Robust distributed calibration of radio interferometers with direction dependent distortions

Virginie Ollier, El Korso, André Ferrari, Remy Boyer, Pascal Larzabal

Robust distributed calibration of radio interferometers with direction dependent distortions

Virginie OLLIER[a,b,*], Mohammed Nabil EL KORSO[c], André FERRARI[d], Rémy BOYER[b], Pascal LARZABAL[a]

[a]*ENS Paris-Saclay, SATIE UMR 8029, 61 avenue du Président Wilson, 94235 Cachan, France*
[b]*Paris-Sud University, L2S UMR 8506, 3 rue Joliot-Curie, 91192 Gif-sur-Yvette, France*
[c]*Paris-Nanterre University, LEME EA 4416, 50 rue de Sèvres, 92410 Ville d'Avray, France*
[d]*Nice Sophia-Antipolis University, Lab. J.L. Lagrange, UMR 7293, Parc Valrose, 06108 Nice, France*

Abstract

In radio astronomy, accurate calibration is of crucial importance for the new generation of radio interferometers. More specifically, because of the potential presence of outliers which affect the measured data, robustness needs to be ensured. On the other hand, calibration is improved by taking advantage of these new instruments and exploiting the known structure of parameters of interest across frequency. Therefore, we propose in this paper an iterative robust multi-frequency calibration algorithm based on a distributed and consensus optimization scheme which aims to estimate the complex gains of the receivers and the directional perturbations caused by the ionosphere. Numerical simulations reveal that the proposed distributed calibration technique outperforms the conventional non-robust algorithm and per-channel calibration.

Keywords: Calibration, robustness, multi-frequency, radio astronomy.

1. Introduction

In radio interferometry, the easiest approach to perform calibration is to consider one frequency bin at a time [1, 2], with a single centralized agent processing the data, leading to suboptimality and computational limitations. However, considering the specific variation of parameters across frequency enables to add some modeling structure on the unknown parameters to estimate and therefore perform more accurately direction dependent calibration [3], notably for the new generation of radio interferometers where multiple frequency sub-bands are present [4]. Exploiting the frequency-dependent response has already been studied in the context of bandpass solutions [5], assuming a smooth response in frequency and also for direction dependent calibration with smooth polynomials [6, 7]. As direction dependent perturbation effects are particularly

significant in observations with new instruments, we focus in this paper on the regime where the receiving elements of the array have a large field-of-view and possibly long baselines, resulting in a direction dependent calibration problem in which receptors see different parts of the ionosphere [2]. We call this the direction dependent distortion regime.

On the one hand, multi-frequency calibration aims to take into account a whole frequency range, *e.g.*, between 30 and 240 MHz for the nominal operating bandpass of the Low Frequency Array (LOFAR) [8, 9]. A computationally efficient way to handle the multiple sub-frequency bands in radio astronomy is to apply distributed and consensus algorithms with a decentralized strategy. In this context, we apply the Alternating Direction Method of Multipliers (ADMM) [10, 6], which is well-suited for large-scale problems as in radio interferometry. This technique is based on decomposition and coordination tasks with a group of computational agents. Each of them has access to a part of the data and finds a solution to a local subproblem, in a restricted frequency interval. Communication with a fusion center enforces consensus among all agents, the goal being to solve a global constrained optimization problem. By collecting and storing data in a distributed way among different computational agents, the global operational cost in the network is substantially reduced.

On the other hand, measurements are frequently affected by the presence of outliers due to interference or weak background sources. As a consequence, the noise can no longer be considered Gaussian as in [6]. In order to propose an alternative to [11], where the noise model is based on the Student's t distribution, in this work, we make use of a compound-Gaussian model which includes heavy-tailed distributions [12]. This is more general and reveals to be more robust [13]. Therefore, to take into account the presence of outliers and the variability of parameters across frequency, we propose here a Multi-frequency Robust Calibration Algorithm (MRCA) based on a compound-Gaussian distribution for modeling the noise and the consensus ADMM approach [10]. As in [14], we aim at estimating physical parameters appearing in the Jones terms [15, 3].

In this paper, we use the following notation: The trace and determinant operators are, respectively, given by $\text{tr}\{\cdot\}$ and $|\cdot|$. The $B \times B$ identity matrix is referred by $\mathbf{I}_B$ and $||\cdot||_2$ denotes the l_2 norm. The symbol $\otimes$ represents the Kronecker product, $\text{vec}(\cdot)$ stacks the columns of a matrix on top of one another, $\text{diag}\{\cdot\}$ converts a vector into a diagonal matrix while $\text{bdiag}\{\cdot\}$ is the block-diagonal operator. Finally, $\Re\{\cdot\}$ and $\Im\{\cdot\}$ are, respectively, the real and imaginary parts, $[\cdot]_k$ refers to the k-th entry of the considered vector, $\arg\{\cdot\}$ is the argument of a complex number and j is the complex number whose square equals -1.

2. Model setup

2.1. Direction dependent distortion regime

An interferometer output consists of visibilities, *i.e.*, correlations of signals measured by two array elements along the corresponding baseline vector. In the noise free case, the measurements are given, for each frequency $f \in \mathcal{F} = \{f_1, \ldots, f_F\}$, by

$$\mathbf{S}_{pq}^{[f]}(\boldsymbol{\theta}^{[f]}) = \sum_{i=1}^{D} \mathbf{J}_{i,p}^{[f]}(\boldsymbol{\theta}^{[f]}) \mathbf{C}_i^{[f]} \mathbf{J}_{i,q}^{[f]^H}(\boldsymbol{\theta}^{[f]}) \tag{1}$$

where $(p,q) \in \{1, \ldots, M\}^2$ is a pair of receivers with $p < q$, M denotes the total number of array elements and D the number of calibrator sources, while $\mathbf{J}_{i,p}^{[f]}(\boldsymbol{\theta}^{[f]})$ refers to the so-called 2×2 Jones matrix [15, 3], accounting for the perturbation effects along the path from the i-th source to the p-th receiver. We note $\boldsymbol{\theta}^{[f]}$ the unknown parameter vector of interest, whose elements are detailed further below. Finally, $\mathbf{C}_i^{[f]}$ is the known source brightness matrix of the i-th calibrator source, describing its polarization state. After vectorization of measurements (1) and considering the noise effect, we obtain the 4×1 data vector $\mathbf{x}_{pq}^{[f]} = \mathrm{vec}\left(\mathbf{S}_{pq}^{[f]}(\boldsymbol{\theta}^{[f]})\right) + \mathbf{n}_{pq}^{[f]}$ where $\mathbf{n}_{pq}^{[f]}$ is the noise vector for baseline (p,q) and $\mathrm{vec}\left(\mathbf{S}_{pq}^{[f]}(\boldsymbol{\theta}^{[f]})\right) = \sum_{i=1}^{D} \mathbf{s}_{i,pq}^{[f]}(\boldsymbol{\theta}^{[f]})$ in which $\mathbf{s}_{i,pq}^{[f]}(\boldsymbol{\theta}^{[f]}) = \left(\mathbf{J}_{i,q}^{[f]^*}(\boldsymbol{\theta}^{[f]}) \otimes \mathbf{J}_{i,p}^{[f]}(\boldsymbol{\theta}^{[f]})\right) \mathbf{c}_i^{[f]}$ and $\mathbf{c}_i^{[f]} = \mathrm{vec}(\mathbf{C}_i^{[f]})$. Finally, the $4B \times 1$ full measurement vector at frequency f, with $B = \frac{M(M-1)}{2}$ the total number of baselines, reads

$$\mathbf{x}^{[f]} = \left[\mathbf{x}_{12}^{[f]^T}, \ldots, \mathbf{x}_{(M-1)M}^{[f]^T}\right]^T = \sum_{i=1}^{D}\left[\mathbf{s}_{i,12}^{[f]^T}(\boldsymbol{\theta}^{[f]}), \ldots, \mathbf{s}_{i,(M-1)M}^{[f]^T}(\boldsymbol{\theta}^{[f]})\right]^T + \left[\mathbf{n}_{12}^{[f]^T}, \ldots, \mathbf{n}_{(M-1)M}^{[f]^T}\right]^T. \tag{2}$$

