JOINT ML CALIBRATION AND DOA ESTIMATION WITH SEPARATED ARRAYS

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

[⋆] SATIE, UMR 8029, ENS Cachan, Université Paris-Saclay, France
[†] LEME, EA 4416, Université Paris-Ouest, Ville d'Avray, France
[⋄] L2S, UMR 8506, Université Paris-Sud, Gif-sur-Yvette, France
[‡] Communication Systems Group, Technische Universität Darmstadt, Darmstadt, Germany

ABSTRACT

This paper investigates parametric direction-of-arrival (DOA) estimation in a particular context: i) each sensor is characterized by an unknown complex gain and ii) the array consists of a collection of subarrays which are substantially separated from each other leading to a structured noise covariance matrix. We propose two iterative algorithms based on the maximum likelihood (ML) estimation method adapted to the context of joint array calibration and DOA estimation. Numerical simulations reveal that the two proposed schemes, the iterative ML (IML) and the modified iterative ML (MIML) algorithms for joint array calibration and DOA estimation, outperform the state of the art methods and the MIML algorithm reaches the Cramér-Rao bound for a low number of iterations.

Index Terms— Direction-of-arrival estimation, calibration, structured noise covariance matrix, maximum likelihood

1. INTRODUCTION

Direction-of-arrival (DOA) estimation [1, 2] is an important topic with a large number of applications: radar, satellite, mobile communications, radio astronomy, geophysics and underwater acoustics [3–5]. In order to achieve high resolution, it is common to use arrays with large aperture and/or a large number of sensors, in a specific noise environment. Considering spatially and temporally uncorrelated zero-mean Gaussian processes is a typical noise assumption but it may be violated in numerous applications, as in the context of sonar, where correlated or colored noise is required [6–9].

We consider here the case where the noise covariance matrix exhibits a particular (bloc-diagonal) structure [10] that differs from the classical assumption: spatially white uniform noise [11, 12] or non-uniform noise [13]. In our paper, we consider DOA estimation in large sensor arrays composed of multiple subarrays. Due to the large spacing between subarrays, we assume that the noise among sensors of different subarrays is statistically spatially independent. In a given subarray, however, the noise is spatially correlated between sensors. This entails a bloc-diagonal structure of the noise covariance matrix, linked to the sparsity of the array.

Apart from this noise assumption, we also consider that in realistic scenarios, due to miscalibration, the individual sensor outputs are generally subject to distortions by constant multiplicative complex factors (gains). These calibration errors are hardware related in our case, leading to different DOA independent sensor gains [10,14]. To precisely estimate these errors, we take advantage of the presence of calibration sources [15, 16] to simultaneously calibrate and estimate

DOAs. Our scenario is general and it can be adapted or extended to some practical applications as in the radio astronomy context [17] where the constant complex sensor gain assumption is common.

We use the conditional/deterministic model [18] for the signal sources. Nevertheless, following the same methodology, we can adapt the proposed algorithms to the case of unconditional/stochastic model [18–21]. The two parametric algorithms we present are based on the maximum likelihood (ML) estimation method, due to its good statistical performances. The size of the unknown parameter vector being large, we perform iterative optimization to make the ML estimation problem computationally tractable. Furthermore, the estimation performances are improved with the introduction of calibration sources in our scenario. To assess the performances [22], the Cramér-Rao bound (CRB) is used.

The notation used through this paper is the following: scalars, vectors and matrices are represented by italic lower-case, boldface lower-case and boldface upper-case symbols, respectively. The symbols $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^H$, $(\cdot)^\dagger$, $\mathrm{tr}\{\cdot\}$ and $\det\{\cdot\}$ denote, respectively, the transpose, the complex conjugate, the hermitian, the pseudo-inverse, the trace and determinant operator. The real and imaginary parts are referred to by $\Re\{\cdot\}$ and $\Im\{\cdot\}$. The operators $\mathrm{bdiag}\{\cdot\}$ and $\mathrm{diag}\{\cdot\}$ represent a bloc-diagonal and a diagonal matrix, respectively. A vector is by default a column vector and $\mathbf{I}$ is the identity matrix. The symbol $\odot$ denotes the Schur-Hadamard product, $\delta(.)$ is the Dirac's delta function and $\mathbf{E}_p$ is a $p \times p$ matrix filled with ones.

