fact a scalar direction-dependent phase given by

$$\mathbf{Z}_{i,p}(\boldsymbol{\alpha}_i) = \exp\left\{j\varphi_{i,p}\right\}\mathbf{I}_2 \quad (8)$$

in which $\varphi_{i,p} = \eta_i u_p + \zeta_i v_p$ where $\boldsymbol{\alpha}_i = [\eta_i, \zeta_i]^T$ is the vector of unknown offsets resulting in a shift of the i -th source direction and $\mathbf{r}_p = [u_p, v_p]^T$ is the vector of known antenna position in units of wavelength.

On top of that, passing through the ionosphere is associated with a rotation of the polarisation plane of each signal source around the line of sight. We call it the ionospheric Faraday rotation matrix $\mathbf{F}_i(\vartheta_i)$ and write it as

$$\mathbf{F}_i(\vartheta_i) = \begin{bmatrix} \cos(\vartheta_i) & -\sin(\vartheta_i) \\ \sin(\vartheta_i) & \cos(\vartheta_i) \end{bmatrix} \quad (9)$$

where ϑ_i is the unknown Faraday rotation angle, assumed identical for all antennas, since the array has a limited spatial extent [44].

• *Instrumental effects* :

Individual antennas are described by electronic complex gains which appear in $\mathbf{G}_p(\mathbf{g}_p) = \text{diag}\{\mathbf{g}_p\}$ with \mathbf{g}_p the unknown electronic gain vector.

Therefore, in this specific structured case, the physical model parameters in (7) are collected in the vector $\boldsymbol{\epsilon}^{3\text{DC}} = \mathbf{P}[\boldsymbol{\theta}_{1,1}^{3\text{DC}T}, \boldsymbol{\theta}_{1,2}^{3\text{DC}T}, \dots, \boldsymbol{\theta}_{D,M}^{3\text{DC}T}]^T$ where \mathbf{P} is an appropriate rearrangement matrix such that $\boldsymbol{\epsilon}^{3\text{DC}} = [\vartheta_1, \dots, \vartheta_D, \mathbf{g}_1^T, \dots, \mathbf{g}_M^T, \boldsymbol{\alpha}_1^T, \dots, \boldsymbol{\alpha}_D^T]^T$.

III. ROBUST CALIBRATION ESTIMATOR

This section is devoted to the design of a robust calibration estimator based on the model (4). As it can be seen from (5), one has to specify the probability density function (pdf) of each texture parameter τ_{pq} in order to obtain the exact ML estimates. Nevertheless, in practical scenarios, such prior knowledge is not available. Consequently, our idea is to make use of a relaxed version of the exact model, i.e.,

we assume deterministic but unknown texture realizations in the estimation process [53,54]. This ensures more flexibility in our algorithm as the texture distribution is not precisely described and avoids any possible model misspecification, which is consistent with our motivation to design a broad robust estimator w.r.t. the presence of outliers. On the other hand, we adopt here an iterative procedure in order to reduce the computational cost. In doing so, the proposed algorithm sequentially updates each block of unknown parameters while fixing the remaining parameters. This leads to a relaxed concentrated ML based calibration estimator for which the expression of the likelihood function, when independency is assumed between measurements, is written as

$$f(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\Omega}) = \prod_{pq} \frac{1}{|\pi\tau_{pq}\boldsymbol{\Omega}|} \exp\left\{-\frac{1}{\tau_{pq}}\mathbf{a}_{pq}^H(\boldsymbol{\theta})\boldsymbol{\Omega}^{-1}\mathbf{a}_{pq}(\boldsymbol{\theta})\right\}, \quad (10)$$

where the vector composed of all texture realizations is $\boldsymbol{\tau} = [\tau_{12}, \tau_{13}, \dots, \tau_{(M-1)M}]^T$ and $\mathbf{a}_{pq}(\boldsymbol{\theta}) = \mathbf{v}_{pq} - \tilde{\mathbf{v}}_{pq}(\boldsymbol{\theta})$. Consequently, the log-likelihood function reads

$$\begin{aligned} \log f(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\Omega}) &= -4B \log \pi \\ &- 4 \sum_{pq} \log \tau_{pq} - B \log |\boldsymbol{\Omega}| - \sum_{pq} \frac{1}{\tau_{pq}} \mathbf{a}_{pq}^H(\boldsymbol{\theta}) \boldsymbol{\Omega}^{-1} \mathbf{a}_{pq}(\boldsymbol{\theta}). \end{aligned} \quad (11)$$

In the following, we present the sequential updates of each block of unknown parameters, namely, $\boldsymbol{\tau}$, $\boldsymbol{\Omega}$ and $\boldsymbol{\theta}$, following the methodology as in [55,56].

1) Derivation of $\hat{\tau}_{pq}$: Taking the derivative of the log-likelihood function in (11) w.r.t. τ_{pq} leads to the following texture estimate

$$\hat{\tau}_{pq} = \frac{1}{4} \mathbf{a}_{pq}^H(\boldsymbol{\theta}) \boldsymbol{\Omega}^{-1} \mathbf{a}_{pq}(\boldsymbol{\theta}). \quad (12)$$

2) Derivation of $\hat{\boldsymbol{\Omega}}$: The derivative of the log-likelihood function w.r.t. the element $[\boldsymbol{\Omega}]_{k,l}$ of the speckle covariance matrix, using classical differential properties [57, p. 2741], leads to

$$-B \text{tr} \left\{ \hat{\boldsymbol{\Omega}}^{-1} \mathbf{e}_k \mathbf{e}_l^T \right\} + \sum_{pq} \frac{1}{\tau_{pq}} \mathbf{a}_{pq}^H(\boldsymbol{\theta}) \hat{\boldsymbol{\Omega}}^{-1} \mathbf{e}_k \mathbf{e}_l^T \hat{\boldsymbol{\Omega}}^{-1} \mathbf{a}_{pq}(\boldsymbol{\theta}) = 0 \quad (13)$$

where the vector \mathbf{e}_k contains zeros except at the k -th position which is equal to unity. The permutation property of the trace operator enables to rewrite (13) as

$$-B \mathbf{e}_l^T \hat{\boldsymbol{\Omega}}^{-1} \mathbf{e}_k + \sum_{pq} \frac{1}{\tau_{pq}} \mathbf{e}_l^T \hat{\boldsymbol{\Omega}}^{-1} \mathbf{a}_{pq}(\boldsymbol{\theta}) \mathbf{a}_{pq}^H(\boldsymbol{\theta}) \hat{\boldsymbol{\Omega}}^{-1} \mathbf{e}_k = 0. \quad (14)$$

This finally leads to the following estimate of the speckle covariance matrix

$$\hat{\boldsymbol{\Omega}} = \frac{1}{B} \sum_{pq} \frac{1}{\tau_{pq}} \mathbf{a}_{pq}(\boldsymbol{\theta}) \mathbf{a}_{pq}^H(\boldsymbol{\theta}). \quad (15)$$

Inserting (12) into (15), we obtain

$$\hat{\boldsymbol{\Omega}}^{h+1} = \frac{4}{B} \sum_{pq} \frac{\mathbf{a}_{pq}(\boldsymbol{\theta}) \mathbf{a}_{pq}^H(\boldsymbol{\theta})}{\mathbf{a}_{pq}^H(\boldsymbol{\theta}) (\hat{\boldsymbol{\Omega}}^h)^{-1} \mathbf{a}_{pq}(\boldsymbol{\theta})} \quad (16)$$

where h denotes the h -th iteration. Due to the introduced constraint, we normalize the estimate of Ω by its trace, as follows

$$\hat{\Omega}^{h+1} = \frac{\hat{\Omega}^{h+1}}{\text{tr}\{\hat{\Omega}^{h+1}\}}. \quad (17)$$

3) Estimation of $\hat{\theta}$: For given Ω and τ , estimating $\hat{\theta}$ is equivalent to the following minimization problem

$$\hat{\theta} = \underset{\theta}{\text{argmin}} \left\{ \sum_{pq} \frac{1}{\tau_{pq}} \mathbf{a}_{pq}^H(\theta) \Omega^{-1} \mathbf{a}_{pq}(\theta) \right\}. \quad (18)$$

In the following, we aim to reduce the computational cost of the minimization procedure in (18) by use of the EM algorithm. For generality, we first adopt the non-structured Jones matrix formulation, which can also be specified depending on the scenario, as shown in Section V.