In the direction dependent distortion regime, a particular decomposition of the Jones matrix is given by [15, 16]

$$\mathbf{J}_{i,p}^{[f]}(\boldsymbol{\theta}^{[f]}) = \mathbf{G}_p^{[f]}(\mathbf{g}_p^{[f]}) \mathbf{H}_{i,p}^{[f]} \mathbf{Z}_{i,p}^{[f]}(\varphi_{i,p}^{[f]}) \mathbf{F}_{i,p}^{[f]}(\vartheta_{i,p}^{[f]}). \tag{3}$$

Specifically, the complex electronic gain matrix is represented by $\mathbf{G}_p^{[f]}(\mathbf{g}_p^{[f]}) = \mathrm{diag}\{\mathbf{g}_p^{[f]}\}$ while $\mathbf{H}_{i,p}^{[f]}$ is an assumed known matrix gathering the geometric delay and beam pattern [15, 3]. In addition, propagation through the ionosphere induces two effects. The first one is a phase delay given by the matrix $\mathbf{Z}_{i,p}^{[f]}(\varphi_{i,p}^{[f]})$ [3], and written as

$$\mathbf{Z}_{i,p}^{[f]}(\varphi_{i,p}^{[f]}) = \exp\left(j\varphi_{i,p}^{[f]}\right)\mathbf{I}_2 \tag{4}$$

where $\varphi_{i,p}^{[f]} \propto \mathrm{TEC}_{i,p}/f$ [17], with $\mathrm{TEC}_{i,p}$ the Total Electron Content defined as the integrated electron density along line of sight i-p. The second effect is the so-called Faraday rotation [16] in (3), given by

$$\mathbf{F}_{i,p}^{[f]}(\vartheta_{i,p}^{[f]}) = \begin{bmatrix} \cos(\vartheta_{i,p}^{[f]}) & -\sin(\vartheta_{i,p}^{[f]}) \\ \sin(\vartheta_{i,p}^{[f]}) & \cos(\vartheta_{i,p}^{[f]}) \end{bmatrix} \tag{5}$$

where the unknown rotation angle $\vartheta_{i,p}^{[f]} \propto \mathrm{RM}_{i,p}/f^2$ [3] and $\mathrm{RM}_{i,p}$ is the Rotation Measure which depends on the magnetic field and the electron density along the path i-p.

Therefore, the $(2MD+2M) \times 1$ complex unknown parameter vector is given by $\boldsymbol{\theta}^{[f]} = [\boldsymbol{\epsilon}^{[f]^T}, \mathbf{g}_1^{[f]^T}, \ldots, \mathbf{g}_M^{[f]^T}]^T$ in which $\boldsymbol{\epsilon}^{[f]} = [\vartheta_{1,1}^{[f]}, \ldots, \vartheta_{D,M}^{[f]}, \exp(j\varphi_{1,1}^{[f]}), \ldots, \exp(j\varphi_{D,M}^{[f]})]^T$ refers to the frequency dependent per-receiver and per-source ionospheric effects.

2.2. Noise modeling as a compound-Gaussian distribution

The presence of outliers has multiple causes in radio astronomy, such as errors in the sky model due to weak sources in the background [11] or man-made Radio Frequency Interference (RFI) [18], leading to statistics, with heavy-tailed distributions, different from the classical Gaussian case [1]. To ensure robustness in the proposed estimator, we adopt a two-scale compound-Gaussian noise modeling given for each baseline by

$$\mathbf{n}_{pq}^{[f]} = \sqrt{\tau_{pq}^{[f]}} \ \boldsymbol{\mu}_{pq}^{[f]}, \tag{6}$$

where the power factor $\tau_{pq}^{[f]}$ is a positive real random variable and the 4×1 vector $\boldsymbol{\mu}_{pq}^{[f]}$ follows a zero-mean complex circular Gaussian distribution, *i.e.*, $\boldsymbol{\mu}_{pq}^{[f]} \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Omega}^{[f]})$ [1]. Therefore, calibration amounts to estimate for each frequency f the $(2MD + 2M) \times 1$ vector $\boldsymbol{\theta}^{[f]}$ describing Jones matrices, as well as $B \times 1$ texture realizations $\boldsymbol{\tau}^{[f]} = [\tau_{12}^{[f]}, \tau_{13}^{[f]}, \ldots, \tau_{(M-1)M}^{[f]}]^T$ and the 4×4 speckle covariance matrix $\boldsymbol{\Omega}^{[f]}$ (which must satisfy, *e.g.*, $\mathrm{tr}\left\{\boldsymbol{\Omega}^{[f]}\right\} = 1$ to avoid the ambiguity with the power factor, the choice of this constraint being arbitrary [19]). In the following, we assume independence of $\mathbf{n}_{pq}^{[f]}$ between baselines and frequencies and no specific prior structure exists *w.r.t.* f. Let us note that the algorithm can be adapted to perform an independent estimation of unknown parameters for each time interval. In order to consider different time scales for the unknown effects, a similar approach to the one proposed in this work for the multi-frequency scenario can be adopted, using specific time variation models or assuming smoothness across time.

3. Description of the proposed estimator

In this section, we introduce the Relaxed Maximum Likelihood (RML) method. Then, the ADMM algorithm is proposed to estimate the frequency dependent parameters in a distributed way. In the iterative procedure, each subset of parameters is updated alternatively, while fixing the remaining parameters.

3.1. Robust estimation of Jones matrices

Robust calibration is based on the model (2) and the compound-Gaussian noise model (6). Estimations are performed iteratively with the ML method similar to that of [19]. Specifically, here, we choose not to specify the probability density function (pdf) of the texture parameters which are assumed unknown and deterministic, leading to the RML. By doing so, we ensure more flexibility and robustness to any prior

---

[1]It is possible to consider a baseline dependent covariance matrix $\boldsymbol{\Omega}_{pq}^{[f]}$. In this case, the proposed algorithm requires a few modifications which are straightforward.