2. OBSERVATION MODEL

We consider D signal sources impinging on a linear (possibly not uniform) array of M sensors. The array response vector for each source $l = 1, \ldots, D$ is defined as [23]

$$\mathbf{a}(\theta_l) = [1, e^{-j2\pi f \frac{d_2}{c} \sin(\theta_l)}, \ldots, e^{-j2\pi f \frac{d_M}{c} \sin(\theta_l)}]^T \quad (1)$$

in which θ_l is the DOA of the lth source, f denotes the carrier frequency, c the propagation speed and d_k the inter-element spacing between the first and the kth sensor. We note as λ the wavelength of the incident wave. The output observation of the full array is given at each snapshot by

$$\mathbf{y}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t) + \mathbf{n}(t), \quad t = 1, \ldots, N \quad (2)$$

where N is the total number of snapshots, $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_D]^T$ is the DOAs vector, $\mathbf{s}(t) = [s_1(t), \ldots, s_D(t)]^T$ the signal source vector, $\mathbf{n}(t) = [n_1(t), \ldots, n_M(t)]^T$ the additive noise vector and $\mathbf{A}(\boldsymbol{\theta}) = [\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_D)]$ the array response matrix. In matrix notation, we have

$$\mathbf{Y} = \mathbf{A}(\boldsymbol{\theta})\mathbf{S} + \mathbf{N} \quad (3)$$

with $\mathbf{Y} = [\mathbf{y}(1), \dots, \mathbf{y}(N)]$, $\mathbf{S} = [\mathbf{s}(1), \dots, \mathbf{s}(N)]$ and $\mathbf{N} = [\mathbf{n}(1), \dots, \mathbf{n}(N)]$. In this work, the different assumptions that we consider are the following:

A1) Calibration sources: In a number of practical applications, the knowledge of one or multiple calibration sources is available [16, 24–26]. Without loss of generality, we consider the first P sources as calibration sources with known DOAs. Thus, the steering matrix is partitioned as

$$\mathbf{A}(\boldsymbol{\theta}) = \big[\mathbf{A}(\boldsymbol{\theta}_K), \mathbf{A}(\boldsymbol{\theta}_U) \big] \qquad (4)$$

in which $\boldsymbol{\theta}_K = [\theta_1, \dots, \theta_P]^T$ represents the known DOAs and $\boldsymbol{\theta}_U = [\theta_{P+1}, \dots, \theta_D]^T$ is the vector of the unknown DOAs. Likewise, the signal source matrix can be written as follows

$$\mathbf{S} = \big[\mathbf{S}_K^T, \mathbf{S}_U^T \big]^T \qquad (5)$$

in which $\mathbf{S}_K = [\mathbf{s}_K(1), \dots, \mathbf{s}_K(N)]$, $\mathbf{S}_U = [\mathbf{s}_U(1), \dots, \mathbf{s}_U(N)]$, $\mathbf{s}_K(t) = [s_1(t), \dots, s_P(t)]^T$ and $\mathbf{s}_U(t) = [s_{P+1}(t), \dots, s_D(t)]^T$.

A2) Complex unknown gains of each sensor: The instrumentation can introduce perturbations such as phase shifts, in particular due to the difference between sensor gains related to, e.g., receiver electronics. To correctly specify the model and avoid inaccurate estimations, calibration needs to be performed. This can be modeled using the following diagonal calibration matrix

$$\mathbf{G} = \mathrm{diag}\{\mathbf{g}\} \qquad (6)$$

where the vector $\mathbf{g} = [g_1, \dots, g_M]^T$ contains the different unknown complex gains [5, 10] which are modeled as DOA independent. Consequently, the observation matrix (3) can be rewritten as

$$\mathbf{Y} = \mathbf{G}\mathbf{A}(\boldsymbol{\theta}_K)\mathbf{S}_K + \mathbf{G}\mathbf{A}(\boldsymbol{\theta}_U)\mathbf{S}_U + \mathbf{N}. \qquad (7)$$