IV. ESTIMATION OF $\hat{\theta}$ FOR NON-STRUCTURED JONES MATRICES

Estimating directly the entries of the Jones matrices avoids specifying any particular physical model, leading to a calibration algorithm which is less sensitive to model errors in comparison with algorithms based on the structured case. Due to the possible large size of θ , a multi-dimensional parameter search needs to be carried out to solve the optimization problem in (18) which requires significant computation time. To reduce this complexity, we make use of the EM algorithm.

The essence of this algorithm relies on a proper parameter vector partitioning as well as an adequate choice of the so-called complete data. As mentioned above, the parameters of interest θ represent the entries of all Jones matrices. Consequently, it is natural to consider the following partition

$$\theta = [\theta_1^T, \dots, \theta_D^T]^T = [\theta_{1,1}^T, \dots, \theta_{1,M}^T, \dots, \theta_{D,1}^T, \dots, \theta_{D,M}^T]^T, \quad (19)$$

for which the vector $\theta_{i,p} \in \mathbb{R}^{8 \times 1}$ is the parametrization of the path from the i -th calibrator source to the p -th sensor, i.e., $\mathbf{J}_{i,p}(\theta) = \mathbf{J}_{i,p}(\theta_{i,p})$.

A. Use of the EM algorithm to solve (18)

The EM algorithm [28]–[30] enables to compute the ML estimates and reduce the computational cost, via the iteration of two steps. The first one is the E-step which reduces, in our scenario, to the computation of the conditional expectation of the complete data given the observed data and the current estimate of parameters [32,58]. Afterwards, the log-likelihood function of this conditional distribution is maximized in the M-step. Therefore, this last step consists in an optimization process which can be performed numerically, for instance with the Levenberg-Marquardt (LM) algorithm [59]–[61], or analytically if closed-form expressions are available. As we show in the following, the optimization step is carried out w.r.t. to $\theta_i \in \mathbb{C}^{4M \times 1}$ instead of $\theta \in \mathbb{C}^{4DM \times 1}$. Therefore, the global multiple source estimation problem is reduced to multiple single source sub-problems.

1) E-step: For the i -th source, we introduce the so-called complete data vector \mathbf{w}_i such that

$$\mathbf{x} = \sum_{i=1}^D \mathbf{w}_i \quad (20)$$

with $\mathbf{w}_i = \mathbf{u}_i(\theta_i) + \mathbf{n}_i$ and $\mathbf{n} = \sum_{i=1}^D \mathbf{n}_i$, in which $\mathbf{n}_i \sim \mathcal{CN}(\mathbf{0}, \beta_i \Psi)$. We have $\sum_{i=1}^D \beta_i = 1$ and Ψ is the covariance matrix of \mathbf{n} . Since $\mathbf{n}_{pq} \sim \mathcal{CN}(\mathbf{0}, \tau_{pq} \Omega)$ and with the independence property, we obtain the following block-diagonal expression for Ψ

$$\Psi = \text{bdiag}\{\tau_{12} \Omega, \dots, \tau_{(M-1)M} \Omega\}. \quad (21)$$

Let us note $\mathbf{w} = [\mathbf{w}_1^T, \dots, \mathbf{w}_D^T]^T$ the complete data vector, whose covariance matrix, denoted as Ξ , has the following form

$$\Xi = \text{bdiag}\{\beta_1 \Psi, \dots, \beta_D \Psi\}. \quad (22)$$

With [62, p. 36], and after some calculus, the expression of the conditional expectation is given by

$$\begin{aligned} \hat{\mathbf{w}}_i &= \text{E}\{\mathbf{w}_i | \mathbf{x}; \theta, \tau, \Omega\} \\ &= \mathbf{u}_i(\theta_i) + \beta_i \left(\mathbf{x} - \sum_{l=1}^D \mathbf{u}_l(\theta_l) \right). \end{aligned} \quad (23)$$

2) M-step: The goal of this step is to estimate θ_i . Once $\hat{\mathbf{w}}_i$ are computed for $i \in \{1, \dots, D\}$, the estimated complete data vector $\hat{\mathbf{w}}$ can be evaluated. The M-step is an optimization problem based on the following likelihood function where \mathbf{w}_i are independent

$$\begin{aligned} f(\hat{\mathbf{w}} | \theta, \tau, \Omega) &= \\ &= \frac{1}{|\pi \Xi|} \exp \left\{ -(\hat{\mathbf{w}} - \mathbf{u}(\theta))^H \Xi^{-1} (\hat{\mathbf{w}} - \mathbf{u}(\theta)) \right\} = \\ &= \prod_{i=1}^D \frac{1}{|\pi \beta_i \Psi|} \exp \left\{ -(\hat{\mathbf{w}}_i - \mathbf{u}_i(\theta_i))^H (\beta_i \Psi)^{-1} (\hat{\mathbf{w}}_i - \mathbf{u}_i(\theta_i)) \right\}. \end{aligned} \quad (24)$$

To obtain an estimation of θ_i , we need to minimize the following cost function

$$\phi_i(\theta_i) = \left(\hat{\mathbf{w}}_i - \mathbf{u}_i(\theta_i) \right)^H (\beta_i \Psi)^{-1} \left(\hat{\mathbf{w}}_i - \mathbf{u}_i(\theta_i) \right). \quad (25)$$

To decrease even more the complexity cost of the proposed robust calibration scheme, we use the BCD algorithm [40,41] in the M-step. Consequently, we obtain analytical solutions for each single source sub-problems in (25), as shown below.

B. Use of the BCD algorithm to minimize (25)

In (25), the optimization is performed w.r.t. θ_i . However, this unknown parameter vector can be partitioned according to the antennas, as expressed in (19). In the following, we perform the optimization of the cost function w.r.t. each $\theta_{i,p}$ (p -th antenna), given the current estimates of all the other $\theta_{i,q}$ with $q \neq p$. This leads to a closed-form expression of $\hat{\theta}_{i,p}$ as function of $\theta_{i,q}$ for $q \neq p$ and the optimization process is repeated for each component vector $\theta_{i,p}$ for $p \in \{1, \dots, M\}$.

If we want to minimize (25) w.r.t. block-coordinate vector $\theta_{i,p}$, we notice that only a subset of $\mathbf{u}_i(\theta_i)$ is actually

dependent on $\theta_{i,p}$, i.e., $\{\mathbf{u}_{i,pq}\}$ for $q > p$, $q \in \{1, \dots, M\}$ and $\{\mathbf{u}_{i,qp}\}$ for $q < p$, $q \in \{1, \dots, M\}$. Therefore, (25) reads

$$\begin{aligned} \phi_i(\theta_{i,p}) = & \sum_{\substack{q=1 \\ q>p}}^M \left(\mathbf{w}_{i,pq} - \mathbf{u}_{i,pq}(\theta_{i,p}) \right)^H (\beta_i \tau_{pq} \Omega)^{-1} \left(\mathbf{w}_{i,pq} - \mathbf{u}_{i,pq}(\theta_{i,p}) \right) + \\ & \sum_{\substack{q=1 \\ q<p}}^M \left(\mathbf{w}_{i,qp} - \mathbf{u}_{i,qp}(\theta_{i,p}) \right)^H (\beta_i \tau_{qp} \Omega)^{-1} \left(\mathbf{w}_{i,qp} - \mathbf{u}_{i,qp}(\theta_{i,p}) \right) + \\ & \text{Constant.} \end{aligned} \quad (26)$$