mismatch *w.r.t.* the unknown pdf of $\boldsymbol{\tau}^{[f]}$. Using independency between measurements, the log-likelihood function of the RML reads

$$\log f\left(\{\mathbf{x}^{[f]}\}_{f\in\mathcal{F}} \mid \{\boldsymbol{\theta}^{[f]}, \boldsymbol{\tau}^{[f]}, \boldsymbol{\Omega}^{[f]}\}_{f\in\mathcal{F}}\right) = -\sum_{f\in\mathcal{F}} l^{[f]}(\boldsymbol{\theta}^{[f]}) - \sum_{f\in\mathcal{F}}\sum_{pq} \log|\pi\tau_{pq}^{[f]}\boldsymbol{\Omega}^{[f]}| \tag{7}$$

in which $l^{[f]}(\boldsymbol{\theta}^{[f]}) = \sum_{pq} \frac{1}{\tau_{pq}^{[f]}} \mathbf{u}_{pq}^{[f]^H}(\boldsymbol{\theta}^{[f]})\boldsymbol{\Omega}^{[f]^{-1}}\mathbf{u}_{pq}^{[f]}(\boldsymbol{\theta}^{[f]})$ with $\mathbf{u}_{pq}^{[f]}(\boldsymbol{\theta}^{[f]}) = \mathbf{x}_{pq}^{[f]} - \sum_{i=1}^{D}\mathbf{s}_{i,pq}^{[f]}(\boldsymbol{\theta}^{[f]})$. Let us recall that for a given frequency f, $\boldsymbol{\theta}^{[f]}$ is the $(2MD + 2M) \times 1$ vector of interest, $\boldsymbol{\tau}^{[f]}$ contains the B texture parameters and $\boldsymbol{\Omega}^{[f]}$ is the 4×4 speckle covariance matrix. After some calculus, the optimization of the log-likelihood function *w.r.t.* the noise parameters leads to the following texture estimate

$$\hat{\tau}_{pq}^{[f]} = \frac{1}{4}\mathbf{u}_{pq}^{[f]^H}(\boldsymbol{\theta}^{[f]})\boldsymbol{\Omega}^{[f]^{-1}}\mathbf{u}_{pq}^{[f]}(\boldsymbol{\theta}^{[f]}) \tag{8}$$

while the speckle covariance matrix reads

$$\hat{\boldsymbol{\Omega}}^{[f]} = \frac{4}{B}\sum_{pq} \frac{\mathbf{u}_{pq}^{[f]}(\boldsymbol{\theta}^{[f]})\mathbf{u}_{pq}^{[f]^H}(\boldsymbol{\theta}^{[f]})}{\mathbf{u}_{pq}^{[f]^H}(\boldsymbol{\theta}^{[f]})\hat{\boldsymbol{\Omega}}^{[f]^{-1}}\mathbf{u}_{pq}^{[f]}(\boldsymbol{\theta}^{[f]})} \tag{9}$$

followed by a normalization step, $\hat{\boldsymbol{\Omega}}^{[f]} \leftarrow \frac{\hat{\boldsymbol{\Omega}}^{[f]}}{\text{tr}\{\hat{\boldsymbol{\Omega}}^{[f]}\}}$, to avoid any ambiguity. Consequently, the proposed algorithm consists in updating alternatively $\{\hat{\boldsymbol{\theta}}^{[f]}\}_{f\in\mathcal{F}}$, $\{\hat{\boldsymbol{\Omega}}^{[f]}\}_{f\in\mathcal{F}}$ and $\{\hat{\boldsymbol{\tau}}^{[f]}\}_{f\in\mathcal{F}}$ in a global iterative loop. Estimation of $\{\boldsymbol{\theta}^{[f]}\}_{f\in\mathcal{F}}$ can be achieved iteratively, for a fixed $\{\hat{\boldsymbol{\tau}}^{[f]}\}_{f\in\mathcal{F}}$ and $\{\hat{\boldsymbol{\Omega}}^{[f]}\}_{f\in\mathcal{F}}$, thanks to the ADMM procedure. More specifically, we need to maximize successively the cost funtion *w.r.t.* the complex gain vector $\mathbf{g}^{[f]} = [\mathbf{g}_1^{[f]^T}, \ldots, \mathbf{g}_M^{[f]^T}]^T$ and on the other hand, *w.r.t.* the direction dependent ionospheric effects $\boldsymbol{\epsilon}^{[f]}$.

3.2. Distributed estimation of the frequency dependent parameters

To estimate the parameter vector $\{\boldsymbol{\theta}^{[f]}\}_{f\in\mathcal{F}}$, with $\{\hat{\boldsymbol{\tau}}^{[f]}\}_{f\in\mathcal{F}}$ and $\{\hat{\boldsymbol{\Omega}}^{[f]}\}_{f\in\mathcal{F}}$ fixed, we use the ADMM procedure [10], which is adapted for distributed optimization problems. In such case, a set of computational agents is considered. Each of them achieves calibration locally for a specific frequency f and transmits the estimates to the fusion center where consensus among all agents is enforced thanks to known constraints. Closed-form expressions are not straightforward to obtain, in order to overcome this difficulty, we rewrite the ionospheric phase delay in (4), for $i \in \{1, \ldots, D\}$ and $p \in \{1, \ldots, M\}$, as $\exp\left(j\varphi_{i,p}^{[f]}\right) = \mathbf{b}^{[f]^T}\mathbf{z}_{i,p}$ and the Faraday rotation angle in (5) as $\vartheta_{i,p}^{[f]} = \frac{1}{f^2}\overline{z}_{i,p}$ *s.t.* $\mathbf{z}_i = [\overline{z}_{i,1}, \ldots, \overline{z}_{i,M}, \mathbf{z}_{i,1}^T, \ldots, \mathbf{z}_{i,M}^T]^T$ is the vector of hidden variables for the i-th source direction, which are frequency independent and need to be estimated. The $N \times 1$ vector $\mathbf{b}^{[f]}$ is defined *s.t.* $[\mathbf{b}^{[f]}]_k = \left(\frac{1}{f}\right)^{k-1}$ where $N - 1$ is the order of the truncated approximation power series of $\exp\left(j\varphi_{i,p}^{[f]}\right)$. As regards the gains, we enforce smoothness with polynomials thanks to

$$\mathbf{g}_p^{[f]} = \tilde{\mathbf{B}}^{[f]}\tilde{\mathbf{z}}_p \tag{10}$$

where $\tilde{\mathbf{B}}^{[f]} = \tilde{\mathbf{b}}^{[f]^T} \otimes \mathbf{I}_2$, in which the $\tilde{N} \times 1$ vector $\tilde{\mathbf{b}}^{[f]}$ is defined *s.t.* $[\tilde{\mathbf{b}}^{[f]}]_k = \left(\frac{f-f_0}{f_0}\right)^{k-1}$ for a given reference frequency f_0 [6]. Finally, the $2\tilde{N} \times 1$ vector $\tilde{\mathbf{z}}_p$ contains the corresponding hidden variables for the p-th receiving element. Consequently, multi-frequency calibration is formulated as

$$\{\hat{\boldsymbol{\theta}}^{[f]}\}_{f \in \mathcal{F}}, \hat{\mathbf{z}} = \underset{\{\boldsymbol{\theta}^{[f]}\}_{f \in \mathcal{F}}, \mathbf{z}}{\operatorname{argmin}} \sum_{f \in \mathcal{F}} l^{[f]}(\boldsymbol{\theta}^{[f]}) \ \text{ s.t. } \ \boldsymbol{\theta}^{[f]} = \mathbf{B}^{[f]} \mathbf{z}$$

