HAL open science

[Technical Report] Multi-frequency calibration for DOA estimation with distributed sensors

Martin Brossard, Virginie Ollier, Mohammed Nabil El Korso, Remy Boyer, Pascal Larzabal

To cite this version:

Martin Brossard, Virginie Ollier, Mohammed Nabil El Korso, Remy Boyer, Pascal Larzabal. [Technical Report] Multi-frequency calibration for DOA estimation with distributed sensors. 2020. hal-02468904

HAL Id: hal-02468904

https://hal.science/hal-02468904

Preprint submitted on 6 Feb 2020

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.

[Technical Report]
Multi-frequency calibration for DOA estimation with distributed sensors

Martin BROSSARD, Virginie OLLIER, Mohammed Nabil EL KORSO, Rémy BOYER, Pascal LARZABAL *

31 Jan 2020

Abstract

In this work, we investigate direction finding in the presence of sensor gain uncertainties and directional perturbations for sensor array processing in a multi-frequency scenario. Specifically, we adopt a distributed optimization scheme in which coherence models are incorporated and local agents exchange information only between connected nodes in the network, i.e., without a fusion center. Numerical simulations highlight the advantages of the proposed parallel iterative technique in terms of statistical and computational efficiency.

Keywords: Calibration, source localization, multi-frequency, sensor array processing, distributed optimization.

1 Introduction

Calibration and Direction-of-Arrival (DoA) estimation is a major issue in array processing [2, 3]. The latter has been studied in several applications, e.g., radar, sonar, satellite, wireless communication and radio interferometric systems [4, 5], where we commonly use largely distributed sensors elements aiming to achieve high resolution. In all these sensor network applications, calibration is required as some parameters are not exactly known due to imperfect instrumentation or propagation conditions [6]. Let us note that calibration algorithms are distinguished by the presence [7] or absence [8] of one or more cooperative sources, named calibrator sources. Indeed, prior source information can be available [6] and consists mainly in the true/nominal

*M. Brossard is with Mines Paristech, France, V. Ollier is with Safran Electronics and Defense, France, M. N. El Korso is with Paris-Nanterre University, France, R. Boyer is with University of Lille, France and P. Larzabal is with ENS Paris-Saclay, France. This work was supported by ANR ASTRID project MARGARITA (ANR-17-ASTR-0015). Some of the contents have been partially published in [1].

directions and powers of calibrator sources (i.e., without any perturbation effects or antenna imperfections). Furthermore, most calibration algorithms are based on the least squares approach, with a sequential procedure updating each parameter alternatively [4]. The least squares estimator is indeed equivalent to the Maximum Likelihood (ML) method under a (unrealistic) Gaussian noise model.

The aim of the proposed methodology here is to estimate successively the unknown sensor gains and phase errors, along with the calibrator and noise parameters, through minimization of a proper weighting cost function. In this work, uncertainties are estimated from the array covariance matrix, since dealing directly with time series data and operating on the signal domain quickly becomes computationally unfeasible for a large number of samples [9]. The scenario under study is general but could be adapted to any practical application as in the radio astronomy context, where the number of parameters to estimate is tremendous and frequency bands are wide.

In the multi-frequency scenario, a suboptimal way to perform calibration is to consider one wavalength bin at a time, with only one centralized processor, which has access to data in the whole available range of wavelengths. In this work, we study an accelerated version based on the scalable form of the Alternating Direction Method of Multipliers (ADMM) [10, 11] with a specific network topology: there is no fusion center and agents exchange information only among themselves. The goal being to reduce the complexity in operation flow and signaling exchanging [12, 13, 14, 15, 16, 17]. For estimation of the directional gains, the compressive sensing framework, especially the sparse representation method, is well-adapted and has already been applied for source localization in fully and partially calibrated arrays [18, 19, 20, 21].

The notation used through this paper is the following: $(.)^*$, $(.)^T$, $(.)^H$, $(.)^{\odot\alpha}$, $\Re(.)$ and $[.]_n$ denote, respectively, the complex conjugate, transpose, Hermitian operator, element-wise raising to α, real part and the n-th element of a vector. The expectation operator is $\mathcal{E}\{.\}$, $\otimes$, $\circ$ and $\odot$ denote, respectively, the Kronecker, the Khatri-Rao and the Hadamard product. The operator diag(.) converts a vector into a diagonal matrix, blkdiag(.) is the block-diagonal operator, whereas vecdiag(.) produces a vector from the main diagonal of a matrix and vec(.) stacks the columns of a matrix on top of one another. The operators $\|.\|_2$ and $\|.\|_{\mathsf{F}}$ refer to the l_2 and Frobenius norms, respectively. Finally, $\mathbf{I}_P$ is the $P \times P$ identity matrix and $|\cdot|$ refers to the cardinality of a set.

2 Model setup

Let us consider Q emitting signal sources and P sensor elements in the array. Each source direction $q \in \{1, \ldots, Q\}$ is defined by a 2-dimensional

vector $\mathbf{d}_q = \left[d_q^l, d_q^m\right]^T$, s.t., all nominal/true known directions, without any disturbances, are stacked in $\mathbf{D}^{\mathrm{K}} = \left[\mathbf{d}_1^{\mathrm{K}}, \ldots, \mathbf{d}_Q^{\mathrm{K}}\right] \in \mathbb{R}^{2\times Q}$. Propagation conditions induce wavelength dependent distortions, leading to apparent source directions $\mathbf{D}_\lambda = [\mathbf{d}_{1,\lambda}, \ldots, \mathbf{d}_{Q,\lambda}]$ different from the true ones. Under the narrowband assumption, the array response matrix reads $\mathbf{A}_{\mathbf{D}_\lambda} = \frac{1}{\sqrt{P}} \exp\left(-j\frac{2\pi}{\lambda}\boldsymbol{\Xi}\mathbf{D}_\lambda\right)$ in which $\boldsymbol{\Xi} = [\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_P]^T \in \mathbb{R}^{P\times 2}$ includes the known Cartesian coordinates describing each sensor location in the array, s.t., for $p \in \{1, \ldots, P\}$, $\boldsymbol{\xi}_p = [x_p, y_p]^T$. Therefore, the $P \times 1$ narrowband signals measured by all antennas is written as follows, for the n-th time sample and wavelength λ,

$$\mathbf{x}_\lambda(n) = \mathbf{G}_\lambda \mathbf{A}_{\mathbf{D}_\lambda} \boldsymbol{\Gamma}_\lambda \mathbf{s}_\lambda(n) + \mathbf{n}_\lambda(n) \tag{1}$$

where the undirectional antenna gains are collected in the complex diagonal matrix $\mathbf{G}_\lambda = \mathrm{diag}\{\mathbf{g}_\lambda\} \in \mathbb{C}^{P\times P}$ and the directional gain responses, assumed identical for all antennas, are modeled by the diagonal matrix $\boldsymbol{\Gamma}_\lambda \in \mathbb{C}^{Q\times Q}$. Finally, $\mathbf{s}_\lambda(n) \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma}_\lambda)$ and $\mathbf{n}_\lambda(n) \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma}_\lambda^n)$ are the i.i.d. calibrator source signal and additive Gaussian thermal noise vectors with their corresponding diagonal covariance matrices $\boldsymbol{\Sigma}_\lambda = \mathrm{diag}\{\boldsymbol{\sigma}_\lambda\} \in \mathbb{R}^{Q\times Q}$ and $\boldsymbol{\Sigma}_\lambda^n = \mathrm{diag}\{\boldsymbol{\sigma}_\lambda^n\} \in \mathbb{R}^{P\times P}$, respectively. From (1), we deduce the following covariance matrix [1]

$$\mathbf{R}_\lambda(\mathbf{p}_\lambda) = \mathcal{E}\left\{\mathbf{x}_\lambda \mathbf{x}_\lambda^{\mathsf{H}}\right\} = \mathbf{E}_{\mathbf{D}_\lambda} \mathbf{M}_\lambda \mathbf{E}_{\mathbf{D}_\lambda}^H + \boldsymbol{\Sigma}_\lambda^n \tag{2}$$