By Constant, we mean the expressions independent of $\theta_{i,p}$. We show in Appendix B that it is possible to write $\mathbf{u}_{i,pq}$ directly as a function of $\theta_{i,p}$, i.e.,

$$\mathbf{u}_{i,pq}(\theta_{i,p}) = \Sigma_{i,q} \theta_{i,p}. \quad (27)$$

In the same way, we have

$$\mathbf{u}_{i,qp}(\theta_{i,p}) = \Upsilon_{i,q} \theta_{i,p}^*. \quad (28)$$

Notation and calculations being introduced in Appendix B, we only present here the results obtained, i.e., the expression of the estimated entries of the Jones matrix associated with the path from the i -th calibrator source to the p -th sensor which is given by

$$\hat{\theta}_{i,p} = \begin{cases} (\Sigma_i^H \mathbf{A}_{i,p} \Sigma_i + \Upsilon_i^H \tilde{\mathbf{A}}_{i,p} \Upsilon_i)^{-1} \times \\ (\Sigma_i^H \mathbf{A}_{i,p} \mathbf{w}_{i,p} + \Upsilon_i^H \tilde{\mathbf{A}}_{i,p} \tilde{\mathbf{w}}_{i,p}) & \text{for } 1 < p < M \\ (\Sigma_i^H \mathbf{A}_{i,p} \Sigma_i)^{-1} \Sigma_i^H \mathbf{A}_{i,p} \mathbf{w}_{i,p} & \text{for } p = 1 \\ (\Upsilon_i^H \tilde{\mathbf{A}}_{i,p} \Upsilon_i)^{-1} \Upsilon_i^H \tilde{\mathbf{A}}_{i,p} \tilde{\mathbf{w}}_{i,p} & \text{for } p = M \end{cases} \quad (29)$$

Therefore, $\theta_{i,p}$ for $p \in \{1, \dots, M\}$ are estimated in an iterative loop. With (25) and (29), it can be proven that the BCD algorithm leads to unique solutions and thus, convergence to at least a local maximizer, is ensured [63]. If the M-step is performed exactly (i.e., the BCD gives the exact minimizer of (25) and consequently, the M step is exactly solved), convergence of the EM algorithm to a stationary point is ensured (to avoid the unusual case of convergence to a saddle point, a proper initialization is required) [30], with a theoretical infinite number of iterations. Finally, in this case, convergence of the concentrated MLE is guaranteed for an infinite number of iterations since the value of the cost function at each step can either improve or maintain but cannot worsen [64]. In practice, only a finite number of iterations is considered in each loop, so we might not attain local convergence. However, we show in section VI-A the relatively good numerical stability of the algorithm.

The scheme of the proposed algorithm is described in Algorithm 1.

V. STRUCTURED JONES MATRICES

We recall that the output of Algorithm 1 is the estimate of each Jones matrix denoted by $\hat{\mathbf{J}}_{i,p}$ for $i \in \{1, \dots, D\}$ and $p \in \{1, \dots, M\}$. In the following, we consider the data model in (7) for the specific 3DC calibration regime, and intend to estimate the unknown parameter vector of interest $\epsilon^{3\text{DC}}$ in a

Algorithm 1: Relaxed concentrated ML based calibration algorithm

```

input   :  $D, M, B, \mathbf{C}_i, \beta_i, \mathbf{x}$ 
output  :  $\hat{\theta}$ 
initialize:  $\hat{\Omega} \leftarrow \Omega_{\text{init}}, \hat{\tau} \leftarrow \tau_{\text{init}}, \hat{\theta} \leftarrow \theta_{\text{init}}$ 
1 while stop criterion unreachable do
2   while stop criterion unreachable do
3     E-step:  $\hat{\mathbf{w}}_i$  obtained from (23),  $i \in \{1, \dots, D\}$ 
4     M-step:  $\hat{\theta}_i$  obtained as follows,  $i \in \{1, \dots, D\}$ 
5       while stop criterion unreachable do
6          $\hat{\theta}_{i,p}$  obtained from (29),
7          $p \in \{1, \dots, M\}$ 
8       end
9     end
10    Obtain  $\hat{\Omega}$  from (16) and (17),
11    Obtain  $\hat{\tau}$  from (12)

```

sequential manner. To do so, we use an iterative estimation procedure by optimizing a cost function w.r.t. one parameter while fixing the others.

1) *Estimation of \mathbf{g}_p* : The diagonal elements of the gain matrix are given by

$$\hat{\mathbf{g}}_p = \underset{\mathbf{g}_p}{\text{argmin}} \kappa(\mathbf{g}_p) \quad (30)$$

where $\kappa(\mathbf{g}_p) = \sum_{i=1}^D \|\hat{\mathbf{J}}_{i,p} - \mathbf{G}_p(\mathbf{g}_p) \mathbf{H}_{i,p} \mathbf{Z}_{i,p} \mathbf{F}_i\|_F^2$. We rewrite the cost function as

$$\kappa(\mathbf{g}_p) = \sum_{i=1}^D \text{Tr} \left\{ \left(\hat{\mathbf{J}}_{i,p} - \mathbf{G}_p(\mathbf{g}_p) \mathbf{R}_{i,p} \right) \left(\hat{\mathbf{J}}_{i,p} - \mathbf{G}_p(\mathbf{g}_p) \mathbf{R}_{i,p} \right)^H \right\} \quad (31)$$

in which $\mathbf{R}_{i,p} = \mathbf{H}_{i,p} \mathbf{Z}_{i,p} \mathbf{F}_i$. The derivation of $\kappa(\mathbf{g}_p)$ w.r.t. the k -th element $[\mathbf{g}_p]_k$ leads to

$$\frac{\partial \kappa(\mathbf{g}_p)}{\partial [\mathbf{g}_p]_k} = \sum_{i=1}^D \text{Tr} \left\{ -\mathbf{e}_k \mathbf{e}_k^T \mathbf{R}_{i,p} \hat{\mathbf{J}}_{i,p}^H + \mathbf{e}_k \mathbf{e}_k^T \mathbf{R}_{i,p} \mathbf{R}_{i,p}^H \mathbf{G}_p^H \right\}. \quad (32)$$

Let us denote $\mathbf{X}_{i,p} = \mathbf{R}_{i,p} \hat{\mathbf{J}}_{i,p}^H$ and $\mathbf{W}_{i,p} = \mathbf{R}_{i,p} \mathbf{R}_{i,p}^H$. From (32), we deduce the equation satisfied by the gain matrix

$$\begin{aligned} \sum_{i=1}^D [\mathbf{X}_{i,p}]_{k,k} &= \sum_{i=1}^D [\mathbf{W}_{i,p} \hat{\mathbf{G}}_p^H]_{k,k} \\ &= \sum_{i=1}^D [\mathbf{W}_{i,p}]_{k,k} [\hat{\mathbf{G}}_p^*]_{k,k} \end{aligned} \quad (33)$$

for $k \in \{1, 2\}$, since \mathbf{G}_p is a diagonal matrix. Therefore, each complex gain element is estimated as

$$[\hat{\mathbf{g}}_p]_k = \left(\sum_{i=1}^D [\mathbf{W}_{i,p}^*]_{k,k} \right)^{-1} \sum_{i=1}^D [\mathbf{X}_{i,p}^*]_{k,k}. \quad (34)$$

2) *Estimation of α_i* : We first need to estimate $\varphi_{i,p}$ (we recall that $\varphi_{i,p} = \eta_i u_p + \zeta_i v_p$). This is done as follows

$$\hat{\varphi}_{i,p} = \underset{\varphi_{i,p}}{\text{argmin}} \tilde{\kappa}(\varphi_{i,p}) \quad (35)$$

where $\tilde{\kappa}(\varphi_{i,p}) = \|\hat{\mathbf{J}}_{i,p} - \mathbf{G}_p \mathbf{H}_{i,p} \mathbf{Z}_{i,p}(\varphi_{i,p}) \mathbf{F}_i\|_F^2$. Taking the derivative of $\tilde{\kappa}(\varphi_{i,p})$ w.r.t. $\varphi_{i,p}$ and setting the result to zero, we obtain

$$\text{Tr} \left\{ j \exp^{-j\hat{\varphi}_{i,p}} \hat{\mathbf{J}}_{i,p} \mathbf{F}_i^H \mathbf{H}_{i,p}^H \mathbf{G}_p^H - j \exp^{j\hat{\varphi}_{i,p}} \mathbf{G}_p \mathbf{H}_{i,p} \mathbf{F}_i \hat{\mathbf{J}}_{i,p}^H \right\} = 0 \quad (36)$$

which leads to

$$\exp \left\{ 2j\hat{\varphi}_{i,p} \right\} = \frac{\text{Tr} \left\{ \mathbf{M}_{i,p} \right\}}{\text{Tr} \left\{ \mathbf{M}_{i,p}^H \right\}} \quad (37)$$

where $\mathbf{M}_{i,p} = \hat{\mathbf{J}}_{i,p} \mathbf{F}_i^H \mathbf{H}_{i,p}^H \mathbf{G}_p^H$.