where $\mathbf{z} = [\overline{z}_{1,1}, \ldots, \overline{z}_{D,M}, \mathbf{z}_{1,1}^T \ldots, \mathbf{z}_{D,M}^T, \tilde{\mathbf{z}}_1^T, \ldots, \tilde{\mathbf{z}}_M^T]^T$ is the augmented vector of hidden variables and the global frequency basis matrix is given by $\mathbf{B}^{[f]} = \text{bdiag}\{\frac{1}{f^2}\mathbf{I}_{DM}, \mathbf{I}_{DM} \otimes \mathbf{b}^{[f]^T}, \mathbf{I}_M \otimes \tilde{\mathbf{B}}^{[f]}\}$. Thus, consensus optimization is associated with the following augmented Lagrangian problem [10] $L\left(\{\boldsymbol{\theta}^{[f]}\}_{f \in \mathcal{F}}, \mathbf{z}, \{\mathbf{y}^{[f]}\}_{f \in \mathcal{F}}\right) = \sum_{f \in \mathcal{F}} L^{[f]}\left(\boldsymbol{\theta}^{[f]}, \mathbf{z}, \mathbf{y}^{[f]}\right)$, where $L^{[f]}\left(\boldsymbol{\theta}^{[f]}, \mathbf{z}, \mathbf{y}^{[f]}\right) = l^{[f]}\left(\boldsymbol{\theta}^{[f]}\right) + h^{[f]}\left(\boldsymbol{\theta}^{[f]}, \mathbf{z}, \mathbf{y}^{[f]}\right)$, in which using [20, 21], it can be proven that

$$h^{[f]}\left(\boldsymbol{\theta}^{[f]}, \mathbf{z}, \mathbf{y}^{[f]}\right) = 2\Re\left\{\mathbf{y}^{[f]^H}\left(\boldsymbol{\theta}^{[f]} - \mathbf{B}^{[f]}\mathbf{z}\right)\right\} + \rho\|\boldsymbol{\theta}^{[f]} - \mathbf{B}^{[f]}\mathbf{z}\|_2^2 \tag{11}$$

and $\mathbf{y}^{[f]} = [\overline{y}_{1,1}^{[f]}, \ldots, \overline{y}_{D,M}^{[f]}, y_{1,1}^{[f]}, \ldots, y_{D,M}^{[f]}, \tilde{\mathbf{y}}_1^{[f]^T}, \ldots, \tilde{\mathbf{y}}_M^{[f]^T}]^T$ denotes the associated Lagrange multipliers for a given f while ρ is a regularization factor. Due to separability in frequency, calibration can be carried out independently by each agent. The three steps of the ADMM algorithm are the following for the t-th iteration

- $\left(\hat{\boldsymbol{\theta}}^{[f]}\right)^{t+1} = \underset{\boldsymbol{\theta}^{[f]}}{\operatorname{argmin}} L^{[f]}\left(\boldsymbol{\theta}^{[f]}, (\hat{\mathbf{z}})^t, \left(\hat{\mathbf{y}}^{[f]}\right)^t\right)$ performed locally by each agent $\tag{12}$

- $(\hat{\mathbf{z}})^{t+1} = \underset{\mathbf{z}}{\operatorname{argmin}} \sum_{f \in \mathcal{F}} h^{[f]}\left(\left(\hat{\boldsymbol{\theta}}^{[f]}\right)^{t+1}, \mathbf{z}, \left(\hat{\mathbf{y}}^{[f]}\right)^t\right)$ performed globally at the fusion center $\tag{13}$

- $\left(\hat{\mathbf{y}}^{[f]}\right)^{t+1} = \left(\hat{\mathbf{y}}^{[f]}\right)^t + \rho\left(\left(\hat{\boldsymbol{\theta}}^{[f]}\right)^{t+1} - \mathbf{B}^{[f]}(\hat{\mathbf{z}})^{t+1}\right)$ performed locally by each agent $\tag{14}$

Solving (13) directly leads to the following closed-form expression

$$\hat{\mathbf{z}} = \left(\rho \sum_{f \in \mathcal{F}}\left(\mathbf{B}^{[f]^H}\mathbf{B}^{[f]}\right)\right)^{-1}\left(\sum_{f \in \mathcal{F}} \mathbf{B}^{[f]^T}\left(\mathbf{y}^{[f]} + \rho\boldsymbol{\theta}^{[f]}\right)\right) \tag{15}$$

which is then broadcasted to all agents as a common global variable. Minimization (12) is performed thanks to an iterative approach. Specifically, we minimize *w.r.t.* $\boldsymbol{\epsilon}^{[f]}$ on the one hand, which is done by minimizing successively *w.r.t.* each $\vartheta_{i,p}^{[f]}$ and $\varphi_{i,p}^{[f]}$ for $i \in \{1, \ldots, D\}$ and $p \in \{1, \ldots, M\}$, and *w.r.t.* $\mathbf{g}^{[f]}$ on the other hand. Estimation of $\vartheta_{i,p}^{[f]}$ is computed thanks to a 1D Newton or gradient descent-type algorithm [22], possibly in parallel for $i \in \{1, \ldots, D\}$ and $p \in \{1, \ldots, M\}$, and requires to solve

$$\hat{\vartheta}_{i,p}^{[f]} = \underset{\vartheta_{i,p}^{[f]}}{\operatorname{argmin}} L^{[f]}\left(\boldsymbol{\theta}^{[f]}, \mathbf{z}, \mathbf{y}^{[f]}\right). \tag{16}$$

Regarding the ionospheric phase delays, minimization (12) for a given $\hat{\vartheta}_{i,p}^{[f]}$ for $i \in \{1, \ldots, D\}$ and $p \in \{1, \ldots, M\}$ leads to

$$\hat{\varphi}_{i,p}^{[f]} = \frac{1}{2}\arg\left\{-\frac{\alpha_{i,p}^{[f]}}{\beta_{i,p}^{[f]}}\right\}, \tag{17}$$

for which notations are specified in Appendix A. For sake of clarity, in the following, dependence *w.r.t.* parameters of interest is dropped and we recall that $\mathbf{g}_p^{[f]} = [[\mathbf{g}_p^{[f]}]_1, [\mathbf{g}_p^{[f]}]_2]^T$. The first-order derivative of the cost function (7) *w.r.t.* $[\mathbf{g}_p^{[f]}]_1$ leads, after some calculus, to

$$\frac{\partial l^{[f]}(\mathbf{g}^{[f]})}{\partial [\mathbf{g}_p^{[f]}]_1} = \sum_{\substack{q=1 \\ q>p}}^{M} \frac{1}{\tau_{pq}^{[f]}} \lambda_{pq}^{[f]} + \sum_{\substack{q=1 \\ q<p}}^{M} \frac{1}{\tau_{qp}^{[f]}} \tilde{\lambda}_{qp}^{[f]} \tag{18}$$

in which the expressions of $\lambda_{pq}^{[f]}$ and $\tilde{\lambda}_{qp}^{[f]}$ are given in Appendix B. After tedious calculations and rearrangement, we obtain the following solution

$$[\mathbf{g}_p^{[f]}]_1 = \frac{a_p^{[f]}}{b_p^{[f]}} \tag{19}$$

where $a_p^{[f]} = -t_p^{[f]^*} - [\mathbf{g}_p^{[f]}]_2 \left(\sum_{k=1}^{2S} [\mathbf{w}_p^{[f]}]_{2k} [\boldsymbol{\omega}_p^{[f]^*}]_{2k} + \sum_{k=1}^{2V} [\boldsymbol{\varsigma}_p^{[f]}]_{2k} [\boldsymbol{\varrho}_p^{[f]^*}]_{2k} \right) - [\tilde{\mathbf{y}}_p^{[f]}]_1 + \rho \tilde{\mathbf{z}}_p^T [\tilde{\mathbf{B}}^{[f]^T}]_{:,1}$ and $b_p^{[f]} = \sum_{k=0}^{2S-1} [\mathbf{w}_p^{[f]}]_{2k+1} [\boldsymbol{\omega}_p^{[f]^*}]_{2k+1} + \sum_{k=0}^{2V-1} [\boldsymbol{\varsigma}_p^{[f]}]_{2k+1} [\boldsymbol{\varrho}_p^{[f]^*}]_{2k+1} + \rho$, for which detailed notations are exposed in Appendix B. Estimation of $[\mathbf{g}_p^{[f]}]_2$ is performed in a similar way, leading to