where $\mathbf{E}_{\mathbf{D}_\lambda} = \mathbf{G}_\lambda \mathbf{A}_{\mathbf{D}_\lambda} \boldsymbol{\Sigma}_\lambda^{1/2}$ and $\mathbf{M}_\lambda = \boldsymbol{\Gamma}_\lambda \boldsymbol{\Gamma}_\lambda^H = \mathrm{diag}\{\mathbf{m}_\lambda\}$. In this context, the calibration problem consists in estimating the parameter vector of interest $\mathbf{p} = [\mathbf{p}_{\lambda_1}^T, \ldots, \mathbf{p}_{\lambda_F}^T]^T$ with F the total number of available wavelengths and $\mathbf{p}_\lambda = [\mathbf{g}_\lambda^T, \mathbf{d}_{1,\lambda}^T, \ldots, \mathbf{d}_{Q,\lambda}^T, \mathbf{m}_\lambda^T, \boldsymbol{\sigma}_\lambda^{\mathrm{n}^T}]^T$. To this end, we exploit sample covariance matrices $\hat{\mathbf{R}}_\lambda$, defined as $\hat{\mathbf{R}}_\lambda = \frac{1}{N}\sum_{n=1}^N \mathbf{x}_\lambda(n)\mathbf{x}_\lambda^H(n)$ for wavelength λ.

In estimation theory, the ML estimator is well-known for its statistical efficiency but not always easy to implement in practice. The Weighting Least Squares approach is an appropriate alternative as it is asymptotically equivalent to the ML for a large number of samples N. Therefore, we wish to minimize the following local cost function, associated to wavelength λ

$$\kappa_\lambda(\mathbf{p}_\lambda) = || \left(\mathbf{R}_\lambda(\mathbf{p}_\lambda) - \hat{\mathbf{R}}_\lambda\right) \odot \boldsymbol{\Omega}_\lambda ||_F^2 \tag{3}$$

where $\boldsymbol{\Omega}_\lambda = (\boldsymbol{\sigma}_\lambda^{\mathrm{n}} \boldsymbol{\sigma}_\lambda^{\mathrm{n}^T})^{\odot -\frac{1}{2}}$. Most sources are assumed buried beneath the noise and antennas are identical in the array with negligible mutual coupling. The aim of the designed calibration algorithm is to minimize the global cost

[1] As in [22], some commonly used assumptions are considered here to overcome scaling ambiguities, such as fixed phase for the first element and one reference source with fixed direction and directional gain/apparent power.

function $\kappa(\mathbf{p}) = \sum_{\lambda\in\Lambda} \kappa_\lambda(\mathbf{p}_\lambda)$ in a parallel and step-wise approach, with $\Lambda = \{\lambda_1, \ldots, \lambda_F\}$ the total set of available wavelengths. Usually, minimization is conducted w.r.t. one specific parameter while fixing the others in $\mathbf{p}_\lambda$ [22].Here, our approach is different: we propose an accelerated version where estimation is performed directly w.r.t. the consensus (hidden) variables, as described in Algorithm 1 and detailed in the following.

3 Description of the proposed estimator

To achieve multi-frequency calibration in the sensor array, coherence is imposed along wavelength subbands for both directional and undirectional gains, by imposing available constraints or enforcing smooth variation. The choice of the basis functions is motivated by the application under analysis and can be adapted accordingly.

3.1 Coherence model for the undirectional antenna gains

To impose coherence along subbands, we introduce a set of smooth wavelength dependent basis functions and express the gains as linear combinations. Let us define $\boldsymbol{\alpha}_p = [\alpha_{1,p}, \ldots, \alpha_{K_g,p}]^T \in \mathbb{C}^{K_g}$, the consensus vector for the p-th sensor with unknown linear coefficients. Therefore, for $p \in \{1, \ldots, P\}$ and $\lambda \in \Lambda$, $[\mathbf{g}_\lambda]_p = \sum_{k=1}^{K_g} b_{k,\lambda}\alpha_{k,p} = \mathbf{b}_\lambda^\mathsf{T}\boldsymbol{\alpha}_p$, in which $\mathbf{b}_\lambda = \left[b_{1,\lambda}, \ldots, b_{K_g,\lambda}\right]^\mathsf{T} \in \mathbb{R}^{K_g}$ stands for the polynomial terms, describing the variation of the undirectional gains w.r.t. wavelength. For instance, we can consider the typical basis function $b_{k,\lambda} = \left(\frac{f-f_0}{f_0}\right)^{k-1}$ in which $f = c/\lambda$ is the studied frequency of interest with c the speed of light and f_0 is the reference frequency [22, 23]. By stacking all vectors $\boldsymbol{\alpha}_p$, we obtain the global consensus vector $\boldsymbol{\alpha} = \left[\boldsymbol{\alpha}_1^\mathsf{T}, \ldots, \boldsymbol{\alpha}_P^\mathsf{T}\right]^\mathsf{T} \in \mathbb{C}^{PK_g}$, leading to

$$\mathbf{g}_\lambda = \mathbf{B}_\lambda\boldsymbol{\alpha}, \tag{4}$$

with $\mathbf{B}_\lambda = \left(\mathbf{I}_P \otimes \mathbf{b}_\lambda^\mathsf{T}\right)$.

3.2 Coherence model for the directional gains

Similarly as for the undirectional gains, the coherence model is defined as follows: let us consider $\boldsymbol{\alpha}_q \in \mathbb{R}^{K_m}$, for $q \in \{1, \ldots, Q\}$, such that for $\lambda \in \Lambda$,

$$[\mathbf{m}_\lambda]_q = \mathbf{b}_{\mathbf{m}_\lambda}^\mathsf{T}\boldsymbol{\alpha}_{\mathbf{m}_q}, \tag{5}$$

in which $\boldsymbol{\alpha}_{\mathbf{m}_q}$ is the vector of hidden variables for the q-th calibrator source, associated to directional gains $\mathbf{m}_\lambda$, while $\mathbf{b}_{\mathbf{m}_\lambda}$ is the corresponding basis vector. As in section 3.1, all $\boldsymbol{\alpha}_{\mathbf{m}_q}$ are stacked in $\boldsymbol{\alpha}_\mathbf{m} = \left[\boldsymbol{\alpha}_{\mathbf{m}_1}^T, \ldots, \boldsymbol{\alpha}_{\mathbf{m}_Q}^\mathsf{T}\right]^\mathsf{T} \in \mathbb{R}^{QK_m}$, finally leading to

$$\mathbf{m}_\lambda = \mathbf{B}_{\mathbf{m}_\lambda}\boldsymbol{\alpha}_\mathbf{m} \tag{6}$$

with $\mathbf{B}_{\mathbf{m}_\lambda} = \left(\mathbf{I}_Q \otimes \mathbf{b}_{\mathbf{m}_\lambda}^{\mathsf{T}}\right)$. We assume identical behavior for all sources but the process can be straightforwardly adapted to different behavior. In [22], the directional gains in $\mathbf{\Gamma}_\lambda$ were assumed inversely proportional to λ but here the algorithm can be adjusted to any general existing models.

3.3 Distributed network with a fusion center

Dealing with large data volumes delivered by advanced sensor array systems requires computationally efficient calibration algorithms, with a huge number of unknowns to solve. To improve both computational cost and estimation accuracy, distributed calibration has been proposed by exploiting data parallelism across frequency. Contrary to a centralized hardware architecture which processes all frequency bands at a single location and is therefore computationally challenging, distributed optimization introduces more than one compute agents and analyzes the data simultaneously across smaller frequency intervals [10]. By distributing the total computations across the network, we gain a significant reduction in operational and energy cost and each agent receives information indirectly across the whole frequency range, thus improving the calibration accuracy. To handle this, let us consider Z computational agents disposed on a network. Each agent has access to some wavelengths $\lambda \in \Lambda_z = \{\lambda_1^z, \ldots, \lambda_{J_z}^z\} \subset \Lambda$. The corresponding unknown parameters in $\mathbf{p}$ are estimated locally and consensus is enforced among agents by imposing constraints in (4) and (6).