In the case of a compact array, we can write for the i -th source

$$\varphi_i^T = \hat{\alpha}_i^T \mathbf{\Lambda} \quad (38)$$

where $\varphi_i = [\hat{\varphi}_{i,1}, \dots, \hat{\varphi}_{i,M}]^T$ and $\mathbf{\Lambda} = \begin{bmatrix} u_1 & \dots & u_M \\ v_1 & \dots & v_M \end{bmatrix}$. Therefore, estimation of the directional shifts due to propagation in the ionosphere is given by

$$\hat{\alpha}_i^T = \frac{\varphi_i^T \mathbf{\Lambda}^H \begin{bmatrix} \sum_{p=1}^M v_p^2 & -\sum_{p=1}^M u_p v_p \\ -\sum_{p=1}^M v_p u_p & \sum_{p=1}^M u_p^2 \end{bmatrix}}{\sum_{p=1}^M u_p^2 \sum_{p=1}^M v_p^2 - (\sum_{p=1}^M u_p v_p)^2}. \quad (39)$$

3) Estimation of ϑ_i : We consider the following minimization problem

$$\hat{\vartheta}_i = \underset{\vartheta_i}{\text{argmin}} \sum_{p=1}^M \|\hat{\mathbf{J}}_{i,p} - \mathbf{G}_p \mathbf{H}_{i,p} \mathbf{Z}_{i,p} \mathbf{F}_i(\vartheta_i)\|_F^2. \quad (40)$$

We assume a large number of antennas M while the number of calibrator sources D is relatively reduced, such that observations outnumber unknown parameters. For each source, the 1D optimization in (40) can be computed in a reasonable computational time through a classical data grid search or a Newton type algorithm.

Finally, the proposed algorithm for the structured Jones matrices case regarding 3DC calibration regime is given in Algorithm 2.

Algorithm 2: Case of structured Jones matrices

input : $D, M, B, \mathbf{C}_i, \beta_i, \mathbf{x}, \hat{\mathbf{J}}_{i,p}$ as output of Algorithm 1, $i \in \{1, \dots, D\}$ and $p \in \{1, \dots, M\}$
output : $\hat{\epsilon}^{3\text{DC}}$
initialize: $\hat{\epsilon}^{3\text{DC}} \leftarrow \epsilon_{\text{init}}^{3\text{DC}}$
1 while stop criterion unreached **do**
1 Obtain $\hat{\vartheta}_i$ from (40), $i \in \{1, \dots, D\}$
2 Obtain $\hat{\mathbf{g}}_p$ from (34), $p \in \{1, \dots, M\}$
3 Obtain $\hat{\alpha}_i = [\eta_i, \zeta_i]^T$ from (39), $i \in \{1, \dots, D\}$
4 end

VI. NUMERICAL SIMULATIONS

In this part, we first aim to assess the statistical performance of Algorithm 1, when the noise model matches our noise assumption, i.e., a SIRP noise modeling. Afterwards, we intend to study our proposed algorithm, in a more realistic scenario where outliers are present. More specifically, the sky is composed of D known bright calibrator sources but also D' weak non-calibrator sources which are absorbed in the noise component and act as outliers. Under such assumption, we compare our scheme with the recently introduced robust calibration approach based on Student's t [35] and with the traditional Gaussian cases [32]. Finally, we apply Algorithm 2 to the introduced 3DC calibration regime where Jones matrices are structured.

A. Numerical results under SIRP noise

The unknown parameters to estimate, $\boldsymbol{\theta}$ given in (19), correspond to the real and imaginary parts of the entries of all Jones matrices. The additive noise in (4) is assumed to follow a SIRP as in (5) and the number of Monte Carlo runs is set to 100.

In order to evaluate the estimation performance of the relaxed concentrated ML based calibration algorithm, we fix the noise distribution, e.g., as a Student's t and make use of the Cramér-Rao bound (CRB) [65]. To do so, each random texture component is supposed to follow an inverse gamma distribution [66]. As an example,

$$\tau_{pq} \sim \mathcal{IG}(\nu/2, \nu/2), \quad (41)$$

with ν degrees of freedom [67] and we choose for example $[\boldsymbol{\Omega}]_{k,l} = \sigma^2 0.9^{|k-l|} \exp^{j\frac{\pi}{2}(k-l)}$.

The covariance inequality principle states that, under quite general/weak conditions, the variance satisfies

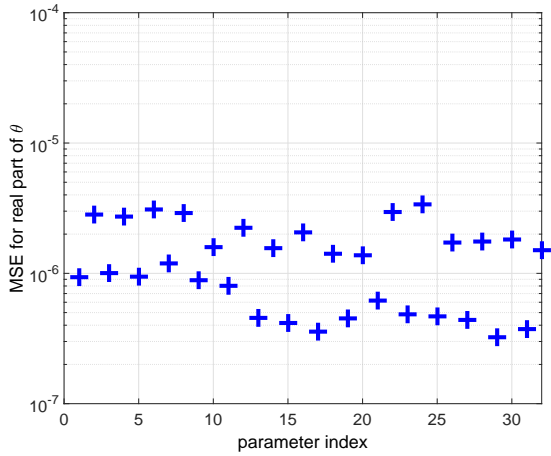
$$\text{MSE}([\hat{\boldsymbol{\theta}}]_k) = \mathbb{E} \left\{ \left([\hat{\boldsymbol{\theta}}]_k - [\boldsymbol{\theta}]_k \right)^2 \right\} \geq [\text{CRB}(\boldsymbol{\theta})]_{k,k} \quad (42)$$

where the CRB is given as the inverse of the Fisher information matrix (FIM) \mathbf{F} . A Slepian-Bangs type formula of the FIM for SIRP observations is given in [68] which can be adapted to our case and reads

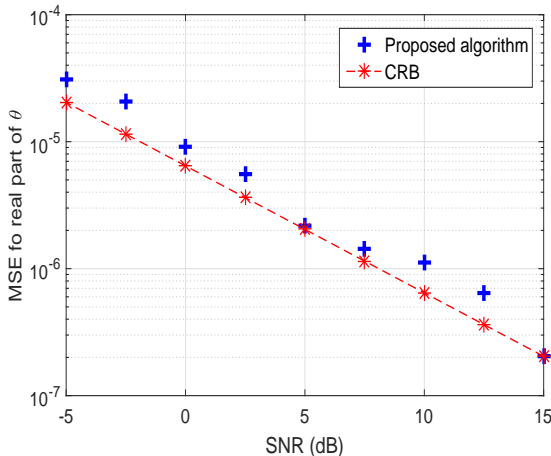
$$[\mathbf{F}]_{k,l} = 2 \frac{\nu+4}{\nu+5} \sum_{pq} \Re \left\{ \frac{\partial \tilde{\mathbf{v}}_{pq}^H(\boldsymbol{\theta})}{\partial [\boldsymbol{\theta}]_k} \boldsymbol{\Omega}^{-1} \frac{\partial \tilde{\mathbf{v}}_{pq}(\boldsymbol{\theta})}{\partial [\boldsymbol{\theta}]_l} \right\}. \quad (43)$$

Note that the noise parameters are decoupled from the parameters of interest. Consequently, only the part corresponding to the latter is kept in the FIM expression.

