Robust calibration of radio interferometers in multi-frequency scenario

V Ollier, M N El Korso, A Ferrari, Remy Boyer, P Larzabal

ROBUST CALIBRATION OF RADIO INTERFEROMETERS IN MULTI-FREQUENCY SCENARIO

V. Ollier[⋆ ⋄] *M. N. El Korso*[†] *A. Ferrari*[‡] *R. Boyer*[⋄] *P. Larzabal*[⋆]

[⋆] SATIE, UMR 8029, ENS Paris-Saclay, Cachan, France
[⋄] L2S, UMR 8506, Université Paris-Sud, Gif-sur-Yvette, France
[†] LEME, EA 4416, Université Paris-Nanterre, Ville d'Avray, France
[‡] Laboratoire J.L. Lagrange, UMR 7293, Université Nice Sophia-Antipolis, France

ABSTRACT

This paper investigates calibration of sensor arrays in the radio astronomy context. Current and future radio telescopes require computationally efficient algorithms to overcome the new technical challenges as large collecting area, wide field of view and huge data volume. Specifically, we study the calibration of radio interferometry stations with significant direction dependent distortions. We propose an iterative robust calibration algorithm based on a relaxed maximum likelihood estimator for a specific context: i) observations are affected by the presence of outliers and ii) parameters of interest have a specific structure depending on frequency. Variation of parameters across frequency is addressed through a distributed procedure, which is consistent with the new radio synthesis arrays where the full observing bandwidth is divided into multiple frequency channels. Numerical simulations reveal that the proposed robust distributed calibration estimator outperforms the conventional non-robust algorithm and/or the mono-frequency case.

***Index Terms*—** Robust calibration, distributed optimization, consensus, direction dependent distortions regime.

1. INTRODUCTION

With a resolution and sensitivity greater than any previous systems in the low observing frequencies, the low frequency array (LOFAR) and the square kilometre array (SKA) impose major challenges in terms of telescope design and data processing [1–3]. Among these challenges, the calibration step is crucial for the advanced phased array radio telescopes due to the large number of receivers and their wide field-of-view. These design features result in direction dependent effects [4], varying over the field-of-view, due, *e.g.*, to ionospheric disturbances, and in gain differences from one receiving element to another. To avoid any calibration error preventing the exploitation of the full sensitivity potential and decreasing the high dynamic range performance of imaging [5], environmental and instrumental unknowns need to be corrected for.

In several studies [6, 7], array signal processing tools are used to perform calibration, especially the maximum likelihood (ML) estimator which is usually introduced under a Gaussian noise model assumption [8]. However, in our application, such noise modeling is not adapted: the sky model is composed of several bright known sources (the calibrator sources) but also many unknown unmodeled sources which give rise to incomplete sky models [9]. Furthermore, radio astronomical measurements are contaminated by radio inteference [10]. All these effects, and others, lead to the presence of outliers in the data set. To take them into account, a robust calibration technique was introduced in [9] where the noise model is specifically described as a Student's t with independent identically distributed entries. To improve calibration, we proposed, in a previous work [11], to adopt a compound-Gaussian modeling [12] which includes a broad range of different distributions and revealed to be more robust [13].

Furthermore, it should be noted that direction independent gains of each sensor and direction dependent perturbations associated to each source are estimated through the calibration process, where the latter parameters are frequency dependent. To exploit the known structure of variation *w.r.t.* frequency [14], calibration is reformulated as a constrained consensus problem and addressed with the alternating direction method of multipliers (ADMM) [15, 16]. To reduce the computational burden, we process the data thanks to a distributed architecture with a network of agents [17]. Decentralized and distributed strategies have already been applied for image reconstruction in radio astronomy [18] and also for calibration [19] in the Gaussian and non-structured case. In this work, we particularly focus on robust calibration of radio interferometry compact stations with direction dependent distortions, named as the 3DC calibration regime (structured case) [13].

The notation used through this paper is the following: symbols $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^H$ denote, respectively, the transpose, the complex conjugate and the Hermitian transpose. The symbol $\otimes$ represents the Kronecker product, $\mathrm{vec}(\cdot)$ stacks the columns of a matrix on top of one another, $\mathrm{diag}\{\cdot\}$ converts a vector into a diagonal matrix and the trace is given by $\mathrm{tr}\{\cdot\}$. The $B \times B$ identity matrix is referred by $\mathbf{I}_B$, $||\cdot||_F$ is the Frobenius norm, while $||\cdot||_2$ denotes the l_2 norm. Finally, j is the complex number whose square equals -1 and $[\cdot]_k$ refers to the k-th entry of the considered vector.