$$[\mathbf{g}_p^{[f]}]_2 = \frac{\tilde{a}_p^{[f]}}{\tilde{b}_p^{[f]}} \tag{20}$$

where $\tilde{a}_p^{[f]} = -t_p^{[f]^*} - [\mathbf{g}_p^{[f]}]_1 \left(\sum_{k=0}^{2S-1} [\mathbf{w}_p^{[f]}]_{2k+1} [\boldsymbol{\omega}_p^{[f]^*}]_{2k+1} + \sum_{k=0}^{2V-1} [\boldsymbol{\varsigma}_p^{[f]}]_{2k+1} [\boldsymbol{\varrho}_p^{[f]^*}]_{2k+1} \right) - [\tilde{\mathbf{y}}_p^{[f]}]_2 + \rho \tilde{\mathbf{z}}_p^T [\tilde{\mathbf{B}}^{[f]^T}]_{:,2}$ and $\tilde{b}_p^{[f]} = \sum_{k=1}^{2S} [\mathbf{w}_p^{[f]}]_{2k} [\boldsymbol{\omega}_p^{[f]^*}]_{2k} + \sum_{k=1}^{2V} [\boldsymbol{\varsigma}_p^{[f]}]_{2k} [\boldsymbol{\varrho}_p^{[f]^*}]_{2k} + \rho$. We introduce the *Matlab* notation $[\cdot]_{:,k}$ to refer to the k-th column of the corresponding matrix. Finally, the global scheme of the proposed MRCA is described in Table 1 while the operation flow and signaling exchange between fusion center and local agents is shown in Figure 1.

4. Numerical simulations

In this part, we compare the performance of the proposed MRCA with a non-robust algorithm based on a Gaussian noise assumption and performed with an alternating least squares (ALS) method. This method consists in updating sequentially the parameters of interest in $\boldsymbol{\theta}^{[f]}$ for $f \in \mathcal{F}$, by minimizing a least squares cost function similar to that of [1, 6]. We also consider a per-channel calibration which means that calibration is performed for each frequency separately without exploiting information across frequency while multi-frequency algorithm refers to distributed calibration obtained with the ADMM procedure. We generate each electronic gain entry $[\mathbf{g}_p^{[f]}]_k$ as a complex circular Gaussian random variable with mean one and variance $\frac{1}{4}$, while ionospheric phase delay and Faraday rotation are randomly generated, as function of the TEC drawn from a uniform distribution $\mathcal{U}(1 \times 10^{17}, 5 \times 10^{17})$ and expressed in m^{-2}. We select $N = \tilde{N} = 6$ and the extended Lagrangian parameter is set at $\rho = 10$. We consider $M = 8$ receiving

7

elements, with hundreds of meters for the longest baseline, $D = 2$ bright signal sources and $D' = 4$ weak background sources. Let us mention that we are considering here notional examples for the purpose of algorithm validation. In Figure 2, we plot the Mean Square Error (MSE) of the real part of one given parameter of vector $\mathbf{g}^{[f_1]}$ as a function of the Signal-to-Interference-plus-Noise Ratio (SINR), for different number of frequencies F which are selected in the range $75 - 125$ MHz and $f_0 = 100$ MHz. The SINR is defined as the ratio of the normalized power of D calibrator sources over the sum of normalized power of D' background sources and a noise factor. For each SINR, we perform 100 Monte-Carlo runs. A few iterations, less than 5 for each loop in Table 1, are sufficient to attain stability in convergence, likewise for the so-called primal and dual residuals (approximately 5 iterations), whose convergence mostly depends on parameters ρ, N, $\tilde{N}$ and initializations [6, 10]. Initial values for $\hat{\mathbf{y}}^{[f]}$ are chosen as a vector full of ones. Better estimation is achieved with the proposed algorithm and the behavior remains the same for other parameters of interest.

In the following Figures, the MeqTrees software [16] is used to generate the observations and ionospheric effects need to be corrected in the presence of Gaussian background noise. We consider $M = 7$ receptors, $D = 2$ and $D' = 16$, taken from the SUMSS survey with a spectral index of 0.7. An image area of 3.5 by 3.5 degrees is shown in Figure 3, with and without the calibrator source, before adding any perturbation. The integration time per visibility point is fixed at 60s, the total synthesis time at 12 hours and the SINR is approximately equal to 4 dB. After calibration, performed with MRCA and different classical methods, and subtraction of the bright calibrator from the data, we show the dirty image, *i.e.*, the corrected residual in Figures 4 and 5, computed using *lwimager*. Corresponding peak flux of the recovered weak sources in Figure 5 are shown in Table 2. As expected, we notice better estimation performance and flux recovering when the proposed robust multi-frequency dependent calibration is performed in regards to the per-channel and non-robust cases, *i.e.*, lower error and more robustness are achieved.

5. Conclusion

In this work, a multi-frequency robust iterative calibration algorithm based on compound-Gaussian modeling and the ADMM approach was described for the so-called direction dependent distortion regime. Numerical simulations highlight the better estimation performance of the proposed technique *w.r.t.* non-robust and/or per-channel cases, and its ability to accurately solve for per-direction and per-antenna TEC and RM physical quantities.

Appendix A.

We specify here the notations which lead to (17). To this end, we need to consider $\frac{\partial l^{[f]}(\boldsymbol{\theta}^{[f]})}{\partial \varphi_{i,p}^{[f]}}$ and $\frac{\partial h^{[f]}(\boldsymbol{\theta}^{[f]})}{\partial \varphi_{i,p}^{[f]}}$. Thus, minimization (12) *w.r.t.* phase delay $\varphi_{i,p}^{[f]}$ leads to

$$\beta_{i,p}^{[f]} \exp\left(j\varphi_{i,p}^{[f]}\right) + \alpha_{i,p}^{[f]} \exp\left(-j\varphi_{i,p}^{[f]}\right) = 0 \tag{A.1}$$