First, in Fig. 2(a), we plot the mean square error (MSE) of the real part of each unknown parameter, obtained with Algorithm 1, for a signal-to-noise ratio SNR = 15dB. We only plot the parameters relative to one given source, the behavior being the same for any source. We also compare the MSE of one given parameter, as a function of the SNR, to its corresponding CRB in Fig. 2(b). This enables to assess the statistical performance of Algorithm 1, and we notice that the MSE approaches the CRB. The small gap between the bound and the algorithm is explained by the relaxed nature hypothesis used in the design of Algorithm 1. Indeed, we



(a)



(b)

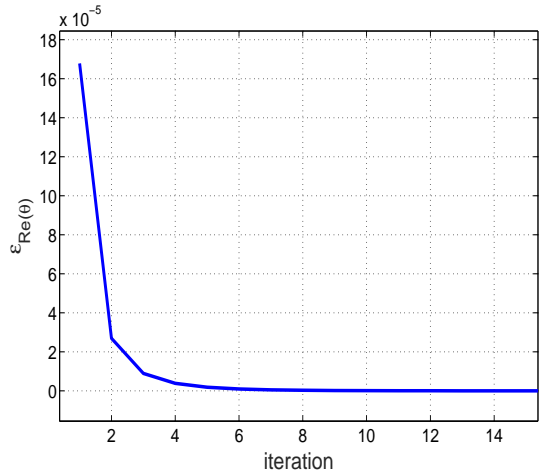
Fig. 2. (a) MSE of the real part of the first 32 unknown parameters for a given SNR, (b) MSE vs. SNR for the real part of a given unknown parameter and the corresponding CRB, for $D = 2$ bright signal sources and $M = 8$ antennas, leading to 128 real unknown parameters of interest to estimate and 224 measurements.

have assumed unknown and deterministic texture parameters when we derived the estimates using Algorithm 1, but in the data model, these parameters are generated as random variables following inverse gamma distribution and the CRB was derived using the prior of the pdf of the texture.

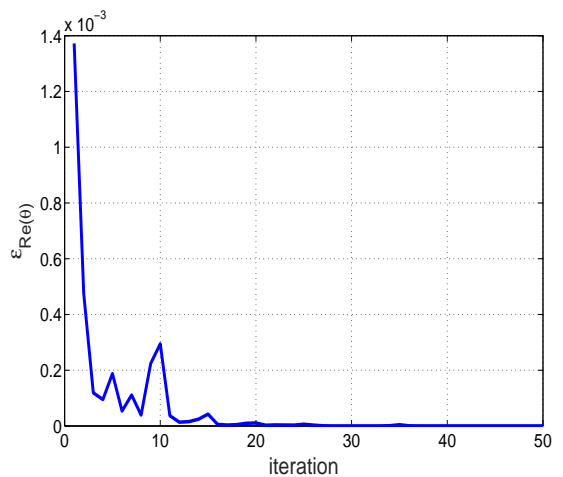
Second, we now aim to investigate numerically the convergence properties of our algorithm, which is composed of 3 loops. In each of these loops, θ is updated at each iteration. We therefore consider the following quantity

$$\epsilon_{\Re\{\theta\}}^h = \|\Re\{\theta^h - \theta^{h-1}\}\|_2^2 \quad (44)$$

where h refers to the h -th iteration. In Fig. 3 we present the convergence rate of loops described in Algorithm 1 at line 1 and 2 (the analysis of convergence of the third loop, given in line 5, has the same behavior as the loop in line 2 and thus, is not reported here). We note that around 5 iterations are required in loop 2 to attain convergence while approximately 20 iterations are needed for the algorithm to be stable, in



(a)



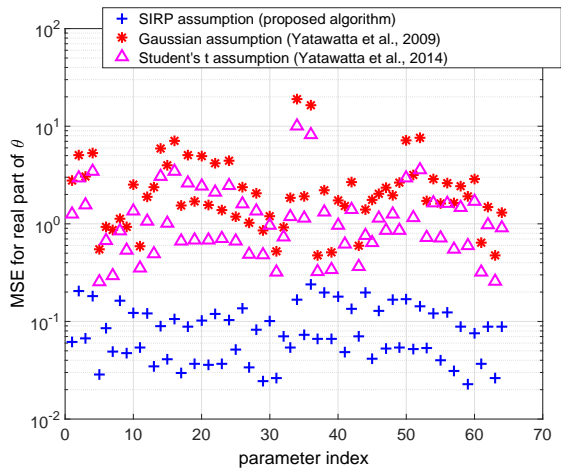
(b)

Fig. 3. $\epsilon_{\Re\{\theta\}}^h$ as function of the h -th iteration, for loop in line 2 (a), and in line 1 (b) from Algorithm 1.

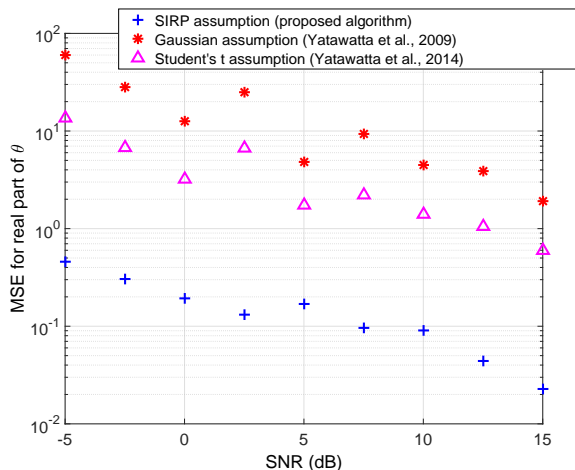
loop 1. Nevertheless, in simulations, we notice that only 3 to 4 iterations are sufficient to get close to the CRB.

B. Numerical results using a realistic model

Let us investigate the robustness of our proposed calibration procedure in a realistic situation, and compare it with the state-of-the-art. To do so, we consider D calibrator sources, D' weak outlier sources and Gaussian background noise in our data model. The real parameters of interest to estimate still correspond to the real and imaginary entries of the Jones matrices associated to the calibrator sources paths. In the following, we compare Algorithm 1 with i) the calibration approach exposed in [35] which assumes a Student's t noise modeling using the so-called expectation-conditional maximization either algorithm [69,70] and ii) the traditional calibration scheme based on zero-mean white Gaussian noise modeling using a least squares approach [32]. The comparative results are plotted in Fig. 4, for the same computation times. We notice better estimation performance with Algorithm 1, since we did not specify any particular noise distribution in



(a)



(b)

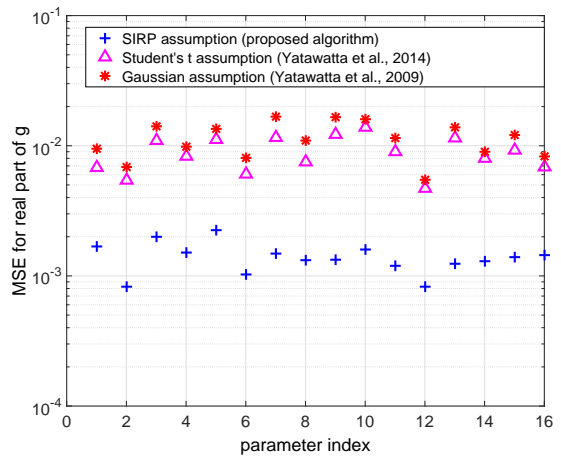
Fig. 4. (a) MSE of the real part of the 64 unknown parameters for a given SNR, (b) MSE vs. SNR for the real part of a given unknown parameter, for $D = 2$, $M = 8$ and $D' = 8$, leading to 128 real parameters of interest to estimate and 224 measurements.

our procedure and a SIRP includes many different types of distributions. Furthermore, no assumption was made about independent entries in the noise vector, thus ensuring more flexibility and robustness.

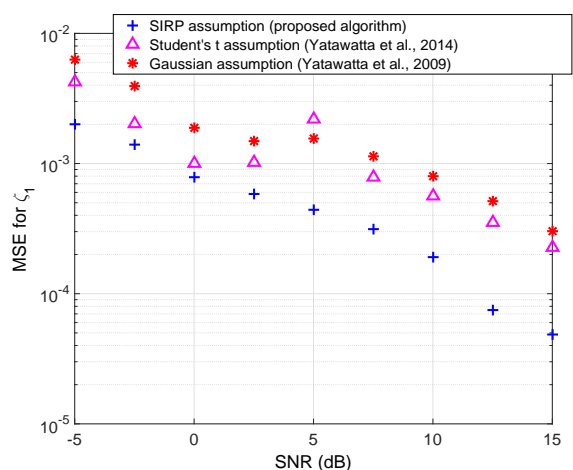
C. Structured case

In order to compare our proposed global algorithm (Algorithm 1 followed by Algorithm 2) with the approach based on Student's t [35] and the Gaussian case [32] which were both introduced in the non-structured case, we apply Algorithm 2 on the output of these two latter algorithms.