2. MODEL SET UP AND BACKGROUND ON ROBUST CALIBRATION ALGORITHM

2.1. Data model

We consider D signal waves emitted by calibrator sources impinging on an array of M antennas. The cross correlation between (noisy free) voltages measured by two antennas p and q, is given by [20]

$$\mathbf{V}_{pq}(\boldsymbol{\theta}) = \sum_{i=1}^{D} \mathbf{J}_{i,p}(\boldsymbol{\theta})\mathbf{C}_i\mathbf{J}_{i,q}^H(\boldsymbol{\theta}) \text{ for } p < q,\ (p,q) \in \{1,\dots,M\}^2, \tag{1}$$

This work was supported by: MAGELLAN (ANR-14-CE23-0004-01), ON FIRE project (Jeunes Chercheurs GDR-ISIS), ANR ASTRID project MARGARITA (ANR-17-ASTR-0015) and HyperFusion project.

where $\mathbf{C}_i$ is a known matrix describing the polarization state of the i-th calibrator source while $\mathbf{J}_{i,p}(\boldsymbol{\theta})$ stands for all the perturbations along the full corresponding signal path (from the i-th source to the p-th antenna) and is referred as a 2×2 Jones matrix [21]. The aim of calibration is to estimate the parameter vector $\boldsymbol{\theta}$. We rewrite (1) as

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

in which $\mathbf{u}_{i,pq}(\boldsymbol{\theta}) = \big(\mathbf{J}_{i,q}^*(\boldsymbol{\theta}) \otimes \mathbf{J}_{i,p}(\boldsymbol{\theta})\big)\, \mathbf{c}_i$ and $\mathbf{c}_i = \text{vec}(\mathbf{C}_i)$. The noisy output observation vector of the full array is $\mathbf{x} = \big[\mathbf{v}_{12}^T, \mathbf{v}_{13}^T, \dots, \mathbf{v}_{(M-1)M}^T\big]^T$ *s.t.* $\mathbf{v}_{pq} = \tilde{\mathbf{v}}_{pq}(\boldsymbol{\theta}) + \mathbf{n}_{pq}$. We note $\mathbf{n}_{pq}$ the noise sample at a particular antenna pair which takes into account background Gaussian noise but also the presence of outliers.

To deal with non-Gaussianity of the noise, we assume a compound-Gaussian noise model since it includes a broad range of heavy-tailed distributions [12, 22]. Its expression is given by

$$\mathbf{n}_{pq} = \sqrt{\tau_{pq}}\, \boldsymbol{\mu}_{pq},$$

where τ_{pq} is the positive texture variable and $\boldsymbol{\mu}_{pq} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Omega})$ is the complex speckle part.

2.2. Robust estimation of Jones matrices

The ML method is used to estimate iteratively parameters of interest $\boldsymbol{\theta}$, the covariance matrix $\boldsymbol{\Omega}$ and all realizations τ_{pq} for $p < q$, $(p, q) \in \{1, \dots, M\}^2$ which are considered unknown and deterministic in the algorithm, leading to a relaxed version of the ML [23]. Results are directly exposed in non-structured calibration algorithm (NSCA) and details can be found in [11, 13]. Let us denote $\mathbf{a}_{pq}(\boldsymbol{\theta}) = \mathbf{v}_{pq} - \tilde{\mathbf{v}}_{pq}(\boldsymbol{\theta})$, $B = \frac{M(M-1)}{2}$ is the number of antenna pairs and we impose $\text{tr}\{\boldsymbol{\Omega}\} = 1$ to remove scaling ambiguities [24]. Estimation of $\boldsymbol{\theta}$ in step 2 of NSCA can be performed numerically or thanks to expectation maximization (EM) [25] and block coordinate descent (BCD) algorithms [26]. Such vector refers here to the entries of all DM Jones matrices (non-structured case) [8, 27].

NSCA: Non-structured calibration algorithm

input : $D, M, B, \{\mathbf{C}_i\}_{i=1,\dots,D}, \mathbf{x}$
output : $\hat{\boldsymbol{\theta}}$
initialize: $\hat{\boldsymbol{\Omega}} \leftarrow \boldsymbol{\Omega}_{\text{init}}$, $\{\hat{\boldsymbol{\tau}}_{pq} \leftarrow \boldsymbol{\tau}_{pq_{\text{init}}}\}_{p<q,(p,q)\in\{1,\dots,M\}^2}$

1 **while** *stop criterion unreached* **do**

2 $\quad \hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}}{\text{argmin}} \left\{ \sum_{pq} \frac{1}{\hat{\tau}_{pq}} \mathbf{a}_{pq}^H(\boldsymbol{\theta}) \hat{\boldsymbol{\Omega}}^{-1} \mathbf{a}_{pq}(\boldsymbol{\theta}) \right\}$

3 $\quad \hat{\boldsymbol{\Omega}} = \frac{4}{B} \sum_{pq} \frac{\mathbf{a}_{pq}(\hat{\boldsymbol{\theta}}) \mathbf{a}_{pq}^H(\hat{\boldsymbol{\theta}})}{\mathbf{a}_{pq}^H(\hat{\boldsymbol{\theta}}) (\hat{\boldsymbol{\Omega}})^{-1} \mathbf{a}_{pq}(\hat{\boldsymbol{\theta}})}$ and $\hat{\boldsymbol{\Omega}} = \frac{\hat{\boldsymbol{\Omega}}}{\text{tr}\{\hat{\boldsymbol{\Omega}}\}}$