in which $\beta_{i,p}^{[f]} = \sum_{\substack{q=1 \\ q>p}}^{M} \frac{j}{\tau_{pq}^{[f]}} \exp\left(-j\varphi_{i,q}^{[f]}\right)\delta_{i,pq}^{[f]} + \sum_{\substack{q=1 \\ q<p}}^{M} \frac{j}{\tau_{qp}^{[f]}} \exp\left(-j\varphi_{i,q}^{[f]}\right)\delta_{i,qp}^{[f]^H} + jy_{i,p}^{[f]^*} - j\rho\mathbf{z}_{i,p}^H\mathbf{b}^{[f]^*}$ and
$\alpha_{i,p}^{[f]} = \sum_{\substack{q=1 \\ q>p}}^{M} \frac{-j}{\tau_{pq}^{[f]}} \exp\left(j\varphi_{i,q}^{[f]}\right)\delta_{i,pq}^{[f]^H} - \sum_{\substack{q=1 \\ q<p}}^{M} \frac{j}{\tau_{qp}^{[f]}} \exp\left(j\varphi_{i,q}^{[f]}\right)\delta_{i,qp}^{[f]} - jy_{i,p}^{[f]} + j\rho\mathbf{b}^{[f]^T}\mathbf{z}_{i,p}$. We also introduce $\delta_{i,pq}^{[f]} = $
$\left(-\mathbf{x}_{pq}^{[f]^H} + \sum_{\substack{k=1 \\ k\neq i}}^{D} \mathbf{s}_{k,pq}^{[f]^H}\right)\mathbf{\Omega}^{[f]^{-1}}\mathbf{d}_{i,pq}^{[f]}$ and $\mathbf{d}_{i,pq}^{[f]} = \left(\mathbf{G}_q^{[f]^*}(\mathbf{g}_q^{[f]})\mathbf{H}_{i,q}^{[f]^*}\mathbf{F}_{i,q}^{[f]}(\vartheta_{i,q}^{[f]})\right) \otimes \left(\mathbf{G}_p^{[f]}(\mathbf{g}_p^{[f]}) \otimes \mathbf{H}_{i,p}^{[f]}\mathbf{F}_{i,p}^{[f]}(\vartheta_{i,p}^{[f]})\right)\mathbf{c}_i^{[f]}$.

Finally, we directly deduce (17) from (A.1).

Appendix B.

We describe here the notations introduced in the estimation of frequency dependent electronic gains. The expressions in (18) are given by

$$\lambda_{pq}^{[f]} = t_{pq}^{[f]} + \mathbf{w}_{pq}^{[f]^H}\left(\mathbf{I}_2 \otimes \mathbf{G}_p^{[f]^*}\right)\boldsymbol{\varpi}_{pq}^{[f]} \tag{B.1}$$

in which $t_{pq}^{[f]} = -\mathbf{x}_{pq}^{[f]^H}\mathbf{m}_{pq}^{[f]}$ with $\mathbf{m}_{pq}^{[f]} = \mathbf{\Omega}^{[f]^{-1}}\left(\mathbf{G}_q^{[f]^*} \otimes \mathbf{E}_1\right)\mathbf{w}_{pq}^{[f]}$ and $\mathbf{E}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Besides, $\mathbf{w}_{pq}^{[f]} = $
$\sum_{i=1}^{D}\left(\mathbf{H}_{i,q}^{[f]^*}\mathbf{Z}_{i,q}^{[f]^*}(\varphi_{i,q})\mathbf{F}_{i,q}^{[f]}(\vartheta_{i,q}^{[f]})\right) \otimes \left(\mathbf{H}_{i,p}^{[f]}\mathbf{Z}_{i,p}^{[f]}(\varphi_{i,p})\mathbf{F}_{i,p}^{[f]}(\vartheta_{i,p}^{[f]})\right)\mathbf{c}_i^{[f]}$ and $\boldsymbol{\varpi}_{pq}^{[f]} = \left(\mathbf{G}_q^{[f]} \otimes \mathbf{I}_2\right)\mathbf{m}_{pq}^{[f]}$. Similarly, we have

$$\tilde{\lambda}_{qp}^{[f]} = \tilde{t}_{qp}^{[f]} + \tilde{\mathbf{w}}_{qp}^{[f]^H}\left(\mathbf{G}_p^{[f]^*} \otimes \mathbf{I}_2\right)\tilde{\boldsymbol{\varpi}}_{qp}^{[f]} \tag{B.2}$$

in which $\tilde{t}_{qp}^{[f]} = -\tilde{\mathbf{w}}_{qp}^{[f]^H}\mathbf{x}_{qp}^{[f]}$ with $\tilde{\mathbf{w}}_{qp}^{[f]} = \mathbf{\Omega}^{[f]^{-1}}\left(\mathbf{E}_1 \otimes \mathbf{G}_q^{[f]}\right)\mathbf{w}_{qp}^{[f]}$ and $\tilde{\boldsymbol{\varpi}}_{qp}^{[f]} = \left(\mathbf{I}_2 \otimes \mathbf{G}_q^{[f]}\right)\mathbf{w}_{qp}^{[f]}$. The derivative in (18) can be written more compactly as

$$\frac{\partial l^{[f]}(\mathbf{g}^{[f]})}{\partial[\mathbf{g}_p^{[f]}]_1} = t_p^{[f]} + \mathbf{w}_p^{[f]^H}\left(\mathbf{I}_S \otimes \mathbf{I}_2 \otimes \mathbf{G}_p^{[f]^*}\right)\boldsymbol{\omega}_p^{[f]} + \tilde{\mathbf{w}}_p^{[f]^H}\left(\mathbf{I}_V \otimes \mathbf{G}_p^{[f]^*} \otimes \mathbf{I}_2\right)\tilde{\boldsymbol{\omega}}_p^{[f]} \tag{B.3}$$

in which $\mathbf{w}_p^{[f]} = [\frac{1}{\tau_{p(p+1)}^{[f]}}\mathbf{w}_{p(p+1)}^{[f]^T}, \ldots, \frac{1}{\tau_{pM}^{[f]}}\mathbf{w}_{pM}^{[f]^T}]^T$, $\tilde{\mathbf{w}}_p^{[f]} = [\frac{1}{\tau_{1p}^{[f]}}\tilde{\mathbf{w}}_{1p}^{[f]^T}, \ldots, \frac{1}{\tau_{(p-1)p}^{[f]}}\tilde{\mathbf{w}}_{(p-1)p}^{[f]^T}]^T$ and $t_p^{[f]} = \sum_{\substack{q=1 \\ q>p}}^{M}\frac{1}{\tau_{pq}^{[f]}}t_{pq}^{[f]} + $
$\sum_{\substack{q=1 \\ q<p}}^{M}\frac{1}{\tau_{qp}^{[f]}}\tilde{t}_{qp}^{[f]}$. Similarly, we introduce $\boldsymbol{\omega}_p^{[f]} = [\boldsymbol{\varpi}_{p(p+1)}^{[f]^T}, \ldots, \boldsymbol{\varpi}_{pM}^{[f]^T}]^T$ and $\tilde{\boldsymbol{\omega}}_p^{[f]} = [\tilde{\boldsymbol{\varpi}}_{1p}^{[f]^T}, \ldots, \tilde{\boldsymbol{\varpi}}_{(p-1)p}^{[f]^T}]^T$. If we are considering the p-th array element, we note $S = (M-p)$ and $V = (p-1)$. Let us define the permutation matrix $\mathbf{P} = [\mathbf{e}_1, \mathbf{e}_3, \mathbf{e}_2, \mathbf{e}_4]$ where the 4×1 vector $\mathbf{e}_k$ has zeros except at the k-th position which is equal to unity. With $\boldsymbol{\varsigma}_p^{[f]} = \left(\mathbf{I}_V \otimes \mathbf{P}^T\right)\tilde{\mathbf{w}}_p^{[f]}$ and $\boldsymbol{\varrho}_p^{[f]} = (\mathbf{I}_V \otimes \mathbf{P})\tilde{\boldsymbol{\omega}}_p^{[f]}$, we obtain the following 1D linear equation

$$\frac{\partial l^{[f]}(\mathbf{g}^{[f]})}{\partial[\mathbf{g}_p^{[f]}]_1} = t_p^{[f]} + [\mathbf{g}_p^{[f]^*}]_1\sum_{k=0}^{2S-1}[\mathbf{w}_p^{[f]^*}]_{2k+1}[\boldsymbol{\omega}_p^{[f]}]_{2k+1} + [\mathbf{g}_p^{[f]^*}]_1\sum_{k=0}^{2V-1}[\boldsymbol{\varsigma}_p^{[f]^*}]_{2k+1}[\boldsymbol{\varrho}_p^{[f]}]_{2k+1}$$
$$+ [\mathbf{g}_p^{[f]^*}]_2\sum_{k=1}^{2S}[\mathbf{w}_p^{[f]^*}]_{2k}[\boldsymbol{\omega}_p^{[f]}]_{2k} + [\mathbf{g}_p^{[f]^*}]_2\sum_{k=1}^{2V}[\boldsymbol{\varsigma}_p^{[f]^*}]_{2k}[\boldsymbol{\varrho}_p^{[f]}]_{2k}. \tag{B.4}$$