Each Jones matrix is generated according to (7) with $\mathbf{g} = [\mathbf{g}_1^T, \dots, \mathbf{g}_M^T]^T$. In all exposed simulations, we consider similar computation times for the three presented patterns. We show the results only for the complex gains, cf. Fig. 5(a), and the source offset ζ_1 , cf. Fig 5(b), due to lack of space, the behavior being the same for the other parameters, i.e., ϑ_1 , ϑ_2 , η_2 , η_1 and ζ_2 . In the case of structured Jones matrices, adapted to the 3DC calibration regime, and in the presence of



(a)



(b)

Fig. 5. (a) MSE of the real part of the 16 complex gains for a given SNR, (b) MSE of ζ_1 vs. SNR, for $D = 2$, $M = 8$, and $D' = 4$, leading to 38 real parameters of interest to estimate and 224 measurements.

outliers, we still notice the better performances of Algorithm 1, compared to the state-of-the-art. This is expected since better estimation of the Jones entries leads to better estimation of the physical parameters describing the structured Jones matrices.

VII. CONCLUSION

In this paper, we proposed a robust calibration technique where perturbation effects are modeled thanks to Jones matrices. To deal with the presence of outliers in our data, the introduced ML estimation method is based on SIRP noise modeling, leading to a relaxed concentrated ML based calibration algorithm. Numerical simulations show that the proposed algorithm is more robust to the presence of outliers in comparison with the state-of-the-art, for both non-structured and structured Jones matrices, with a reasonable computational complexity.

APPENDIX A

We describe here the corresponding expressions of (12) and (16) when we assume a different Ω_{pq} for $p < q$, $p, q \in$

$\{1, \dots, M\}$. In this case, the log-likelihood function is written as

$$\begin{aligned} \log f(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\tau}, \boldsymbol{\Omega}_{12}, \boldsymbol{\Omega}_{13}, \dots, \boldsymbol{\Omega}_{(M-1)M}) &= -4B \log \pi \\ &- 4 \sum_{pq} \log \tau_{pq} - \sum_{pq} \log |\boldsymbol{\Omega}_{pq}| - \sum_{pq} \frac{1}{\tau_{pq}} \mathbf{a}_{pq}^H(\boldsymbol{\theta}) \boldsymbol{\Omega}_{pq}^{-1} \mathbf{a}_{pq}(\boldsymbol{\theta}). \end{aligned} \quad (45)$$

For each antenna pair, the texture estimate is given by

$$\hat{\tau}_{pq} = \frac{1}{4} \mathbf{a}_{pq}^H(\boldsymbol{\theta}) \boldsymbol{\Omega}_{pq}^{-1} \mathbf{a}_{pq}(\boldsymbol{\theta}) \quad (46)$$

while the speckle covariance estimate reads

$$\hat{\boldsymbol{\Omega}}_{pq}^{h+1} = 4 \frac{\mathbf{a}_{pq}(\boldsymbol{\theta}) \mathbf{a}_{pq}^H(\boldsymbol{\theta})}{\mathbf{a}_{pq}^H(\boldsymbol{\theta}) \left(\hat{\boldsymbol{\Omega}}_{pq}^h \right)^{-1} \mathbf{a}_{pq}(\boldsymbol{\theta})}. \quad (47)$$

The remainder of the algorithm is straightforwardly obtained using (47).

APPENDIX B

We present here the steps to obtain (29). Firstly, for sake of clarity, let us denote $\mathbf{c}_i = [c_{i1}, c_{i2}, c_{i3}, c_{i4}]^T$ to refer to the four entries of the vectorization of source coherency matrix \mathbf{C}_i .

Likewise, for the i -th source, we write $\mathbf{J}_{i,p}(\boldsymbol{\theta}_{i,p}) = \begin{bmatrix} p_{i1} & p_{i2} \\ p_{i3} & p_{i4} \end{bmatrix}$ for the p -th antenna and $\mathbf{J}_{i,q}(\boldsymbol{\theta}_{i,q}) = \begin{bmatrix} q_{i1} & q_{i2} \\ q_{i3} & q_{i4} \end{bmatrix}$ for the q -th antenna, i.e., $\boldsymbol{\theta}_{i,p} = [p_{i1}, p_{i2}, p_{i3}, p_{i4}]^T$ and $\boldsymbol{\theta}_{i,q} = [q_{i1}, q_{i2}, q_{i3}, q_{i4}]^T$. Using these latter notation, we obtain (27) where

$$\boldsymbol{\Sigma}_{i,q} = \begin{bmatrix} \alpha_{i,q} & \beta_{i,q} & 0 & 0 \\ 0 & 0 & \alpha_{i,q} & \beta_{i,q} \\ \gamma_{i,q} & \rho_{i,q} & 0 & 0 \\ 0 & 0 & \gamma_{i,q} & \rho_{i,q} \end{bmatrix} \quad (48)$$

in which $\alpha_{i,q} = q_{i1}^* c_{i1} + q_{i2}^* c_{i3}$, $\beta_{i,q} = q_{i1}^* c_{i2} + q_{i2}^* c_{i4}$, $\gamma_{i,q} = q_{i3}^* c_{i1} + q_{i4}^* c_{i3}$ and $\rho_{i,q} = q_{i3}^* c_{i2} + q_{i4}^* c_{i4}$.

We also obtain (28) where

$$\boldsymbol{\Upsilon}_{i,q} = \begin{bmatrix} \lambda_{i,q} & \mu_{i,q} & 0 & 0 \\ \nu_{i,q} & \xi_{i,q} & 0 & 0 \\ 0 & 0 & \lambda_{i,q} & \mu_{i,q} \\ 0 & 0 & \nu_{i,q} & \xi_{i,q} \end{bmatrix} \quad (49)$$

in which $\lambda_{i,q} = q_{i1} c_{i1} + q_{i2} c_{i2}$, $\mu_{i,q} = q_{i1} c_{i3} + q_{i2} c_{i4}$, $\nu_{i,q} = q_{i3} c_{i1} + q_{i4} c_{i2}$ and $\xi_{i,q} = q_{i3} c_{i3} + q_{i4} c_{i4}$.

Finally, the cost function in (26) can be written as

$$\begin{aligned} \phi_i(\boldsymbol{\theta}_{i,p}) &= \left(\mathbf{w}_{i,p} - \mathbf{u}_{i,p}(\boldsymbol{\theta}_{i,p}) \right)^H \mathbf{A}_{i,p} \left(\mathbf{w}_{i,p} - \mathbf{u}_{i,p}(\boldsymbol{\theta}_{i,p}) \right) + \\ &\left(\tilde{\mathbf{w}}_{i,p} - \tilde{\mathbf{u}}_{i,p}(\boldsymbol{\theta}_{i,p}) \right)^H \tilde{\mathbf{A}}_{i,p} \left(\tilde{\mathbf{w}}_{i,p} - \tilde{\mathbf{u}}_{i,p}(\boldsymbol{\theta}_{i,p}) \right) + \text{Constant} \end{aligned} \quad (50)$$

where $\mathbf{w}_{i,p} = [\mathbf{w}_{i,p(p+1)}^T, \dots, \mathbf{w}_{i,pM}^T]^T$, $\mathbf{u}_{i,p}(\boldsymbol{\theta}_{i,p}) = [\mathbf{u}_{i,p(p+1)}^T(\boldsymbol{\theta}_{i,p}), \dots, \mathbf{u}_{i,pM}^T(\boldsymbol{\theta}_{i,p})]^T$ and $\mathbf{A}_{i,p} = \text{bdiag}\{\beta_i \tau_{p(p+1)} \boldsymbol{\Omega}, \dots, \beta_i \tau_{pM} \boldsymbol{\Omega}\}^{-1}$.

Furthermore, we have $\tilde{\mathbf{w}}_{i,p} = [\mathbf{w}_{i,1p}^T, \dots, \mathbf{w}_{i,(p-1)p}^T]^T$, $\tilde{\mathbf{u}}_{i,p}(\boldsymbol{\theta}_{i,p}) = [\mathbf{u}_{i,1p}^T(\boldsymbol{\theta}_{i,p}), \dots, \mathbf{u}_{i,(p-1)p}^T(\boldsymbol{\theta}_{i,p})]^T$ and $\tilde{\mathbf{A}}_{i,p} = \text{bdiag}\{\beta_i \tau_{1p} \boldsymbol{\Omega}^*, \dots, \beta_i \tau_{(p-1)p} \boldsymbol{\Omega}^*\}^{-1}$.