To start with, let us focus on estimation of the undirectionnal sensor gains in section 3.1. We define $\boldsymbol{\alpha}^z$ as the local copy of the common optimization variable $\boldsymbol{\alpha}$ for the z-th agent and we note $\{\boldsymbol{\alpha}^z\}_{\mathcal{Z}} = \{\boldsymbol{\alpha}^1, \ldots, \boldsymbol{\alpha}^Z\}$ the set of all $\boldsymbol{\alpha}^z$ in the network. Calibration is reformulated as the following constrained problem

$$\hat{\boldsymbol{\alpha}} = \underset{\boldsymbol{\alpha}, \{\boldsymbol{\alpha}^z\}_{\mathcal{Z}}}{\operatorname{argmin}} \sum_{z=1}^{Z} \kappa^z\left(\boldsymbol{\alpha}^z\right) \quad \text{subject to } \boldsymbol{\alpha}^z = \boldsymbol{\alpha} \text{ for } z \in \{1, \ldots, Z\} \tag{7}$$

where $\kappa^z\left(\boldsymbol{\alpha}^z\right)$ is the cost function for the z-th agent, i.e., for $\lambda \in \Lambda_z$, which depends on the local variable $\boldsymbol{\alpha}^z$ and is associated to data $\left\{\hat{\mathbf{R}}_\lambda\right\}_{\lambda \in \Lambda_z}$. To solve this problem, we use the augmented Lagrangian, given by [24] $L\left(\{\boldsymbol{\alpha}^z\}_{\mathcal{Z}}, \boldsymbol{\alpha}, \{\mathbf{y}^z\}_{\mathcal{Z}}\right) = \sum_{z=1}^{Z} \kappa^z\left(\boldsymbol{\alpha}^z\right) + \Re\left\{\mathbf{y}^{z\mathsf{H}}\left(\boldsymbol{\alpha}^z - \boldsymbol{\alpha}\right)\right\} + \frac{\rho}{2}\left\|\boldsymbol{\alpha}^z - \boldsymbol{\alpha}\right\|_2^2$ where $\{\mathbf{y}^z\}_{\mathcal{Z}}$ are the Z Lagrange multipliers and ρ is the regularization term. We resort to the consensus ADMM in the scaled form by introducing the scaled dual variable $\mathbf{u}^z = \frac{1}{\rho}\mathbf{y}^z$ [10]. The three updates of the iterative

algorithm are therefore given by

$$\boldsymbol{\alpha}^{z[t+1]} = \underset{\boldsymbol{\alpha}^z}{\operatorname{argmin}}\, \kappa^z\left(\boldsymbol{\alpha}^z\right) + \frac{\rho}{2}\|\boldsymbol{\alpha}^z - \boldsymbol{\alpha}^{[t]} + \mathbf{u}^{z[t]}\|_2^2 = \underset{\boldsymbol{\alpha}^z}{\operatorname{argmin}}\, \tilde{L}^z\left(\boldsymbol{\alpha}^z, \boldsymbol{\alpha}^{[t]}, \mathbf{u}^{z[t]}\right) \tag{8}$$

$$\boldsymbol{\alpha}^{[t+1]} = \underset{\boldsymbol{\alpha}}{\operatorname{argmin}} \sum_{z=1}^{Z} \|\boldsymbol{\alpha}^{z[t+1]} - \boldsymbol{\alpha} + \mathbf{u}^{z[t]}\|_2^2 \tag{9}$$

$$\mathbf{u}^{z[t+1]} = \mathbf{u}^{z[t]} + \left(\boldsymbol{\alpha}^{z[t+1]} - \boldsymbol{\alpha}^{[t+1]}\right) \tag{10}$$

where t is the iteration counter. Minimization (9) leads to the following average, computed at the fusion center and sent to all agents in the network,

$$\hat{\boldsymbol{\alpha}} = \frac{1}{Z}\sum_{z=1}^{Z}\left(\boldsymbol{\alpha}^z + \mathbf{u}^z\right), \tag{11}$$

from which the undirectional gains can be directly deduced with (4). The local minimization step in (8) is the computationally most expensive one. To this end, we adopt an iterative approach and notice that the problem is separable w.r.t. each $\boldsymbol{\alpha}^z$, i.e., w.r.t. each agent. Let us assume $\boldsymbol{\alpha}^z$ and $(\boldsymbol{\alpha}^z)^*$ as two independent variables [25]. We then minimize $\tilde{L}^z\left(\boldsymbol{\alpha}^z, (\boldsymbol{\alpha}^z)^*, \boldsymbol{\alpha}, \mathbf{u}^z\right)$ w.r.t. $\boldsymbol{\alpha}^z$, considering $(\boldsymbol{\alpha}^z)^*$ as fixed and neglecting the diagonal elements in the cost function. In this case, the local cost function becomes separable w.r.t. the sub-vectors of $\boldsymbol{\alpha}^z$, i.e., $\boldsymbol{\alpha}^z = \left[\boldsymbol{\alpha}_1^{z\mathsf{T}}, \ldots, \boldsymbol{\alpha}_P^{z\mathsf{T}}\right]^{\mathsf{T}}$, where $\boldsymbol{\alpha}_p^z$ is the local consensus vector for the p-th sensor at the z-th agent. The following decompositions w.r.t. the sensor elements are also possible

$$\kappa^z(\boldsymbol{\alpha}^z) = \sum_{p=1}^{P} \kappa_p^z(\boldsymbol{\alpha}_p^z) \tag{12}$$

and $\tilde{L}^z\left(\boldsymbol{\alpha}^z; \boldsymbol{\alpha}, \mathbf{u}^z\right) = \sum_{p=1}^{P} \tilde{L}_p^z\left(\boldsymbol{\alpha}_p^z; \boldsymbol{\alpha}_p, \mathbf{u}_p^z\right)$ with $\tilde{L}_p^z\left(\boldsymbol{\alpha}_p^z; \boldsymbol{\alpha}_p, \mathbf{u}_p^z\right) = \kappa_p^z(\boldsymbol{\alpha}_p^z) + \frac{\rho}{2}\|\boldsymbol{\alpha}_p^z - \boldsymbol{\alpha}_p + \mathbf{u}_p^z\|_2^2$ where $\kappa_p^z(\boldsymbol{\alpha}_p^z)$ corresponds to the cost function for the p-th row of $\left\{\hat{\mathbf{R}}_\lambda\right\}_{\lambda\in\lambda_z}$, which only depends on $\boldsymbol{\alpha}_p^z$ since the remaining parameters are considered as fixed in this step. Let us define the operator $\mathcal{S}_p(.)$, that converts to a vector the p-th row of a matrix and removes the p-th element of this selected vector. We also introduce the quantity $\mathbf{R}_\lambda^{\mathrm{K}} = \mathbf{A}_{\mathbf{D}_\lambda}\boldsymbol{\Sigma}_\lambda\mathbf{M}_\lambda\mathbf{A}_{\mathbf{D}_\lambda}^{\mathsf{H}}$ (reference source model) and the following vectors

$$\hat{\mathbf{r}}_p^\lambda = \mathcal{S}_p\left(\hat{\mathbf{R}}_\lambda\right) \odot \boldsymbol{\omega}_p^\lambda, \qquad \mathbf{z}_p^\lambda = \mathcal{S}_p\left(\mathbf{R}_\lambda^{\mathrm{K}}\,\mathrm{diag}\left(\mathbf{B}_\lambda(\boldsymbol{\alpha}^z)^*\right)\right) \odot \boldsymbol{\omega}_p^\lambda \tag{13}$$

in which $\boldsymbol{\omega}_p^\lambda = \mathcal{S}_p\left(\boldsymbol{\Omega}_\lambda\right)$. In addition, let us consider the $J_z \times K_g$ matrix $\mathbf{B}^z = \left[\mathbf{b}_{\lambda_1^z}, \ldots, \mathbf{b}_{\lambda_{J_z}^z}\right]^{\mathsf{T}}$, $\hat{\mathbf{r}}_p^z = \left[\hat{\mathbf{r}}_p^{\lambda_1^z\mathsf{T}}, \ldots, \hat{\mathbf{r}}_p^{\lambda_{J_z}^z\mathsf{T}}\right]^{\mathsf{T}} \in \mathbb{C}^{(P-1)J_z\times 1}$, $\mathbf{Z}_p^z =$