4 $\quad \{\hat{\tau}_{pq} = \frac{1}{4} \mathbf{a}_{pq}^H(\hat{\boldsymbol{\theta}}) \hat{\boldsymbol{\Omega}}^{-1} \mathbf{a}_{pq}(\hat{\boldsymbol{\theta}})\}_{p<q,(p,q)\in\{1,\dots,M\}^2}$

5 **end**

3. CALIBRATION IN 3DC REGIME

3.1. 3DC calibration regime

We consider calibration of radio interferometry stations where antennas are enclosed in a compact array and direction dependent effects are dominant [28]. If we intend to calibrate the data at different frequencies, then a particular decomposition of the Jones matrix is [29, 30]

$$\mathbf{J}_{i,p}^{[f]}(\boldsymbol{\theta}_{i,p}^{[f]}) = \mathbf{G}_p(\mathbf{g}_p) \mathbf{H}_{i,p}^{[f]} \mathbf{Z}_{i,p}^{[f]}(\boldsymbol{\alpha}_i^{[f]}) \mathbf{F}_i^{[f]}(\vartheta_i^{[f]}) \tag{2}$$

for $i \in \{1, \dots, D\}$, $p \in \{1, \dots, M\}$, $f \in \mathcal{F} = \{f_1, \dots, f_F\}$ and $\boldsymbol{\theta}_{i,p}^{[f]} = [\vartheta_i^{[f]}, \mathbf{g}_p^T, \boldsymbol{\alpha}_i^{[f]^T}]^T$. Each individual Jones term and its corresponding effect is described in the following:

3.1.1. Ionospheric effects

While travelling through a charged medium as the ionosphere, the incoming wave is affected by a phase delay due to refraction, and written as [31]

$$\mathbf{Z}_{i,p}^{[f]}(\boldsymbol{\alpha}_i^{[f]}) = \exp\Big(j\varphi_{i,p}^{[f]}\Big) \mathbf{I}_2$$

where $\varphi_{i,p}^{[f]} = \eta_i^{[f]} u_p^{[f]} + \zeta_i^{[f]} v_p^{[f]}$ in which $\boldsymbol{\alpha}_i^{[f]} = [\eta_i^{[f]}, \zeta_i^{[f]}]^T \propto f^{-2}$ [32] stands for the apparent position shift in the source location and $\mathbf{r}_p^{[f]} = [u_p^{[f]}, v_p^{[f]}]^T$ is the known position vector of the p-th antenna in wavelength units. Therefore, we deduce that $\varphi_{i,p}^{[f]} \propto f^{-1}$ [33]. Propagation through the upper part of the atmosphere also results in a rotation of the polarization plane, called Faraday rotation [34],

$$\mathbf{F}_i^{[f]}(\vartheta_i^{[f]}) = \begin{bmatrix} \cos(\vartheta_i^{[f]}) & -\sin(\vartheta_i^{[f]}) \\ \sin(\vartheta_i^{[f]}) & \cos(\vartheta_i^{[f]}) \end{bmatrix}$$

where the Faraday rotation angle $\vartheta_i^{[f]} \propto f^{-2}$ [29] is the same for all antennas due to the compact geometry of the array [28].

3.1.2. Instrumental effects

Receiver electronics introduce direction and frequency independent effects [29], leading to a complex sensor gain matrix, noted $\mathbf{G}_p(\mathbf{g}_p) = \text{diag}\{\mathbf{g}_p\}$.

Finally, let us note that $\mathbf{H}_{i,p}^{[f]}$ is a known matrix given by electromagnetic modeling and a priori information (calibrator sources and antenna positions) [21]. Let us recall that the combined effect of all individual perturbations along a particular signal path is represented by a global Jones matrix, as shown in (2).

3.2. Mono-frequency case

In 3DC regime, calibration amounts to estimate the physical parameters $\boldsymbol{\theta}_{i,p}$ which appear in the structured Jones matrices (2). In the mono-frequency case, we use an alternating least squares approach to estimate unknowns sequentially [35], by optimizing a cost function *w.r.t.* one parameter while fixing the others. The global frame of the structured calibration algorithm (SCA) is exposed hereafter, for one given i and p, and details can be found in [13]. Let us note that $\hat{\boldsymbol{\varphi}}_i = [\hat{\varphi}_{i,1}, \dots, \hat{\varphi}_{i,M}]^T$ and $\boldsymbol{\Lambda} = [\mathbf{r}_1, \dots, \mathbf{r}_M]$.