A3) Geometry of sensor subarrays: In our scenario, the sensor array is constituted of a set of L subarrays. Due to the large intersubarray distances with respect to the signal wavelength [27, 28], the noise is considered statistically independent between subarrays. Nevertheless, for a given subarray, sensors being closely spaced, the noise is assumed to be spatially correlated [10]. Thus, the noise covariance matrix denoted by $\boldsymbol{\Omega}$ has the following bloc-diagonal structure

$$\boldsymbol{\Omega} = \mathrm{bdiag}\{\boldsymbol{\Omega}_1, \dots, \boldsymbol{\Omega}_L\} \qquad (8)$$

in which $\boldsymbol{\Omega}_i$ is a $M_i \times M_i$ square matrix where M_i is the number of sensors in the ith subarray, such that $\sum_{i=1}^{L} M_i = M$.

Vector of unknown parameters: Let us consider a deterministic/conditional model for the signal sources and zero-mean complex circular Gaussian noise so that

$$\mathbf{y}(t) \sim CN\big(\mathbf{G}\mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t), \boldsymbol{\Omega}\big). \qquad (9)$$

Consequently, the vector of unknown parameters is given by

$$\begin{aligned} \boldsymbol{\eta} = & [\boldsymbol{\theta}_U^T, \mathbf{s}_U(1)^T, \dots, \mathbf{s}_U(N)^T, \{[\boldsymbol{\Omega}_1]_{h_1, l_1}\}_{l_1 \geq h_1}, \dots, \\ & \{[\boldsymbol{\Omega}_L]_{h_L, l_L}\}_{l_L \geq h_L}, \mathbf{g}^T]^T \end{aligned} \qquad (10)$$

in which for $i = 1, \dots, L$ and $h_i, l_i = 1, \dots, M_i$, $\{[\boldsymbol{\Omega}_i]_{h_i, l_i}\}_{l_i \geq h_i}$ represent all the non-zero elements in and above the diagonal of the noise covariance matrix.

3. PROPOSED ALGORITHMS

In this section, we propose two schemes for joint array calibration and DOA estimation, based on an iterative ML algorithm [10,13,29]. The iterative procedure allows us to obtain a closed-form expression of the unknown complex gains, the unknown signal sources and the structured noise covariance matrix. Indeed, the log-likelihood function is optimized w.r.t. each unknown parameter, while fixing the others. The different closed-form expressions obtained are mutually dependent and require an iterative updating procedure with initialization. For the estimation of the unknown DOAs, an optimization procedure needs to be performed.

The presence of calibration sources and the iterative procedure allow us to reduce a $(D + 2ND + \sum_{i=1}^{L} M_i^2 + 2M)$-dimensional optimization problem to a $(D - P)$-dimensional optimization problem. The main difference between the two proposed schemes lies in the estimation of the calibration matrix as it will be explained in the following.

3.1. Iterative ML (IML) algorithm for joint array calibration and DOA estimation

Let us denote $L(\boldsymbol{\eta})$ the log-likelihood function. Omitting the constant term, it becomes

$$L(\boldsymbol{\eta}) = -N \log\Big(\det\{\boldsymbol{\Omega}\} \Big) - \mathrm{tr}\Big\{ \mathbf{V}^H \boldsymbol{\Omega}^{-1} \mathbf{V} \Big\} \qquad (11)$$

in which

$$\mathbf{V} = \mathbf{Y} - \mathbf{G}\mathbf{A}(\boldsymbol{\theta})\mathbf{S}. \qquad (12)$$

1) Estimation of $\boldsymbol{\Omega}$: We take the derivative of $L(\boldsymbol{\eta})$ with respect to the elements $[\boldsymbol{\Omega}_i]_{h_i, l_i}$ for $h_i, l_i = 1, \dots, M_i$ and $i = 1, \dots, L$. During this operation, all the other unknown parameters remain fixed. We obtain for such derivation