$\text{blkdiag}\left(\mathbf{z}_p^{\lambda_1^z},\ldots,\mathbf{z}_p^{\lambda_{J_z}^z}\right) \in \mathbb{C}^{(P-1)J_z\times J_z}$ and $\tilde{\mathbf{Z}}_p^z = \mathbf{Z}_p^z\mathbf{B}^z$. We can thus write $\kappa_p^z(\boldsymbol{\alpha}_p^z)$ in (12) as $\kappa_p^z(\boldsymbol{\alpha}_p^z) = \left\|\hat{\mathbf{r}}_p^z - \tilde{\mathbf{Z}}_p^z\boldsymbol{\alpha}_p^z\right\|_2^2$ and finally obtain the following estimate

$$\hat{\boldsymbol{\alpha}}_p^z = \left(2\tilde{\mathbf{Z}}_p^{z\mathsf{H}}\tilde{\mathbf{Z}}_p^z + \rho\mathbf{I}_{K_g}\right)^{-1}\left(2\tilde{\mathbf{Z}}_p^{z\mathsf{H}}\hat{\mathbf{r}}_p^z + \rho\left(\boldsymbol{\alpha}_p - \mathbf{u}_p^z\right)\right). \tag{14}$$

3.4 Distributed network with no fusion center

We consider a specific formulation of the ADMM where every node in the network performs calibration locally and consensus is only reached with clearly identified neighbours without fusion center [12]. We note $\mathcal{N}_z$ the index set that corresponds to the neighbours of the z-th agent. The considered network architecture is exposed in Figure 1 where for example, $\mathcal{N}_3 = \{2,4\}$. We define the quantity $(\cdot)^{z,y}$ as the copy available at the z-th agent, transferred to the y-th agent. In such context, the minimization problem becomes

$$\begin{aligned}\hat{\boldsymbol{\alpha}} = &\underset{\{\boldsymbol{\alpha}^z,\boldsymbol{\beta}^{z,y},\forall y\in\mathcal{N}_z\}_{\mathcal{Z}}}{\text{argmin}} \sum_{z=1}^{Z}\kappa^z\left(\boldsymbol{\alpha}^z\right)\\ &\text{subject to } \boldsymbol{\alpha}^z = \boldsymbol{\beta}^{z,y},\quad \boldsymbol{\beta}^{y,z} = \boldsymbol{\beta}^{z,y}, \forall y\in\mathcal{N}_z,\ \text{for } z\in\{1,\ldots,Z\}\end{aligned} \tag{15}$$

where the auxiliary variables $\boldsymbol{\beta}^{z,y}$ impose consensus contraints on two neighboring agents and are meant to be local copies of $\boldsymbol{\alpha}$. The decentralized strategy enables to cooperatively minimize a sum of local objective functions, the final aim being to converge to a common value, with fast convergence speed and good estimation performance [26]. To obtain a more compact form of the problem in (15), we define $\boldsymbol{\beta}^z = \left[\{\boldsymbol{\beta}^{z,y}\}_{y\in\mathcal{N}_z}\right]$ and $\boldsymbol{\beta} = \left[\{\boldsymbol{\beta}^z\}_{z\in\{1,\ldots,Z\}}\right]$, leading to

$$\hat{\boldsymbol{\alpha}} = \underset{\{\boldsymbol{\alpha}^z,\boldsymbol{\beta}^z\}_{\mathcal{Z}}}{\text{argmin}} \sum_{z=1}^{Z}\kappa^z\left(\boldsymbol{\alpha}^z\right)\quad \text{subject to } \mathbf{H}^z\boldsymbol{\alpha}^z = \boldsymbol{\beta}^z,\ \text{for } z\in\{1,\ldots,Z\},\quad \boldsymbol{\beta}\in\mathcal{B} \tag{16}$$

with $\mathcal{B} = \left\{\boldsymbol{\beta}|\boldsymbol{\beta}^{z,y} = \boldsymbol{\beta}^{y,z}, \forall y\in\mathcal{N}_z,\ \text{for } z\in\{1,\ldots,Z\}\right\}$ and $\mathbf{H}^z = \mathbf{1}_{N_z\times 1}\otimes\mathbf{I}_{K_gP}$ where $N_z = |\mathcal{N}_z|$. As in section 3.3, the scaled version of the ADMM leads to

$$\boldsymbol{\alpha}^{z[t+1]} = \underset{\boldsymbol{\alpha}^z}{\text{argmin}}\,\kappa^z\left(\boldsymbol{\alpha}^z\right) + \frac{\rho_z^{[t+1]}}{2}\|\mathbf{H}^z\boldsymbol{\alpha}^z - \boldsymbol{\beta}^{z[t]} + \mathbf{u}^{z[t]}\|_2^2 = \underset{\boldsymbol{\alpha}^z}{\text{argmin}}\,\tilde{L}^z\left(\boldsymbol{\alpha}^z,\boldsymbol{\beta}^{z[t]},\mathbf{u}^{z[t]}\right) \tag{17}$$

$$\{\boldsymbol{\beta}^{z[t+1]}\}_{\mathcal{Z}} = \underset{\{\boldsymbol{\beta}^z\}_{\mathcal{Z}}\in\mathcal{B}}{\text{argmin}}\,L\left(\{\boldsymbol{\alpha}^{z[t+1]},\boldsymbol{\beta}^z,\mathbf{u}^{z[t]}\}_{\mathcal{Z}}\right) \tag{18}$$

$$\mathbf{u}^{z[t+1]} = \mathbf{u}^{z[t]} + \left(\mathbf{H}^z\boldsymbol{\alpha}^{z[t+1]} - \boldsymbol{\beta}^{z[t+1]}\right) \tag{19}$$

and through decomposition of the problem in (17) w.r.t. sensor dependence, we obtain

$$\hat{\boldsymbol{\alpha}}_p^z = \left(2\tilde{\mathbf{Z}}_p^{z\mathsf{H}}\tilde{\mathbf{Z}}_p^z + \rho N_z \mathbf{I}_{K_g}\right)^{-1}\left(2\tilde{\mathbf{Z}}_p^{z\mathsf{H}}\hat{\mathbf{r}}_p^z + \rho \mathbf{H}_p^{z\mathsf{H}}\left(\boldsymbol{\beta}_p^z - \mathbf{u}_p^z\right)\right) \tag{20}$$

with $\mathbf{H}_p^z = \mathbf{1}_{N_z\times 1} \otimes \mathbf{I}_{K_g}$. The selected variables $\boldsymbol{\beta}_p^z$ and $\mathbf{u}_p^z$ are obtained from $\boldsymbol{\beta}^z$ and $\mathbf{u}^z$ via an appropriate selection matrix. After considering the projection onto $\mathcal{B}$ and denoting the messages passed between the agent as

$$\boldsymbol{\gamma}^{z[t+1]} = \left[\{\boldsymbol{\gamma}^{z,y[t+1]}\}_{y\in\mathcal{N}_z}\right] = \mathbf{H}^z\boldsymbol{\alpha}^{z[t+1]} + \mathbf{u}^{z[t]}, \tag{21}$$

we solve (18) thanks to

$$\boldsymbol{\beta}^{z,y[t+1]} = \frac{1}{2}\left(\boldsymbol{\gamma}^{y,z[t+1]} + \boldsymbol{\gamma}^{z,y[t+1]}\right). \tag{22}$$

The steps of the proposed distributed method for calibration of sensor gains are exposed in Algorithm 1.2.