Considering (B.4) and $\frac{\partial h^{[f]}(\mathbf{g}^{[f]})}{\partial[\mathbf{g}_p^{[f]}]_1}$ leads to (19). Estimation of $[\mathbf{g}_p^{[f]}]_2$ is similar, except that

$\mathbf{m}_{pq}^{[f]} = \mathbf{\Omega}^{[f]^{-1}}\left(\mathbf{G}_q^{[f]^*} \otimes \mathbf{E}_2\right)\mathbf{w}_{pq}^{[f]}$ and $\tilde{\mathbf{w}}_{qp}^{[f]} = \mathbf{\Omega}^{[f]^{-1}}\left(\mathbf{E}_2 \otimes \mathbf{G}_q^{[f]}\right)\mathbf{w}_{qp}^{[f]}$ where $\mathbf{E}_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$.

References

[1] S. J. Wijnholds, A.-J. van der Veen, Multisource self-calibration for sensor arrays, IEEE Transactions on Signal Processing 57 (9) (2009) 3512–3522.

[2] S. J. Wijnholds, S. van der Tol, R. Nijboer, A.-J. van der Veen, Calibration challenges for future radio telescopes, IEEE Signal Processing Magazine 27 (1) (2010) 30–42.

[3] O. M. Smirnov, Revisiting the radio interferometer measurement equation. II. Calibration and direction-dependent effects, Astronomy & Astrophysics 527 (2011) A107.

[4] M. de Vos, A. W. Gunst, R. Nijboer, The LOFAR telescope: System architecture and signal processing, Proceedings of the IEEE 97 (8) (2009) 1431–1437.

[5] C. M. Trott, R. B. Wayth, Spectral calibration requirements of radio interferometers for Epoch of Reionisation science with the SKA, Publications of the Astronomical Society of Australia 33 (2016) e019.

[6] S. Yatawatta, Distributed radio interferometric calibration, Monthly Notices of the Royal Astronomical Society 449 (4) (2015) 4506–4514.

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

[8] M. P. Van Haarlem, et al., LOFAR: The LOw-Frequency ARray, Astronomy & Astrophysics 556 (2013) A2.

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

[10] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, Distributed optimization and statistical learning via the alternating direction method of multipliers, Foundations and Trends® in Machine Learning 3 (1) (2011) 1–122.

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

[12] E. Ollila, D. E. Tyler, V. Koivunen, H. Poor, Complex elliptically symmetric distributions survey, new results and applications, IEEE Transactions on Signal Processing 60 (11) (2012) 5597–5625.

[13] V. Ollier, M. N. El Korso, R. Boyer, P. Larzabal, M. Pesavento, Robust calibration of radio interferometers in non-Gaussian environment, IEEE Transactions on Signal Processing 65 (21) (2017) 5649–5660.

[14] C. Tasse, Nonlinear Kalman filters for calibration in radio interferometry, Astronomy & Astrophysics 566 (2014) A127.

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

[16] J. E. Noordam, O. M. Smirnov, The MeqTrees software system and its use for third-generation calibration of radio interferometers, Astronomy & Astrophysics 524 (2010) A61.

[17] S. van der Tol, A.-J. van der Veen, Ionospheric calibration for the LOFAR radio telescope, in: International Symposium on Signals, Circuits and Systems (ISSCS), Vol. 2, 2007, pp. 1–4.

[18] J. Raza, A.-J. Boonstra, A.-J. Van der Veen, Spatial filtering of RF interference in radio astronomy, IEEE Signal Processing Letters 9 (2) (2002) 64–67.

[19] X. Zhang, M. N. El Korso, M. Pesavento, MIMO radar target localization and performance evaluation under SIRP clutter, Signal Processing Journal 130 (2017) 217–232.

[20] L. Li, X. Wang, G. Wang, Alternating Direction Method of Multipliers for separable convex optimization of real functions in complex variables, Mathematical Problems in Engineering 2015 (Article ID 104531) (2015) 1–14.

[21] B. Jiang, S. Ma, S. Zhang, Alternating direction method of multipliers for real and complex polynomial optimization models, Optimization 63 (6) (2014) 883–898.

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

Table 1: MRCA

---

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

$\{\hat{\boldsymbol{\Omega}}^{[f]} \leftarrow \boldsymbol{\Omega}_{\text{init}}^{[f]}\}_{f\in\mathcal{F}}, \{\hat{\tau}_{pq}^{[f]} \leftarrow \tau_{pq_{\text{init}}}^{[f]}\}_{f\in\mathcal{F},\ p<q,\ (p,q)\in\{1,...,M\}^2}$