3.3. Multi-frequency calibration algorithm

The aim of calibration in a multi-frequency scenario is to estimate the parameter vector of interest $\boldsymbol{\epsilon} = [\boldsymbol{\epsilon}^{[f_1]^T}, \dots, \boldsymbol{\epsilon}^{[f_F]^T}, \mathbf{g}^T]^T$ where $\boldsymbol{\epsilon}^{[f]} = [\vartheta_1^{[f]}, \dots, \vartheta_D^{[f]}, \boldsymbol{\alpha}_1^{[f]^T}, \dots, \boldsymbol{\alpha}_D^{[f]^T}]^T$ and $\mathbf{g} = [\mathbf{g}_1^T, \dots, \mathbf{g}_M^T]^T$. To do so, we introduce the following cost function

$$l^{[f]}(\boldsymbol{\epsilon}^{[f]}) = \sum_{i=1}^{D} l_i^{[f]}(\boldsymbol{\epsilon}_i^{[f]}) \tag{3}$$

SCA: Structured calibration algorithm

input : D, M, B, $\mathbf{C}_i$, $\mathbf{x}$, $\hat{\mathbf{J}}_{i,p}$ as output of NSCA

output : $\hat{\boldsymbol{\theta}}_{i,p}$

initialize: $\hat{\boldsymbol{\theta}}_{i,p} \leftarrow \boldsymbol{\theta}_{i,p_{\text{init}}}$

while *stop criterion unreached* **do**

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

$\hat{\mathbf{g}}_p = \operatorname{argmin}_{\mathbf{g}_p} \sum_{i=1}^{D} ||\hat{\mathbf{J}}_{i,p} - \mathbf{G}_p(\mathbf{g}_p) \mathbf{H}_{i,p} \mathbf{Z}_{i,p} \mathbf{F}_i||_F^2$

$\hat{\varphi}_{i,p} = \operatorname{argmin}_{\varphi_{i,p}} ||\hat{\mathbf{J}}_{i,p} - \mathbf{G}_p \mathbf{H}_{i,p} \mathbf{Z}_{i,p}(\varphi_{i,p}) \mathbf{F}_i||_F^2$

$\hat{\boldsymbol{\alpha}}_i^T = \dfrac{\hat{\boldsymbol{\varphi}}_i^T \boldsymbol{\Lambda}^H \begin{bmatrix} \sum_{p=1}^M v_p^2 & -\sum_{p=1}^M u_p v_p \\ -\sum_{p=1}^M v_p u_p & \sum_{p=1}^M u_p^2 \end{bmatrix}}{\sum_{p=1}^M u_p^2 \sum_{p=1}^M v_p^2 - (\sum_{p=1}^M u_p v_p)^2}$

end

in which $l_i^{[f]}(\boldsymbol{\epsilon}_i^{[f]}) = \sum_{p=1}^{M} ||\hat{\mathbf{J}}_{i,p}^{[f]} - \mathbf{G}_p \mathbf{H}_{i,p}^{[f]} \mathbf{Z}_{i_p}^{[f]}(\boldsymbol{\alpha}_i^{[f]}) \mathbf{F}_i^{[f]}(\vartheta_i^{[f]})||_F^2$ with $\boldsymbol{\epsilon}_i^{[f]} = [\vartheta_i^{[f]}, \boldsymbol{\alpha}_i^{[f]^T}]^T$. Prior information on $\hat{\mathbf{J}}_{i,p}^{[f]}$ for $i \in \{1,\ldots,D\}$, $p \in \{1,\ldots,M\}$ and $f \in \mathcal{F}$ is provided by the output of NSCA.

We wish to distributedly solve the following constrained optimization problem thanks to a network of agents

$$\{\hat{\boldsymbol{\epsilon}}^{[f]}\}_{f\in\mathcal{F}}, \hat{\mathbf{z}} = \underset{\boldsymbol{\epsilon}^{[f_1]},\ldots,\boldsymbol{\epsilon}^{[f_F]},\mathbf{z}}{\operatorname{argmin}} \sum_{f\in\mathcal{F}} l^{[f]}(\boldsymbol{\epsilon}^{[f]}) \tag{4}$$
$$\textit{s.t.}\;\; \boldsymbol{\epsilon}_i^{[f]} = \mathbf{B}^{[f]} \mathbf{z}_i, \;\; i \in \{1,\ldots,D\}, f \in \mathcal{F}$$

where $\mathbf{B}^{[f]} = \frac{1}{f^2}\mathbf{I}_3$ is the known frequency model, $\mathbf{z}_i$ is an unknown associated global variable, independent *w.r.t.* frequency and shared by all agents, and $\mathbf{z} = [\mathbf{z}_1^T, \ldots, \mathbf{z}_D^T]^T$. To solve (4), we use a consensus optimization scheme as in the ADMM procedure [16]. Instead of considering the original objective function (3), we study the following augmented Lagrangian

$$L(\boldsymbol{\epsilon}^{[f_1]}, \ldots, \boldsymbol{\epsilon}^{[f_F]}, \mathbf{z}, \mathbf{y}^{[f_1]}, \ldots, \mathbf{y}^{[f_F]}) = \sum_{f\in\mathcal{F}} \sum_{i=1}^{D} L_i^{[f]}\left(\boldsymbol{\epsilon}_i^{[f]}, \mathbf{z}_i, \mathbf{y}_i^{[f]}\right)$$