3.5 Estimation of directional gains

In this section, we describe the part of the algorithm dedicated to the estimation of DoA $\mathbf{D}_\lambda$ and directional gains $\mathbf{m}_\lambda$, for fixed sensor gains, with a sparse and distributed implementation. Assuming a sparse observed scene, we define dictionaries of steering matrices for $q \in \{1,\dots Q\}$ and $\lambda \in \Lambda$, as $\tilde{\mathbf{A}}_\lambda = \left[\tilde{\mathbf{A}}_{1,\lambda},\dots,\tilde{\mathbf{A}}_{Q,\lambda}\right] \in \mathbb{C}^{P\times N_g}$, where $N_g = \sum_{q=1}^{Q} N_q$ denotes the total number of directions on the grid. The sparse vectors in $\tilde{\mathbf{m}}_\lambda = \left[\tilde{\mathbf{m}}_{1,\lambda}^{\mathsf{T}},\dots,\tilde{\mathbf{m}}_{Q,\lambda}^{\mathsf{T}}\right]^{\mathsf{T}} \in \mathbb{R}^{N_g}$, contain the corresponding squared direction dependent gains. The covariance model is rewritten as $\mathbf{R}_\lambda = \tilde{\mathbf{E}}_\lambda\tilde{\mathbf{M}}_\lambda\tilde{\mathbf{E}}_\lambda^{\mathsf{H}} + \boldsymbol{\Sigma}_\lambda^{\mathrm{n}}$, in which $\tilde{\mathbf{M}}_\lambda = \mathrm{diag}(\tilde{\mathbf{m}}_\lambda) = \left(\mathbf{I}_{N_g}\otimes\mathbf{b}_\lambda^T\right)\mathrm{blkdiag}\left(\boldsymbol{\alpha}_1,\dots,\boldsymbol{\alpha}_{N_g}\right)$, $\tilde{\mathbf{E}}_\lambda = \mathbf{G}_\lambda\tilde{\mathbf{A}}_\lambda\tilde{\boldsymbol{\Sigma}}_\lambda^{\frac{1}{2}}$ and $\tilde{\boldsymbol{\Sigma}}_\lambda = \mathrm{blkdiag}\left(\mathbf{I}_{N_1}\left[\boldsymbol{\sigma}_\lambda\right]_1,\dots,\mathbf{I}_{N_Q}\left[\boldsymbol{\sigma}_\lambda\right]_Q\right)$. To handle the DoA estimation and satisfy both sparsity and positivity requirements, we use the Distributed Iterative Hard Thresholding (IHT) [27, 28]. But contrary to [22], the following hard-thresholding operator $\mathcal{H}_1\left(\sum_{\lambda\in\Lambda}\left(\check{\mathbf{V}}_\lambda^{q\mathsf{T}}\check{\hat{\mathbf{r}}}_\lambda^q\right)^{\odot 2}\right)$ is considered to provide access to the DoA of the q-th source, and a first estimate of the directional gain $\check{\mathbf{m}}_q^z$. The quantity $(\cdot)^q$ refers to the q-th column of a matrix, the expression $\check{(\cdot)}$ discards the elements corresponding to the diagonal of $\hat{\mathbf{R}}_\lambda$ and the hard thresholding operator $\mathcal{H}_s(.)$ keeps the s-largest components and sets the remaining entries equal to zero. Finally, thanks to (5) and dealing with the consensus variables as in section 3.4, the minimization problem becomes

$$\hat{\boldsymbol{\alpha}}_{\mathbf{m}_q} = \underset{\{\boldsymbol{\alpha}_{\mathbf{m}_q}^{z,z},\{\boldsymbol{\alpha}_{\mathbf{m}_q}^{z,y}\}_{y\in\mathcal{N}_z}\}_{\mathcal{Z}}}{\mathrm{argmin}} \sum_{z=1}^{Z}\eta_q^z\left(\boldsymbol{\alpha}_{\mathbf{m}_q}^{z,z}\right) \quad \text{subject to } \boldsymbol{\alpha}_{\mathbf{m}_q}^{z,z} = \boldsymbol{\alpha}_{\mathbf{m}_q}^{y,z}, \forall y\in\mathcal{N}_z, \text{ for } z\in\{1,\dots,Z\} \tag{23}$$

where we benefit from the previous hard-thresholding estimate to define $\eta_q^z\left(\boldsymbol{\alpha}_{\mathbf{m}_q}\right) = \sum_{\lambda\in\Lambda_z}\left\|\check{m}_{q,\lambda} - \mathbf{b}_{\mathbf{m}_\lambda}^T\boldsymbol{\alpha}_{\mathbf{m}_q}\right\|_2^2 = \left\|\check{\mathbf{m}}_q^z - \mathbf{B}_{\mathbf{m}}^z\boldsymbol{\alpha}_{\mathbf{m}_q}\right\|_2^2$ with $\check{\mathbf{m}}_q^z = [\check{m}_{q,\lambda_1^z},\ldots,\check{m}_{q,\lambda_{J_z}^z}]^T$ and $\mathbf{B}_{\mathbf{m}}^z = \left[\mathbf{b}_{\mathbf{m}_{\lambda_1^z}},\ldots,\mathbf{b}_{\mathbf{m}_{\lambda_{J_z}^z}}\right]^\mathsf{T}$. As previously, we impose consensus between neighbours thanks to some auxiliary variables but due to lack of space, we only present here the resulting local update for $\boldsymbol{\alpha}_{\mathbf{m}_q}^z$,

$$\hat{\boldsymbol{\alpha}}_{\mathbf{m}_q}^z = \left(2\mathbf{B}_{\mathbf{m}}^{z\mathsf{T}}\mathbf{B}_{\mathbf{m}}^z + \rho^z\mathbf{H}_{\mathbf{m}}^{z^\mathsf{T}}\mathbf{H}_{\mathbf{m}}^z\right)^{-1}\left(2\mathbf{B}_{\mathbf{m}}^{z\mathsf{T}}\check{\mathbf{m}}_q^z + \rho^z\mathbf{H}_{\mathbf{m}}^{z^\mathsf{T}}\left(\boldsymbol{\beta}_{\mathbf{m}}^z - \mathbf{u}_{\mathbf{m}}^z\right)\right) \quad (24)$$

where $\mathbf{H}_{\mathbf{m}}^z = \mathbf{1}_{N_z\times 1}\otimes\mathbf{I}_{K_m\times K_m}$. From $\hat{\boldsymbol{\alpha}}_{\mathbf{m}_q}^z$, we obtain an estimate of $[\mathbf{m}_\lambda]_q$ and process the next source, as shown in Algorithm 1.3.

4 Numerical simulations

In order to evaluate the method, we consider realistic simulations for the radio astronomy context where the new generation of phased array systems such as the Low Frequency Array (LOFAR) and the Square Kilometre Array (SKA) requires the development of new advanced signal processing techniques for calibration purpose [4, 29]. Indeed, lack of calibration leads to dramatic effects and distortions in the reconstructed images. We consider $P = 60$ antennas spread over a five-armed spiral [30, 31], which corresponds to the LOFAR's Initial Test Station. Let us assume a sky model with $Q = 3$ strong calibrator sources and $Q^{\mathrm{U}} = 8$ weak unknown sources in the background. The reference frequency f_0 is set to 30 MHz and we consider frequencies ranging from 29.6 MHz to 30.4 MHz, with $Z = 3$ agents in the network and $N_z = 2$. The polynomial orders are chosen as $K_g = K_m = 3$. The consensus variables $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}_{\mathbf{m}}$ are initialized as zeros and the squared directional gains are generated thanks to power law functions $(\lambda/\lambda_0)^{k-1}$ for $k\in\{1,\ldots,K_m\}$.

4.1 Influence of the number of frequency channels

First of all, we investigate the statistical performance of the proposed distributed algorithm as a function of the number of samples N or the Signal-to-Noise Ratio (SNR). The SNR is defined as the ratio between the sum of apparent powers for all Q sources and the noise power. Results are averaged for 100 Monte-Carlo runs. In Figure 2, we plot the three following cases: $F = 3$ and each agent handles one frequency, i.e., $J_z = 1$ (green curve), $F = 9$ with $J_z = 3$ (blue curve) and $F = 27$ with $J_z = 9$ (red curve). In Figure 2 (a), we plot the Root Mean Square Error (RMSE) as a function of N for the undirectional gains $\mathbf{g}_\lambda$, defined as $\epsilon_{\mathrm{RMSE}}^{\mathbf{g}_\lambda} = \frac{1}{\sqrt{PF}}\sum_{\lambda\in\Lambda}\left\|\mathbf{B}_\lambda\hat{\boldsymbol{\alpha}} - \mathbf{B}_\lambda\boldsymbol{\alpha}\right\|_2$, for fixed SNR $= -36$ dB. A similar figure is presented in Figure 2 (b), for the source directions $\mathbf{D}_\lambda$, as a function of the SNR and fixed $N = 2^8$. We illustrate the performance by comparing with the mono-calibration scenario

where each agent handles one single frequency, independently. We notice that mono-calibration is clearly improved, by using a distributed procedure where the whole information is flowing through the entire network.