We make use of (27) in what follows

$$\mathbf{u}_{i,p}(\boldsymbol{\theta}_{i,p}) = \begin{bmatrix} \mathbf{u}_{i,p(p+1)}(\boldsymbol{\theta}_{i,p}) \\ \vdots \\ \mathbf{u}_{i,pM}(\boldsymbol{\theta}_{i,p}) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Sigma}_{i,p+1} \boldsymbol{\theta}_{i,p} \\ \vdots \\ \boldsymbol{\Sigma}_{i,M} \boldsymbol{\theta}_{i,p} \end{bmatrix} = \boldsymbol{\Sigma}_i \boldsymbol{\theta}_{i,p} \quad (51)$$

where $\boldsymbol{\Sigma}_i = [\boldsymbol{\Sigma}_{i,p+1}^T, \dots, \boldsymbol{\Sigma}_{i,M}^T]^T$. Likewise, we use (28) in

$$\tilde{\mathbf{u}}_{i,p}(\boldsymbol{\theta}_{i,p}) = \begin{bmatrix} \mathbf{u}_{i,1p}^*(\boldsymbol{\theta}_{i,p}) \\ \vdots \\ \mathbf{u}_{i,(p-1)p}^*(\boldsymbol{\theta}_{i,p}) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Upsilon}_{i,1}^* \boldsymbol{\theta}_{i,p} \\ \vdots \\ \boldsymbol{\Upsilon}_{i,p-1}^* \boldsymbol{\theta}_{i,p} \end{bmatrix} = \boldsymbol{\Upsilon}_i \boldsymbol{\theta}_{i,p} \quad (52)$$

in which $\boldsymbol{\Upsilon}_i = [\boldsymbol{\Upsilon}_{i,1}^{*T}, \dots, \boldsymbol{\Upsilon}_{i,p-1}^{*T}]^T$. Inserting (51) and (52) into (50) and taking the derivative w.r.t. $\boldsymbol{\theta}_{i,p}$ leads to the expressions in (29), using the fact that $\mathbf{A}_{i,p}$ and $\tilde{\mathbf{A}}_{i,p}$ are Hermitian.

REFERENCES

- [1] A. R. Thompson, J. M. Moran, and G. W. Swenson Jr., *Interferometry and synthesis in radio astronomy*, 2nd ed. John Wiley & Sons, 2001.
- [2] M. P. Van Haarlem *et al.*, "LOFAR: The LOw-Frequency ARray," *Astronomy & Astrophysics*, vol. 556, no. A2, 2013.
- [3] P. E. Dewdney, P. J. Hall, R. T. Schilizzi, and T. J. L. Lazio, "The Square Kilometre Array," *Proceedings of the IEEE*, vol. 97, no. 8, pp. 1482–1496, 2009.
- [4] S. J. Wijnholds and A.-J. van der Veen, "Fundamental imaging limits of radio telescope arrays," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 5, pp. 613–623, 2008.
- [5] U. Rau, S. Bhatnagar, M. A. Voronkov, and T. J. Cornwell, "Advances in calibration and imaging techniques in radio interferometry," *Proceedings of the IEEE*, vol. 97, no. 8, pp. 1472–1481, 2009.
- [6] M. de Vos, A. W. Gunst, and R. Nijboer, "The LOFAR telescope: System architecture and signal processing," *Proceedings of the IEEE*, vol. 97, no. 8, pp. 1431–1437, 2009.
- [7] A.-J. van der Veen and S. J. Wijnholds, "Signal processing tools for radio astronomy," in *Handbook of Signal Processing Systems*. Springer, 2013, pp. 421–463.
- [8] D. A. Mitchell, L. J. Greenhill, R. B. Wayth, R. J. Sault, C. J. Lonsdale, R. J. Cappallo, M. F. Morales, and S. M. Ord, "Real-time calibration of the Murchison Widefield Array," *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 5, pp. 707–717, 2008.
- [9] S. J. Wijnholds, A.-J. van der Veen, F. De Stefani, E. La Rosa, and A. Farina, "Signal processing challenges for radio astronomical arrays," in *International Conference on Acoustics, Speech and Signal Processing*, Florence, Italy, 2014, pp. 5382–5386.
- [10] S. J. 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, no. 1, pp. 30–42, 2010.
- [11] D. R. Fuhrmann, "Estimation of sensor gain and phase," *IEEE Transactions on Signal Processing*, vol. 42, no. 1, pp. 77–87, 1994.
- [12] B. C. Ng and C. M. S. See, "Sensor-array calibration using a maximum-likelihood approach," *IEEE Transactions on Antennas and Propagation*, vol. 44, no. 6, pp. 827–835, 1996.
- [13] B. P. Ng, J. P. Lie, M. H. Er, and A. Feng, "A practical simple geometry and gain/phase calibration technique for antenna array processing," *IEEE Transactions on Antennas and Propagation*, vol. 57, no. 7, pp. 1963–1972, 2009.
- [14] J. T. H. Lo and S. L. Marple, "Eigenstructure methods for array sensor localization," in *International Conference on Acoustics, Speech, and Signal Processing*, Dallas, Texas, 1987, pp. 2260–2263.
- [15] B. C. Ng and W. Ser, "Array shape calibration using sources in known locations," in *ICCS/ISITA: Communications on the Move*, Singapore, 1992, pp. 836–840.
- [16] B. C. Ng and A. Nehorai, "Active array sensor localization," *Signal Processing*, vol. 44, no. 3, pp. 309–327, 1995.
- [17] Y. Rockah and P. Schultheiss, "Array shape calibration using sources in unknown locations—part i: Far-field sources," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, no. 3, pp. 286–299, 1987.