where $L_i^{[f]}\left(\boldsymbol{\epsilon}_i^{[f]}, \mathbf{z}_i, \mathbf{y}_i^{[f]}\right) = l_i^{[f]}\left(\boldsymbol{\epsilon}_i^{[f]}\right) + h_i^{[f]}\left(\boldsymbol{\epsilon}_i^{[f]}, \mathbf{z}_i, \mathbf{y}_i^{[f]}\right)$ and

$$h_i^{[f]}\left(\boldsymbol{\epsilon}_i^{[f]}, \mathbf{z}_i, \mathbf{y}_i^{[f]}\right) = \mathbf{y}_i^{[f]^T}\left(\boldsymbol{\epsilon}_i^{[f]} - \mathbf{B}^{[f]}\mathbf{z}_i\right) + \frac{\rho}{2}||\boldsymbol{\epsilon}_i^{[f]} - \mathbf{B}^{[f]}\mathbf{z}_i||_2^2.$$

We note $\mathbf{y}^{[f]} = [\mathbf{y}_1^{[f]^T}, \ldots, \mathbf{y}_D^{[f]^T}]^T$ the associated Lagrange parameters (or dual variables) and $\rho > 0$ a penalty factor. We notice separability of the Lagrangian *w.r.t.* source direction but above all, separability *w.r.t.* frequency, meaning that each agent solves a subproblem locally at a given frequency. The ADMM consists in updating sequentially the three following quantities:

- $$\left(\hat{\boldsymbol{\epsilon}}_i^{[f]}\right)^{t+1} = \underset{\boldsymbol{\epsilon}_i^{[f]}}{\operatorname{argmin}}\, L_i^{[f]}\left(\boldsymbol{\epsilon}_i^{[f]}, (\hat{\mathbf{z}}_i)^t, \left(\hat{\mathbf{y}}_i^{[f]}\right)^t\right) \tag{5}$$
  performed locally by each agent for $i \in \{1, \ldots, D\}$

- $$(\hat{\mathbf{z}}_i)^{t+1} = \underset{\mathbf{z}_i}{\operatorname{argmin}} \sum_{f\in\mathcal{F}} L_i^{[f]}\left(\left(\hat{\boldsymbol{\epsilon}}_i^{[f]}\right)^{t+1}, \mathbf{z}_i, \left(\hat{\mathbf{y}}_i^{[f]}\right)^t\right) \tag{6}$$
  performed globally for $i \in \{1, \ldots, D\}$

- $$\left(\hat{\mathbf{y}}_i^{[f]}\right)^{t+1} = \left(\hat{\mathbf{y}}_i^{[f]}\right)^t + \rho\left(\left(\hat{\boldsymbol{\epsilon}}_i^{[f]}\right)^{t+1} - \mathbf{B}^{[f]}(\hat{\mathbf{z}}_i)^{t+1}\right) \tag{7}$$
  performed locally by each agent for $i \in \{1, \ldots, D\}$

where t is the iteration counter. Minimization (6) needs access to local solutions from all agents, *i.e.*, at all frequencies, and leads to the following closed-form expression [16, 19],

$$\hat{\mathbf{z}}_i = \left(\sum_{f\in\mathcal{F}} \rho \mathbf{B}^{[f]^T}\mathbf{B}^{[f]}\right)^{-1}\left(\sum_{f\in\mathcal{F}} \mathbf{B}^{[f]^T}(\mathbf{y}_i^{[f]} + \rho\boldsymbol{\epsilon}_i^{[f]})\right). \tag{8}$$

Minimization (5) is addressed iteratively. To this end, we compute the gradient of $L_i^{[f]}(\boldsymbol{\epsilon}_i^{[f]}, \mathbf{z}_i, \mathbf{y}_i^{[f]})$ *w.r.t.* $\boldsymbol{\epsilon}_i^{[f]}$, which induces

$$\frac{\partial L_i^{[f]}(\boldsymbol{\epsilon}_i^{[f]}, \mathbf{z}_i, \mathbf{y}_i^{[f]})}{\partial \vartheta_i^{[f]}} = [\mathbf{y}_i^{[f]}]_1 + \rho(\vartheta_i^{[f]} - \frac{1}{f^2}[\mathbf{z}_i]_1) + \sum_{p=1}^{M} \operatorname{tr}\left\{\mathbf{S}_{i,p}^{[f]} + \mathbf{S}_{i,p}^{[f]^H}\right\} \tag{9}$$

where $\mathbf{S}_{i,p}^{[f]} = -\mathbf{G}_p \mathbf{H}_{i,p}^{[f]} \mathbf{Z}_{i,p}^{[f]} \frac{\partial \mathbf{F}_i^{[f]}(\vartheta_i^{[f]})}{\partial \vartheta_i^{[f]}} \hat{\mathbf{J}}_{i,p}^{[f]^H}$, and