4.2 Influence of the network architecture

We aim to show the advantages of the proposed distributed network with no fusion center and only exchange of local information between neighboring agents, in terms of complexity. With similar number of iterations in all loops of the algorithm, different estimation performance are attained in Figure 2 (a) while similar RMSE is reachable in Figure 2 (b) but with an additional computational cost if there is a fusion center (an increase of at least a factor 5 in computing time).

4.3 Convergence analysis

We illustrate the convergence behavior of the proposed algorithm by analyzing the following residuals as function of the iteration number. Depending on the iteration in Algorithm 1, we plot the primal residual as a function of the iteration number of Algorithm 1.2, defined as $\epsilon_p^{[t]} = \frac{1}{\sqrt{PK_gZN_z}} \sum_{z=1}^{Z} \left\| \mathbf{H}^z \boldsymbol{\alpha}^{z[t]} - \boldsymbol{\beta}^{z[t]} \right\|_2$. Likewise, we also study the different estimates between agents through $\epsilon_{\text{DIFF}}^{[t]} = \frac{1}{\sqrt{PK_gZ(Z-1)}} \sum_{z,z'=1}^{Z} \left\| \boldsymbol{\alpha}^{z[t]} - \boldsymbol{\alpha}^{z'[t]} \right\|_2$. Similar statistical behavior can be obtained for corresponding residuals in Algorithm 1.3

5 Conclusion

In this work, we proposed an iterative algorithm for parallel calibration, applied in a general context of sensor array processing: complex electronic gains are imprecisely known and propagation disturbances lead to deviations in the source locations. In order to reduce the communication overhead, the specific variation of parameters across wavelength is exploited in a distributed network with no fusion center and local exchange of information between adjacent connected nodes. The two main steps of the algorithm are based on the scalable form of the ADMM and distributed IHT procedures. We highlighted the effectiveness and time efficiency of the proposed method using simulated data, even in the presence of non-calibrator sources at unknown directions.

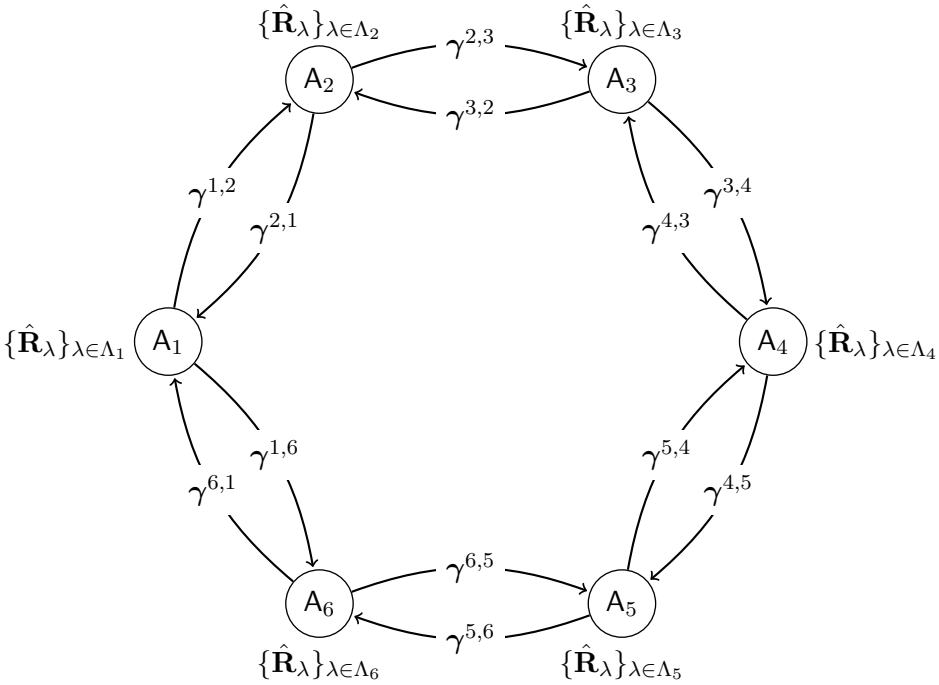


Figure 1: Example of distributed network with no fusion center.

Algorithm 1: Proposed calibration algorithm

Input: $\left\{\hat{\mathbf{R}}_\lambda\right\}_{\lambda\in\Lambda}, \mathbf{D}^{\mathrm{K}}, \eta_{\mathbf{p}}$;
Initialize: set
$i = 0, \left\{\mathbf{g}_\lambda = \mathbf{g}_\lambda^{[0]}, \mathbf{D}_\lambda = \mathbf{D}^{\mathrm{K}}, \mathbf{m}_\lambda = \mathbf{m}_\lambda^{[0]}, \boldsymbol{\Omega}_\lambda = \mathbf{1}_{P\times P}\right\}_{\lambda\in\Lambda}$;
repeat
$i = i + 1$;
Estimate in parallel $\left\{\mathbf{g}_\lambda^{[i]}\right\}_{\lambda\in\Lambda}$ with **Algorithm 1.2**;
Estimate in parallel $\left\{\mathbf{D}_\lambda^{[i]}, \mathbf{m}_\lambda^{[i]}, \boldsymbol{\sigma}_\lambda^{\mathrm{n}[i]}\right\}_{\lambda\in\Lambda}$ with **Algorithm 1.3**;
Update locally $\left\{\boldsymbol{\Omega}_\lambda^{[i]}\right\}_{\lambda\in\Lambda}$;
until $\left\|\mathbf{p}^{[i-1]} - \mathbf{p}^{[i]}\right\|_2 \leq \left\|\mathbf{p}^{[i]}\right\|_2 \eta_{\mathbf{p}}$;
Output: $\hat{\mathbf{p}} = \left[\mathbf{p}_{\lambda_1}^{[i]\mathsf{T}}, \ldots, \mathbf{p}_{\lambda_F}^{[i]\mathsf{T}}\right]^{\mathsf{T}}$;

References

[1] M. Brossard, V. Ollier, M. N. El Korso, R. Boyer, P. Larzabal, Fast decentralized multi-frequency calibration for doa estimation with distributed sensors, in: DAT 2020, Proceeding Conf. Algiers, Algeria, 2020.

[2] P. Stoica, K. C. Sharman, Maximum likelihood methods for direction-of-arrival estimation, IEEE Transactions on Acoustics, Speech and Signal Processing 38 (7) (1990) 1132–1143.

[3] S. Vorobyov, A. B. Gershman, K. M. Wong, Maximum likelihood direction-of-arrival estimation in unknown noise fields using sparse sensor arrays, IEEE Transactions on Signal Processing 53 (1) (2005) 34–43.

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

[5] L. C. Godara, Application of antenna arrays to mobile communications, part II: Beam-forming and direction-of-arrival considerations, Proceedings of the IEEE 85 (8) (1997) 1195–1245.

[6] B. C. Ng, C. M. S. See, Sensor-array calibration using a maximum-likelihood approach, IEEE Transactions on Antennas and Propagation 44 (6) (1996) 827–835.

[7] B. C. Ng, A. Nehorai, Active array sensor localization, Elsevier Signal processing 44 (3) (1995) 309–327.

[8] A. J. Weiss, B. Friedlander, Array shape calibration using sources in unknown locations-a maximum likelihood approach, IEEE Transactions on Acoustics, Speech, and Signal Processing 37 (12) (1989) 1958–1966.