- [18] A. J. Weiss and B. Friedlander, "Array shape calibration using sources in unknown locations—a maximum likelihood approach," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 12, pp. 1958–1966, 1989.
- [19] B. Friedlander and A. J. Weiss, "Direction finding in the presence of mutual coupling," *IEEE Transactions on Antennas and Propagation*, vol. 39, no. 3, pp. 273–284, 1991.
- [20] M. P. Wylie, S. Roy, and H. Messer, "Joint DOA estimation and phase calibration of linear equispaced (LES) arrays," *IEEE Transactions on Signal Processing*, vol. 42, no. 12, pp. 3449–3459, 1994.
- [21] L. Qiong, G. Long, and Y. Zhongfu, "An overview of self-calibration in sensor array processing," in *6th International Symposium on Antennas, Propagation and EM Theory*, Beijing, China, 2003, pp. 279–282.
- [22] B. P. Flanagan and K. L. Bell, "Array self-calibration with large sensor position errors," *Signal Processing*, vol. 81, no. 10, pp. 2201–2214, 2001.
- [23] J. Baars, R. Genzel, I. Pauliny-Toth, and A. Witzel, "The absolute spectrum of CAS A—an accurate flux density scale and a set of secondary calibrators," *Astronomy and Astrophysics*, vol. 61, no. 1, pp. 99–106, 1977.
- [24] A.-J. Boonstra and A.-J. van der Veen, "Gain calibration methods for radio telescope arrays," *IEEE Transactions on Signal Processing*, vol. 51, no. 1, pp. 25–38, 2003.
- [25] S. J. Wijnholds and A.-J. van der Veen, "Multisource self-calibration for sensor arrays," *IEEE Transactions on Signal Processing*, vol. 57, no. 9, pp. 3512–3522, 2009.
- [26] S. J. Wijnholds, "Fish-eye observing with phased array radio telescopes," Ph.D. dissertation, Technische Universiteit Delft, Delft, The Netherlands, 2010.
- [27] S. Salvini and S. J. Wijnholds, "Fast gain calibration in radio astronomy using alternating direction implicit methods: Analysis and applications," *Astronomy & Astrophysics*, vol. 571, p. A97, 2014.
- [28] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 39, no. 1, pp. 1–38, 1977.
- [29] T. K. Moon, "The Expectation-Maximization algorithm," *IEEE Signal Processing Magazine*, vol. 13, no. 6, pp. 47–60, 1996.
- [30] G. J. McLachlan and T. Krishnan, *The EM algorithm and extensions*, 2nd ed. John Wiley & Sons, 2008.
- [31] J. A. Fessler and A. O. Hero, "Space-alternating generalized expectation-maximization algorithm," *IEEE Transactions on Signal Processing*, vol. 42, no. 10, pp. 2664–2677, 1994.
- [32] S. Yatawatta, S. Zaroubi, G. de Bruyn, L. Koopmans, and J. Noordam, "Radio interferometric calibration using the SAGE algorithm," in *13th Digital Signal Processing Workshop and 5th Signal Processing Education Workshop*, Marco Island, FL, 2009, pp. 150–155.
- [33] J. Raza, A.-J. Boonstra, and A.-J. Van der Veen, "Spatial filtering of RF interference in radio astronomy," *IEEE Signal Processing Letters*, vol. 9, no. 2, pp. 64–67, 2002.
- [34] S. Van Der Tol and A.-J. Van Der Veen, "Performance analysis of spatial filtering of RF interference in radio astronomy," *IEEE Transactions on Signal Processing*, vol. 53, no. 3, pp. 896–910, 2005.
- [35] S. Yatawatta and S. Kazemi, "Robust radio interferometric calibration," in *International Conference on Acoustics, Speech and Signal Processing*, Florence, Italy, 2014, pp. 5429–5433.
- [36] A.-J. Boonstra, "Radio frequency interference mitigation in radio astronomy," Ph.D. dissertation, Technische Universiteit Delft, Delft, The Netherlands, 2005.
- [37] E. Jay, "Détection en environnement non gaussien," Ph.D. dissertation, Université de Cergy Pontoise, 2002.
- [38] K. Yao, "Spherically invariant random processes: Theory and applications," in *Communications, Information and Network Security*, V. Bhargava et al., Eds. Dordrecht, The Netherlands: Kluwer Academic: Springer, 2003, pp. 315–331.
- [39] J. Wang, A. Dogandic, and A. Nehorai, "Maximum likelihood estimation of compound-gaussian clutter and target parameters," *IEEE Transactions on Signal Processing*, vol. 54, no. 10, pp. 3884–3898, 2006.
- [40] J. Friedman, T. Hastie, H. Höfling, R. Tibshirani et al., "Pathwise coordinate optimization," *The Annals of Applied Statistics*, vol. 1, no. 2, pp. 302–332, 2007.
- [41] M. Hong, M. Razaviyayn, Z.-Q. Luo, and J.-S. Pang, "A unified algorithmic framework for block-structured optimization involving big data: With applications in machine learning and signal processing," *IEEE Signal Process. Mag.*, vol. 33, no. 1, pp. 57–77, 2016.
- [42] J. P. Hamaker, J. D. Bregman, and R. J. Sault, "Understanding radio polarimetry. I. Mathematical foundations," *Astronomy and Astrophysics Supplement Series*, vol. 117, no. 1, pp. 137–147, 1996.
- [43] O. M. Smirnov, "Revisiting the radio interferometer measurement equation. I. A full-sky Jones formalism," *Astronomy & Astrophysics*, vol. 527, no. A106, 2011.
- [44] J. E. Noordam, "The measurement equation of a generic radio telescope, AIPS++ implementation note nr 185," ASTRON, Tech. Rep., 1996.
- [45] J. E. Noordam and O. M. Smirnov, "The MeqTrees software system and its use for third-generation calibration of radio interferometers," *Astronomy & Astrophysics*, vol. 524, no. A61, 2010.
- [46] C. Lonsdale, "Calibration approaches," LFD memo 015, Tech. Rep., 2004.
- [47] L. M. Ker, "Radio AGN evolution with low frequency radio surveys," Ph.D. dissertation, The University of Edinburgh, 2012.
- [48] C. D. Nunhokee, "Link between ghost artefacts, source suppression and incomplete calibration sky models," Ph.D. dissertation, Rhodes University, Rhodes, South Africa, 2015.
- [49] X. Zhang, M. N. El Korso, and M. Pesavento, "MIMO radar target localization and performance evaluation under SIRP clutter," *Signal Processing Journal*, vol. 130, pp. 217–232, 2017.
- [50] S. Yatawatta, "Reduced ambiguity calibration for LOFAR," *Experimental Astronomy*, vol. 34, no. 1, pp. 89–103, 2012.
- [51] S. Kazemi, P. Hurlley, O. Öçal, and G. Cherubini, "Blind calibration for radio interferometry using convex optimization," in *3rd International Workshop on Compressed Sensing Theory and its Applications to Radar, Sonar and Remote Sensing*, Pisa, Italy, 2015, pp. 164–168.
- [52] S. van der Tol, B. D. Jeffs, and A.-J. van der Veen, "Self-calibration for the LOFAR radio astronomical array," *IEEE Transactions on Signal Processing*, vol. 55, no. 9, pp. 4497–4510, 2007.
- [53] E. Conte, A. De Maio, and G. Ricci, "Recursive estimation of the covariance matrix of a compound-Gaussian process and its application to adaptive CFAR detection," *IEEE Transactions on Signal Processing*, vol. 50, no. 8, pp. 1908–1915, 2002.
- [54] F. Pascal, Y. Chitour, J.-P. Ovarlez, P. Forster, and P. Larzabal, "Covariance structure maximum-likelihood estimates in compound Gaussian noise: Existence and algorithm analysis," *IEEE Transactions on Signal Processing*, vol. 56, no. 1, pp. 34–48, 2008.
- [55] X. Zhang, M. N. El Korso, and M. Pesavento, "Maximum likelihood and maximum a posteriori direction-of-arrival estimation in the presence of SIRP noise," in *International Conference on Acoustics, Speech and Signal Processing*, Shanghai, China, 2016.
- [56] V. Ollier, M. N. El Korso, R. Boyer, P. Larzabal, and M. Pesavento, "Relaxed concentrated MLE for robust calibration of radio interferometers," in *24th European Signal Processing Conference*, Budapest, Hungary, 2016.
- [57] A. Hjørungnes and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 2740–2746, 2007.
- [58] M. Feder and E. Weinstein, "Parameter estimation of superimposed signals using the EM algorithm," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 36, no. 4, pp. 477–489, 1988.
- [59] K. Madsen, H. B. Nielsen, and O. Tingleff, "Methods for non-linear least squares problems," Informatics and Mathematical Modelling, Technical University of Denmark, Tech. Rep., 2004.
- [60] J. Nocedal and S. J. Wright, *Numerical optimization*, 2nd ed. Springer Science & Business Media, 2006.
- [61] H. Gavin, "The Levenberg-Marquardt method for nonlinear least squares curve-fitting problems," Duke University, 2011.
- [62] T. W. Anderson, *An introduction to multivariate statistical analysis*. Wiley New York, 1958, vol. 2.
- [63] D. P. Bertsekas, *Nonlinear programming*. Athena scientific Belmont, 1999.
- [64] S. Vorobyov, A. B. Gershman, and K. M. Wong, "Maximum likelihood direction-of-arrival estimation in unknown noise fields using sparse sensor arrays," *IEEE Trans. Signal Processing*, vol. 53, no. 1, pp. 34–43, 2005.
- [65] P. Stoica and R. L. Moses, *Spectral analysis of signals*. Pearson Prentice Hall, Upper Saddle River, NJ, 2005.
- [66] A. Balleri, A. Nehorai, and J. Wang, "Maximum likelihood estimation for compound-Gaussian clutter with inverse gamma texture," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 2, pp. 775–780, 2007.
- [67] E. Ollila, D. E. Tyler, V. Koivunen, and H. Poor, "Complex elliptically symmetric distributions survey, new results and applications," *IEEE Transactions on Signal Processing*, vol. 60, no. 11, pp. 5597–25 625, 2012.
- [68] O. Besson and Y. I. Abramovich, "On the Fisher information matrix for multivariate elliptically contoured distributions," *IEEE Signal Processing Letters*, vol. 20, no. 11, pp. 1130–1133, 2013.

-
- [69] C. Liu and D. B. Rubin, "ML estimation of the t distribution using EM and its extensions, ECM and ECME," *Statistica Sinica*, vol. 5, no. 1, pp. 19–39, 1995.
- [70] S. Li, H. Wang, and T. Chai, "A t-distribution based particle filter for target tracking," in *American Control Conference*, Minneapolis, MN, 2006.