$$\frac{\partial L_i^{[f]}(\boldsymbol{\epsilon}_i^{[f]}, \mathbf{z}_i, \mathbf{y}_i^{[f]})}{\partial \eta_i^{[f]}} = [\mathbf{y}_i^{[f]}]_2 + \rho(\eta_i^{[f]} - \frac{1}{f^2}[\mathbf{z}_i]_2) + \sum_{p=1}^{M} \operatorname{tr}\left\{\mathbf{D}_{i,p}^{[f]} + \mathbf{D}_{i,p}^{[f]^H}\right\} \tag{10}$$

where $\mathbf{D}_{i,p}^{[f]} = j u_p^{[f]} \mathbf{Z}_{i,p}^{[f]^*} \mathbf{M}_{i,p}^{[f]}$ and $\mathbf{M}_{i,p}^{[f]} = \hat{\mathbf{J}}_{i,p}^{[f]} \mathbf{F}_i^{[f]^T} \mathbf{H}_{i,p}^{[f]^H} \mathbf{G}_p^H$. Likewise, we have

$$\frac{\partial L_i^{[f]}(\boldsymbol{\epsilon}_i^{[f]}, \mathbf{z}_i, \mathbf{y}_i^{[f]})}{\partial \zeta_i^{[f]}} = [\mathbf{y}_i^{[f]}]_3 + \rho(\zeta_i^{[f]} - \frac{1}{f^2}[\mathbf{z}_i]_3) + \sum_{p=1}^{M} \operatorname{tr}\left\{\mathbf{V}_{i,p}^{[f]} + \mathbf{V}_{i,p}^{[f]^H}\right\} \tag{11}$$

where $\mathbf{V}_{i,p}^{[f]} = j v_p^{[f]} \mathbf{Z}_{i,p}^{[f]^*} \mathbf{M}_{i,p}^{[f]}$. Using eqs. (9) to (11), we obtain $\hat{\boldsymbol{\epsilon}}_i^{[f]}$ with a root-finding algorithm or thanks to standard numerical optimization tools as Newton or gradient descent-type algorithm [36].

Estimation of the gains $\mathbf{g}_p$ is done as minimization of the following least squares cost function

$$\kappa(\mathbf{g}_p) = \sum_{f\in\mathcal{F}} \sum_{i=1}^{D} ||\hat{\mathbf{J}}_{i,p}^{[f]} - \mathbf{G}_p(\mathbf{g}_p) \mathbf{H}_{i,p}^{[f]} \mathbf{Z}_{i,p}^{[f]} \mathbf{F}_i^{[f]}||_F^2$$

leading to the following estimate of each complex gain element for $k \in \{1, 2\}$

$$[\hat{\mathbf{g}}_p]_k = \left(\sum_{f\in\mathcal{F}} \sum_{i=1}^{D} [\mathbf{W}_{i,p}^{[f]^*}]_{k,k}\right)^{-1} \sum_{f\in\mathcal{F}} \sum_{i=1}^{D} [\mathbf{X}_{i,p}^{[f]^*}]_{k,k} \tag{12}$$

where $\mathbf{X}_{i,p}^{[f]} = \mathbf{R}_{i,p}^{[f]} \hat{\mathbf{J}}_{i,p}^{[f]^H}$ and $\mathbf{W}_{i,p}^{[f]} = \mathbf{R}_{i,p}^{[f]} \mathbf{R}_{i,p}^{[f]^H}$ in which $\mathbf{R}_{i,p}^{[f]} = \mathbf{H}_{i,p}^{[f]} \mathbf{Z}_{i,p}^{[f]} \mathbf{F}_i^{[f]}$. The scheme of the proposed multi-frequency structured calibration algorithm (MSCA) is described below.

4. NUMERICAL SIMULATIONS

In this section, we compare the performance of the proposed MSCA with the mono-frequency case, *i.e.*, SCA in which frequency diversity is not taken into account. We recall that radio astronomy observations are affected by the presence of outliers. Thus, we also compare our robust approach with an algorithm based on a classical Gaussian noise assumption [8], which amounts to solve a non-linear least squares problem. First, in Fig. 1, we plot the mean square error (MSE) of $\eta_1^{[f_1]}$ as a function of the signal-to-noise ratio (SNR), the