Algorithm 1.2: Distributed estimation of consensus variables for undirectional gains

Input: $\left\{\hat{\mathbf{R}}_\lambda\right\}_{\lambda\in\Lambda}, \mathbf{p}^{[i-1]}, \eta_{\boldsymbol{\alpha}}$;

Initialize: set $t=0, \boldsymbol{\alpha}^z = \boldsymbol{\alpha}^{[i-1]}$, $\mathbf{R}^{\mathrm{K}}_\lambda = \mathbf{A}_{\mathbf{D}^{[i-1]}_\lambda}\boldsymbol{\Sigma}_\lambda\mathbf{M}^{[i-1]}_\lambda\mathbf{A}^{\mathsf{H}}_{\mathbf{D}^{[i-1]}_\lambda}$;

while *stop criterion unreached* **do**

$t = t+1$;

Estimate locally $\boldsymbol{\alpha}^{z[t]}$ with **Algorithm 1.2.2**;

Calculate locally $\boldsymbol{\gamma}^{z[t]}$ with (21);

$\Rightarrow$ Broadcast values $\boldsymbol{\gamma}^{z,y[t]}$ to region $y \in \mathcal{N}_z$;

$\Leftarrow$ Receive values $\boldsymbol{\gamma}^{y,z[t]}$ from region $y \in \mathcal{N}_z$;

Estimate locally $\boldsymbol{\beta}^{z[t]}$ with (22) ;

Update locally $\mathbf{u}^{z[t]}$ with (19);

Update possibly $\rho^{[t]}_z$ with [10, 32];

Algorithm 1.2.2: local estimation of $\boldsymbol{\alpha}^z$

Input: $\left\{\hat{\mathbf{R}}_\lambda, \mathbf{R}^{\mathrm{K}}_\lambda,\right\}_{\lambda\in\Lambda_z}, \boldsymbol{\alpha}^{z[t-1]}, \boldsymbol{\beta}^{z[t-1]}, \mathbf{u}^{z[t-1]}, \eta_{\boldsymbol{\alpha}^z}$;

Initialize: set $t^z = 0, \boldsymbol{\alpha}^z = \boldsymbol{\alpha}^{z[t-1]}$;

while $\left\|\boldsymbol{\alpha}^{z[t^z-1]} - \boldsymbol{\alpha}^{z[t^z]}\right\|_2 \geq \left\|\boldsymbol{\alpha}^{z[t^z]}\right\|_2 \eta_{\boldsymbol{\alpha}^z}$ **do**

$t^z = t^z + 1$;

for $p \in \{1, \ldots, P\}$ **do**

Update $\tilde{\mathbf{Z}}^z_p$;

Estimate $\hat{\boldsymbol{\alpha}}^{z[t^z]}_p$ with (20) ;

Update $(\hat{\boldsymbol{\alpha}}^z_p)^*$;

Output: $\hat{\boldsymbol{\alpha}}^z = \boldsymbol{\alpha}^{z[t^z]}$;

[9] S. J. Wijnholds, S. Chiarucci, Blind calibration of phased arrays using sparsity constraints on the signal model, in: 24th European Signal Processing Conference (EUSIPCO), 2016, pp. 270–274.

[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. Kazemi, S. Yatawatta, S. Zaroubi, Clustered calibration: an improvement to radio interferometric direction-dependent self-calibration, Monthly Notices of the Royal Astronomical Society 430 (2) (2013) 1457–1472.

Algorithm 1.3: Distributed estimation of $\{\mathbf{m}_\lambda, \mathbf{D}_\lambda, \boldsymbol{\sigma}^{\mathrm{n}}_\lambda\}_{\lambda\in\Lambda}$

Input: $\left\{\hat{\mathbf{R}}_\lambda\right\}_{\lambda\in\Lambda}, \mathbf{p}^{[i-1]}, \eta_{\boldsymbol{\alpha}_\mathbf{m}}, \eta_\mathbf{D}, \hat{\mathbf{r}}_\lambda = \mathrm{vec}\left(\hat{\mathbf{R}}_\lambda \odot \boldsymbol{\Omega}_\lambda\right)$,
$\mathbf{V}_\lambda = (\boldsymbol{\Sigma}^{\mathrm{n}}_\lambda)^{-\frac{1}{2}}\,\tilde{\mathbf{E}}^*_\lambda \otimes (\boldsymbol{\Sigma}^{\mathrm{n}}_\lambda)^{-\frac{1}{2}}\,\tilde{\mathbf{E}}_\lambda$;

Init: set $k=0$, $\mathbf{g}^{[k]}_\lambda = \mathbf{g}^{[i]}_\lambda$, $\mathbf{M}^{[k]}_\lambda = \mathbf{M}^{[i-1]}_\lambda$, $\mathbf{D}^{[k]}_\lambda = \mathbf{D}^{[i-1]}_\lambda, \boldsymbol{\sigma}^{\mathrm{n}[k]}_\lambda = \boldsymbol{\sigma}^{\mathrm{n}[i-1]}_\lambda$;

while $\left\|\boldsymbol{\alpha}^{[k-1]}_\mathbf{m} - \boldsymbol{\alpha}^{[k]}_\mathbf{m}\right\|_2 \geq \left\|\boldsymbol{\alpha}^{[k]}_\mathbf{m}\right\|_2 \eta_{\boldsymbol{\alpha}_\mathbf{m}}$ *and* $\sum_{\lambda\in\Lambda}\left\|\mathbf{D}^{[k-1]}_\lambda - \mathbf{D}^{[k]}_\lambda\right\|_\mathsf{F} \geq \eta_\mathbf{D}$ **do**

1 $k = k+1$;

for $q \in \{1,\ldots,Q\}$ **do**

foreach $\mathsf{A}_z, z \in \{1,\ldots,Z\}$ **do**

foreach $\lambda \in \Lambda_z$ **do**

2 Calculate locally the residual $\check{\hat{\mathbf{r}}}^q_\lambda$ as indicated in [22];

3 Compute $\mathcal{H}_1\left(\sum_{\lambda\in\Lambda}\left(\check{\mathbf{V}}^{q\mathsf{T}}_\lambda\,\check{\hat{\mathbf{r}}}^q_\lambda\right)^{\odot 2}\right)$ as indicated in [22];

4 Deduce $\hat{\mathbf{d}}_{q,\lambda}$ and $\check{\mathbf{m}}^z_q$ for each wavelength and agent;

5 Estimate $[\mathbf{m}_\lambda]_q$ with similar procedure than **Algorithm 1.2** and (24);

6 Estimate locally $\boldsymbol{\sigma}^{\mathrm{n}}_\lambda$ as indicated in [22] ;

Output: $\left\{\hat{\mathbf{m}}_\lambda, \hat{\mathbf{D}}_\lambda, \hat{\boldsymbol{\sigma}}^{\mathrm{n}}_\lambda\right\}_{\lambda\in\Lambda}$;

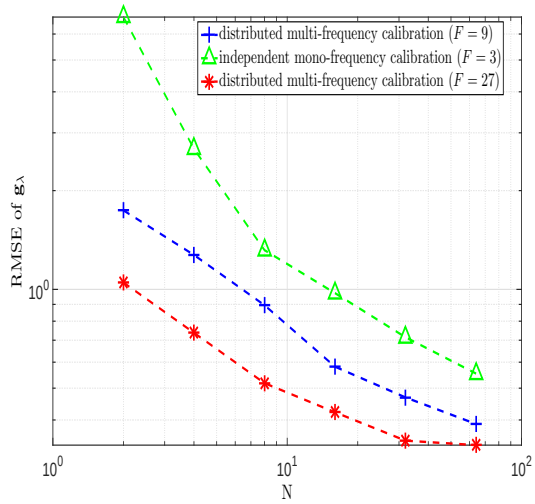
[12] T. Erseghe, A distributed and scalable processing method based upon ADMM, IEEE Signal Processing Letters 19 (9) (2012) 563–566.

[13] W. Shi, Q. Ling, K. Yuan, G. Wu, W. Yin, On the linear convergence of the ADMM in decentralized consensus optimization., IEEE Transactions on Signal Processing 62 (7) (2014) 1750–1761.