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

2   **while** *stop criterion unreached* **do**

3   **while** *stop criterion unreached* **do**

4   Obtain $\{\hat{\vartheta}_{i,p}^{[f]}\}_{i=1,...,D,\ p=1,...,M}$ locally from (16)

5   Obtain $\{\hat{\varphi}_{i,p}^{[f]}\}_{i=1,...,D,\ p=1,...,M}$ locally from (17)

6   Obtain $\hat{\mathbf{g}}^{[f]}$ locally from (19) and (20)

7   **end**

8   Obtain $\hat{\mathbf{z}}$ globally from (15)

9   Obtain $\hat{\mathbf{y}}^{[f]}$ locally from (14)

10   **end**

11   Obtain $\{\hat{\boldsymbol{\Omega}}^{[f]}\}_{f\in\mathcal{F}}$ from (9)

12   Obtain $\{\hat{\tau}_{pq}^{[f]}\}_{f\in\mathcal{F},\ p<q,\ (p,q)\in\{1,...,M\}^2}$ from (8)

13 end

---

input: $\left\{\mathbf{x}^{[f]}, \hat{\boldsymbol{\theta}}^{[f]}, \hat{\boldsymbol{\Omega}}^{[f]}\right\}_{f\in\mathcal{F}}$, $\{\hat{\tau}_{pq}^{[f]}\}_{f\in\mathcal{F},\ p<q,\ (p,q)\in\{1,...,M\}^2}$
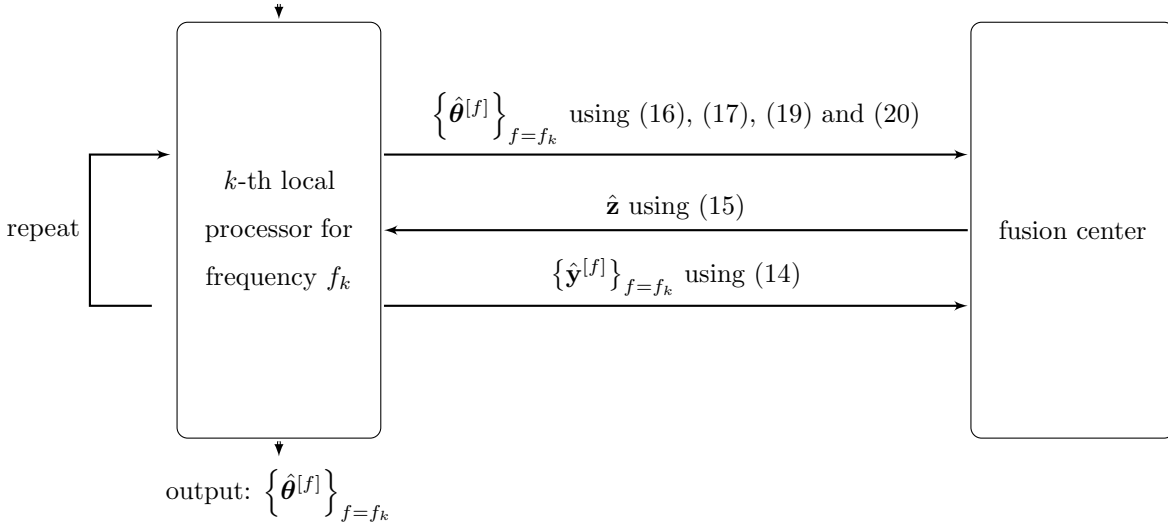

Figure 1: Operation flow and signaling exchange between the k-th local processor and the fusion center. The three arrows in the center are performed sequentially and repeated, which corresponds to the loop from line 2 to line 10 of Table 1.
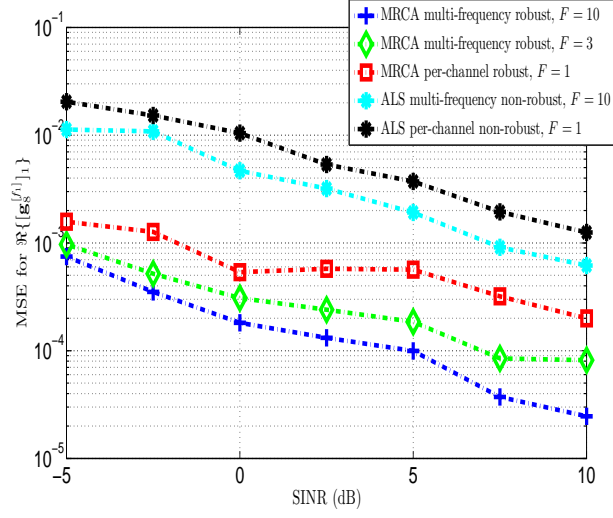
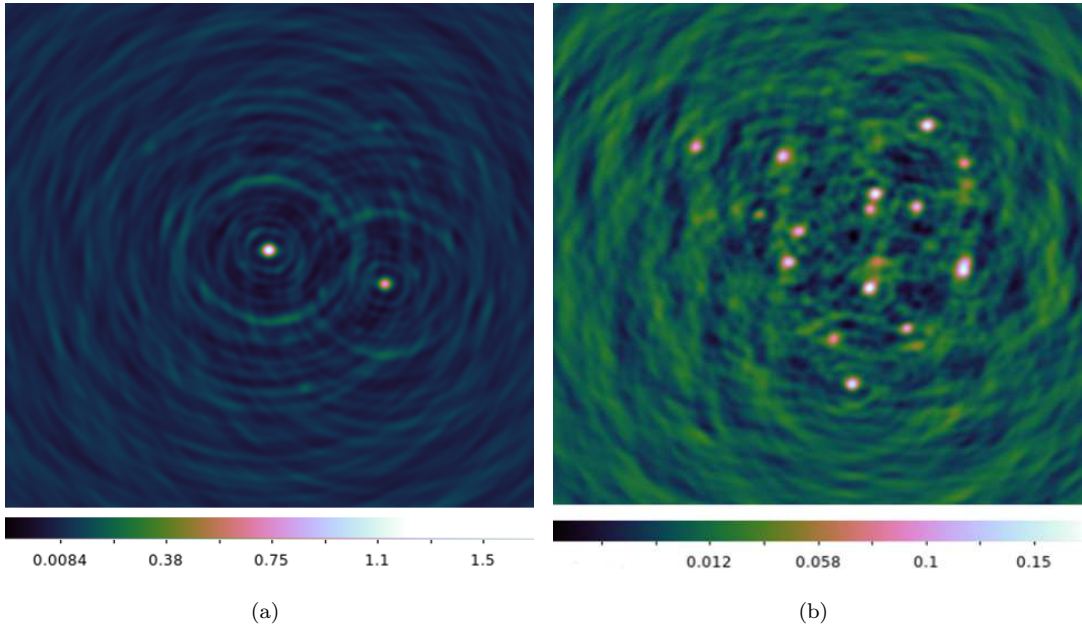Figure 2: MSE of the real part of a given complex gain vs. SINR.

(a)

(b)

Figure 3: Dirty image representing (a) the $D = 2$ calibrator sources and $D' = 16$ background sources which are hardly visible, (b) only the background sources.

Table 2: Recovered peak flux (in Jy) for the three sources in Figure 5.

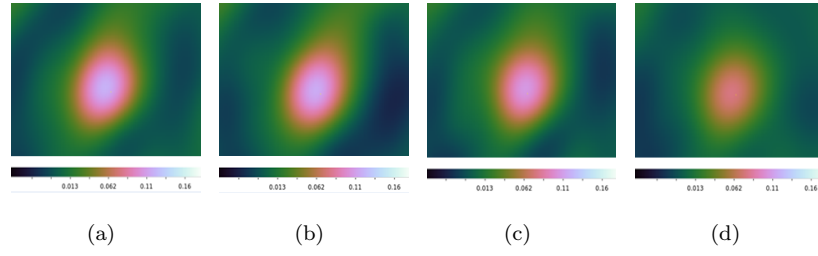| Position	Number 1	Number 2	Number 3
(a) True	0.167113	0.108398	0.122119
(b)	0.162519	0.0910741	0.1111702
(c)	0.136886	0.0882433	0.10028
(d)	0.115676	0.0831246	0.082365

12

Figure 4: Corrected residual for (a) ideal case, (b) MRCA, (c) multi-frequency non-robust calibration and (d) per-channel non-robust calibration, around one of the weakest background sources. The peak flux intensity is respectively given by (a) true 0.11856, (b) 0.117225, (c) 0.109427 and (d) 0.0808339 Jy.
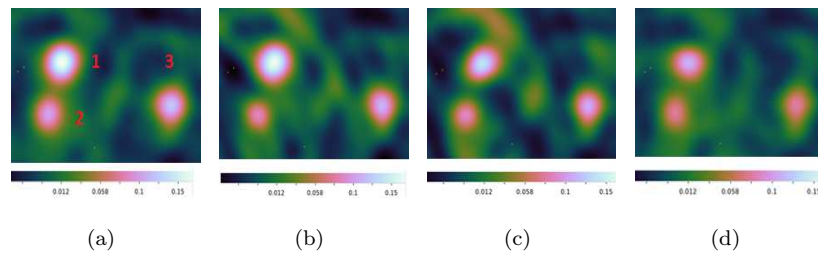

Figure 5: Corrected residual for (a) ideal case, (b) MRCA (c) multi-frequency non-robust calibration and (d) per-channel non-robust calibration in a small area surrounding three background sources.