MSCA: Multi-frequency structured calibration algorithm

input : D, M, F, $\mathbf{r}_p^{[f]}$, $\mathbf{H}_{i,p}^{[f]}$, $\hat{\mathbf{J}}_{i,p}^{[f]}$ as output of NSCA for $i \in \{1, \dots, D\}$, $p \in \{1, \dots, M\}$ and $f \in \mathcal{F}$

output : $\hat{\boldsymbol{\epsilon}}$

initialize: $\hat{\boldsymbol{\epsilon}} \leftarrow \boldsymbol{\epsilon}_{\text{init}}$, $\hat{\mathbf{z}} \leftarrow \mathbf{z}_{\text{init}}$, $\{\hat{\mathbf{y}}^{[f]} \leftarrow \mathbf{y}_{\text{init}}^{[f]}\}_{f \in \mathcal{F}}$

while *stop criterion unreached* **do**
  while *stop criterion unreached* **do**
    while *stop criterion unreached* **do**
      Obtain $\{\hat{\vartheta}_i^{[f]}\}_{i=1,\dots,D}$ locally with (9)
      Obtain $\{\hat{\eta}_i^{[f]}\}_{i=1,\dots,D}$ locally with (10)
      Obtain $\{\hat{\zeta}_i^{[f]}\}_{i=1,\dots,D}$ locally with (11)
    end
    Obtain $\{\hat{\mathbf{z}}_i\}_{i=1,\dots,D}$ globally with (8)
    Obtain $\{\hat{\mathbf{y}}_i^{[f]}\}_{i=1,\dots,D}$ locally with (7)
  end
  Obtain $\{\hat{\mathbf{g}}_p\}_{p=1,\dots,M}$ with (12)
end

behavior being the same for any other parameter of $\boldsymbol{\epsilon}$. We compare the estimation performance for different number of frequencies F and notice better statistical performance when multi-frequency robust calibration is performed. Robust calibration based on the Student's t [9] is not exposed in the simulations due to a different model which is not adapted to the 3DC regime.

In the following figures, we use Meqtrees [29] to generate the data model, the observations and compare its least squares solver to MSCA. Here, we choose to correct for Faraday rotation matrices, which are the only introduced perturbations in the observations. We consider $M = 7$ antennas (KAT-7 instrument), $D = 1$, $D' = 16$ weak realistic background sources taken from the SUMSS survey using a spectral index of 0.7. The full duration of the observation is 12 hours, for 60 seconds integration time per data sample. After calibration and subtraction of the bright calibrator source, a dirty image, namely the corrected residual, is constructed with Meqtrees using *lwimager*. Fig. 2 gives the corrected residual image at 895 MHz in a small area surrounding the calibrator, whose position corresponds to the red cross. Fig. 3 gives the recovered flux for one of the D' weak outlier sources. Therefore, we notice better flux estimation of weak background sources and better calibrator removal using joint frequency dependent calibration with MSCA compared to a frequency independent calibration.

5. CONCLUSION

This paper introduces a robust ML based calibration technique in the context of radio interferometry. Robustness is addressed thanks to a compound-Gaussian noise modeling and a particular scenario is studied, *i.e.*, the 3DC calibration regime. Variation of parameters *w.r.t.* frequency imposes additional constraints which are considered in an ADMM-based distributed algorithm. We show in the simulations the advantages of the proposed algorithm in regards to standard mono-frequency and/or non-robust scenario.

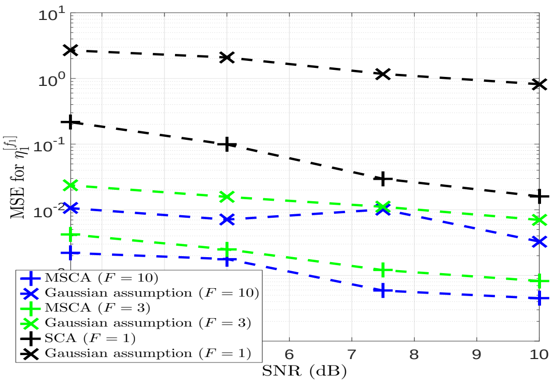
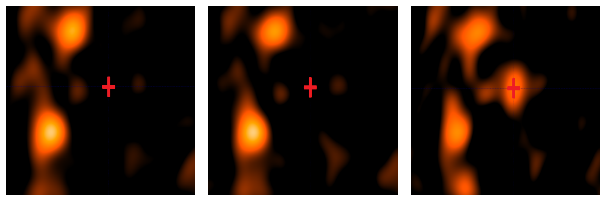


Fig. 1. MSE of $\eta_1^{[f_1]}$ vs. SNR, for $D = 2$ bright signal sources, $M = 8$ antennas and 4 weak outlier sources.



(a) (b) (c)

Fig. 2. Corrected residual for (a) ideal case, (b) MSCA and (c) Meqtrees solver at 895 MHz. The red cross corresponds to the location of the calibrator source.

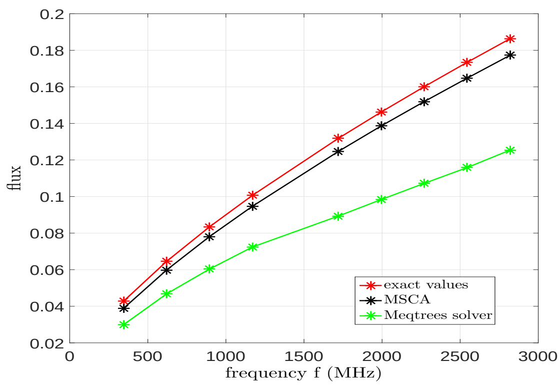


Fig. 3. Recovered flux of one weak outlier source after calibration for $F = 9$.