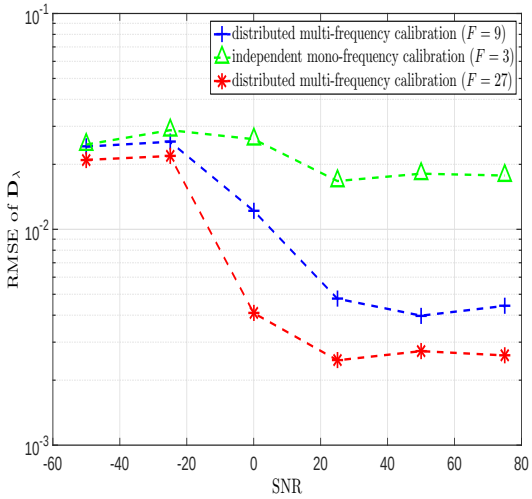
[14] J. F. C. Mota, J. M. F. Xavier, P. M. Q. Aguiar, M. Puschel, D-ADMM: a communication-efficient distributed algorithm for separable optimization, IEEE Transactions on Signal Processing 61 (10) (2013) 2718–2723.

[15] V. Ollier, M. N. El Korso, A. Ferrari, R. Boyer, P. Larzabal, Robust distributed calibration of radio interferometers with direction dependent distortions, accepted for publication by Elsevier Signal Processing.

[16] S. Yatawatta, Fine tuning consensus optimization for distributed radio interferometric calibration, in: 24th European Signal Processing Conference (EUSIPCO), Budapest, Hungary, 2016.



(a)

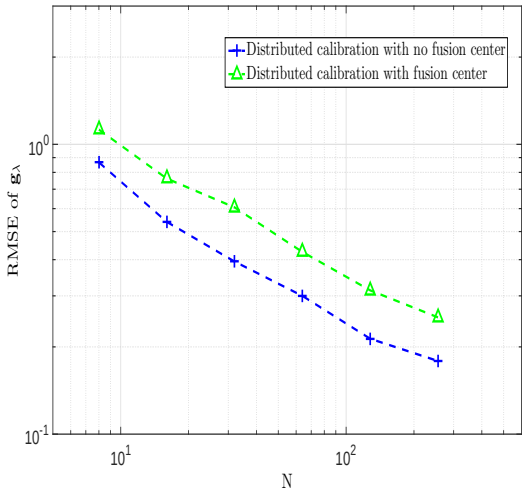


(b)

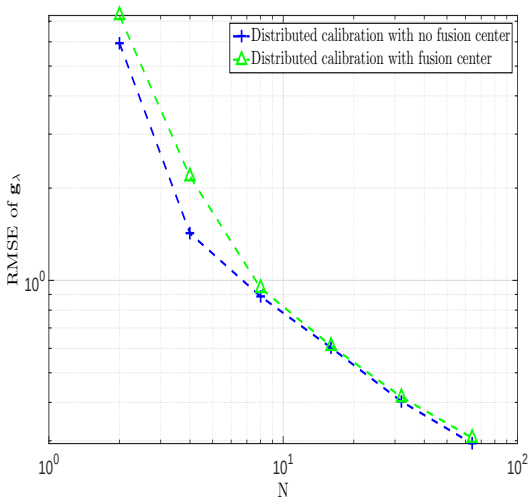
Figure 2: (a) RMSE on the undirectional gains as function of the number of samples N, (b) RMSE on the apparent source directions as function of the SNR.

[17] S. Yatawatta, S. Kazemi, S. Zaroubi, GPU accelerated nonlinear optimization in radio interferometric calibration, in: Innovative Parallel Computing (InPar), IEEE, San Jose, CA, 2012, pp. 1–6.

[18] D. Malioutov, M. Cetin, A. S. Willsky, A sparse signal reconstruction perspective for source localization with sensor arrays, IEEE Transac-



(c)

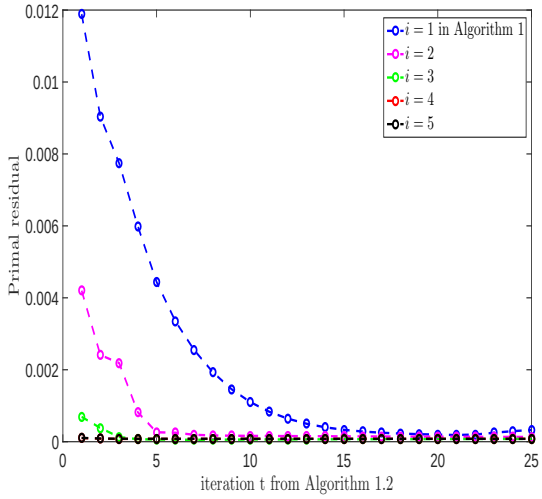


(d)

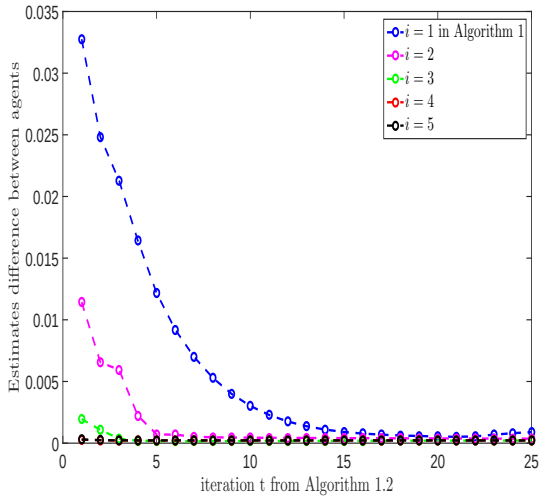
Figure 3: (a) Statistical comparison between different network topologies for same computational cost, (b) Statistical comparison between different network topologies for different computational cost.

tions on Signal Processing 53 (8) (2005) 3010–3022.

[19] C. Steffens, P. Parvazi, M. Pesavento, Direction finding and array calibration based on sparse reconstruction in partly calibrated arrays, in: 8th Sensor Array and Multichannel Signal Processing Workshop (IEEE SAM), 2014, pp. 21–24.



(e)



(f)

Figure 4: (a) Primal residual ϵ_p and (b) estimates difference ϵ_{DIFF} of the local consensus variables among agents as function of the iteration t in Algorithm 2, for different values of the iteration i in Algorithm 1.

[20] M. Haardt, M. Pesavento, F. Roemer, M. N. El Korso, Subspace methods and exploitation of special array structures, in: Electronic Reference in Signal Processing:Array and Statistical Signal Processing (M. Virberg, ed.), Vol. 3, Academic Press Library in Signal Processing, Elsevier Ltd., 2014.

[21] V. Ollier, M. N. El Korso, R. Boyer, P. Larzabal, M. Pesavento, Joint ML calibration and DOA estimation with separated arrays, in: International Conference on Acoustics, Speech and Signal Processing (IEEE ICASSP), Shanghai, China, 2016.

[22] 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.

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

[24] 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.

[25] S. Salvini, S. J. Wijnholds, Fast gain calibration in radio astronomy using alternating direction implicit methods: Analysis and applications, Astronomy & Astrophysics 571 (A97).

[26] T. Erseghe, D. Zennaro, E. Dall'Anese, L. Vangelista, Fast consensus by the alternating direction multipliers method, IEEE Transactions on Signal Processing 59 (11) (2011) 5523–5537.

[27] T. Blumensath, M. E. Davies, Normalized iterative hard thresholding: Guaranteed stability and performance, IEEE Journal of selected topics in signal processing 4 (2) (2010) 298–309.

[28] S. Patterson, Y. C. Eldar, I. Keidar, Distributed compressed sensing for static and time-varying networks, IEEE Transactions on Signal Processing 62 (19) (2014) 4931–4946.

[29] V. Ollier, M. N. El Korso, R. Boyer, P. Larzabal, M. Pesavento, Relaxed concentrated MLE for robust calibration of radio interferometers, in: 24th European Signal Processing Conference (EUSIPCO), Budapest, Hungary, 2016, pp. 280–284.

[30] S. J. Wijnholds, J. D. Bregman, A. J. Boonstra, Sky noise limited snapshot imaging in the presence of RFI with Lofar's Initial Test Station, Experimental Astronomy 17 (1-3) (2004) 35–42.

[31] 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.

[32] E. Ghadimi, A. Teixeira, I. Shames, M. Johansson, Optimal parameter selection for the alternating direction method of multipliers (ADMM): quadratic problems, IEEE Transactions on Automatic Control 60 (3) (2015) 644–658.