6. REFERENCES

[1] 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 (IEEE ICASSP)*, Florence, Italy, 2014, pp. 5382–5386.

[2] C. Ferrari *et al.*, "French SKA White Book - The French community towards the Square Kilometre Array," *Published by the SKA-France Coordination in collaboration with AS SKA-LOFAR, arXiv:1712.06950v2*, 2018.

[3] P. Zarka *et al.*, "NenUFAR: Instrument description and science case," in *International Conference on Antenna Theory and Techniques (IEEE ICATT)*, Kharkiv, Ukraine, 2015.

[4] B. D. Jeffs, S. van der Tol, and A.-J. van der Veen, "Direction

dependent self calibration of large distributed sensor arrays," in *IEEE International Conference on Acoustics, Speech and Signal Processing (IEEE ICASSP)*, Toulouse, France, 2006, pp. IV – 1069–IV – 1072.

[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] 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.

[7] 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.

[8] 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.

[9] S. Yatawatta and S. Kazemi, "Robust radio interferometric calibration," in *International Conference on Acoustics, Speech and Signal Processing (IEEE ICASSP)*, Florence, Italy, 2014, pp. 5392–5396.

[10] 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.

[11] 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 (EUSIPCO)*, Budapest, Hungary, 2016, pp. 280–284.

[12] 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. 5597–5625, 2012.

[13] V. Ollier, M. N. El Korso, R. Boyer, P. Larzabal, and M. Pesavento, "Robust calibration of radio interferometers in non-Gaussian environment," *IEEE Transactions on Signal Processing*, vol. 65, no. 21, pp. 5649–5660, 2017.

[14] M. Brossard, M. N. El Korso, M. Pesavento, R. Boyer, P. Larzabal, and S. J. Wijnholds, "Parallel multi-wavelength calibration algorithm for radio astronomical arrays," *Elsevier Signal Processing Journal*, vol. 145, pp. 258–271, 2018.

[15] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and distributed computation: numerical methods*. Prentice Hall, 1989.

[16] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[17] T.-H. Chang, M. Hong, and X. Wang, "Multi-agent distributed optimization via inexact consensus ADMM," *IEEE Transactions on Signal Processing*, vol. 63, no. 2, pp. 482–497, 2015.

[18] A. Ferrari, D. Mary, R. Flamary, and C. Richard, "Distributed image reconstruction for very large arrays in radio astronomy," in *8th Sensor Array and Multichannel Signal Processing Workshop (IEEE SAM)*, 2014, pp. 389–392.

[19] S. Yatawatta, "Distributed radio interferometric calibration," *Monthly Notices of the Royal Astronomical Society*, vol. 449, no. 4, pp. 4506–4514, 2015.

[20] L. M. Ker, "Radio AGN evolution with low frequency radio surveys," Ph.D. dissertation, The University of Edinburgh, 2012.

[21] A. R. Thompson, J. M. Moran, and G. W. Swenson Jr., *Interferometry and synthesis in radio astronomy, 2nd ed.* New York: John Wiley & Sons, 2001.

[22] J. Wang, A. Dogandzic, 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.

[23] 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.

[24] 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.

[25] G. J. McLachlan and T. Krishnan, *The EM algorithm and extensions, 2nd ed.* John Wiley & Sons, 2008.

[26] 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 Processing Magazine*, vol. 33, no. 1, pp. 57–77, 2016.

[27] C. D. Nunhokee, "Link between ghost artefacts, source suppression and incomplete calibration sky models," Ph.D. dissertation, Rhodes University, Rhodes, South Africa, 2015.

[28] 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.

[29] 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, p. A61, 2010.

[30] S. Yatawatta, "Reduced ambiguity calibration for LOFAR," *Experimental Astronomy*, vol. 34, no. 1, pp. 89–103, 2012.

[31] O. M. Smirnov, "Revisiting the radio interferometer measurement equation. II. Calibration and direction-dependent effects," *Astronomy & Astrophysics*, vol. 527, p. A107, 2011.

[32] A. Cohen and H. Röttgering, "Probing fine-scale ionospheric structure with the Very Large Array radio telescope," *The Astronomical Journal*, vol. 138, no. 2, pp. 439–447, 2009.

[33] S. Van der Tol, "Bayesian estimation for ionospheric calibration in radio astronomy," Ph.D. dissertation, TU Delft, Delft University of Technology, 2009.

[34] K. Davies, *Ionospheric radio*. London: Peter Peregrinus, 1990.

[35] S. J. Wijnholds, "Fish-eye observing with phased array radio telescopes," Ph.D. dissertation, Technische Universiteit Delft, Delft, The Netherlands, 2010.

[36] J. Nocedal and S. J. Wright, *Numerical optimization, 2nd ed.* Springer Science & Business Media, 2006.