$$\begin{aligned} & \frac{\partial L(\boldsymbol{\eta})}{\partial [\boldsymbol{\Omega}_i]_{h_i, l_i}} = \\ & - \mathrm{tr}\{ N\boldsymbol{\Omega}^{-1}\mathbf{e}_{i,h_i}\mathbf{e}_{i,l_i}^T - \mathbf{V}^H \boldsymbol{\Omega}^{-1}\mathbf{e}_{i,h_i}\mathbf{e}_{i,l_i}^T \boldsymbol{\Omega}^{-1}\mathbf{V}\} = \\ & - N\mathbf{e}_{i,l_i}^T \boldsymbol{\Omega}^{-1}\mathbf{e}_{i,h_i} + \mathbf{e}_{i,l_i}^T \boldsymbol{\Omega}^{-1}\mathbf{V}\mathbf{V}^H \boldsymbol{\Omega}^{-1}\mathbf{e}_{i,h_i} \end{aligned} \qquad (13)$$

where $[\mathbf{e}_{i,h_i}]_j = \delta(j - h_i)$ for $j, h_i = 1, \dots, M_i$ and $i = 1, \dots, L$. Equating (13) to zero, we obtain the estimations, $[\hat{\boldsymbol{\Omega}}_i]_{h_i, l_i}$, of all the non-zero elements of $\boldsymbol{\Omega}$. Due to the particular geometry of sensor subarrays, the exact covariance matrix is structured as in (8). Consequently, we introduce $\mathbf{E} = \mathrm{bdiag}\{\mathbf{E}_{M_1}, \dots, \mathbf{E}_{M_L}\}$ in order to impose this structure, and the estimation of $\boldsymbol{\Omega}$ becomes

$$\hat{\boldsymbol{\Omega}}_{\mathrm{IML}} = \frac{1}{N}(\mathbf{V}\mathbf{V}^H) \odot \mathbf{E}. \qquad (14)$$

One can note that the algorithm can be straightforwardly extended to the case of other (sparse) colored noise models.

2) Estimation of $\mathbf{G}$: We develop the second part of the r.h.s. of (11) as follows

$$\begin{aligned} \mathrm{tr}\{\mathbf{V}^H \boldsymbol{\Omega}^{-1}\mathbf{V}\} = \mathrm{tr}\{ & \mathbf{Y}^H \boldsymbol{\Omega}^{-1}\mathbf{Y} - \mathbf{Y}^H \boldsymbol{\Omega}^{-1}\mathbf{G}\mathbf{A}(\boldsymbol{\theta})\mathbf{S} - \\ & \mathbf{S}^H \mathbf{A}(\boldsymbol{\theta})^H \mathbf{G}^H \boldsymbol{\Omega}^{-1}\mathbf{Y} + \mathbf{S}^H \mathbf{A}(\boldsymbol{\theta})^H \mathbf{G}^H \boldsymbol{\Omega}^{-1}\mathbf{G}\mathbf{A}(\boldsymbol{\theta})\mathbf{S}\}. \end{aligned} \qquad (15)$$

Consequently, the derivation of $L(\boldsymbol{\eta})$ with respect to the elements g_i, for $i = 1, \dots, M$, has the following form

$$\begin{aligned} \frac{\partial L(\boldsymbol{\eta})}{\partial g_i} = \mathrm{tr}\{ & \mathbf{Y}^H \boldsymbol{\Omega}^{-1}\mathbf{e}_i\mathbf{e}_i^T \mathbf{A}(\boldsymbol{\theta})\mathbf{S} \\ & - \mathbf{S}^H \mathbf{A}(\boldsymbol{\theta})^H \mathbf{G}^H \boldsymbol{\Omega}^{-1}\mathbf{e}_i\mathbf{e}_i^T \mathbf{A}(\boldsymbol{\theta})\mathbf{S}\} \end{aligned} \qquad (16)$$

where $[\mathbf{e}_i]_j = \delta(i - j)$, for $i, j = 1, \ldots, M$. Let us denote $\mathbf{Z}_1 = \mathbf{A}(\boldsymbol{\theta})\mathbf{S}\mathbf{Y}^H\boldsymbol{\Omega}^{-1}$ and $\mathbf{Z}_2 = \mathbf{A}(\boldsymbol{\theta})\mathbf{S}\mathbf{S}^H\mathbf{A}(\boldsymbol{\theta})^H$. Equating (16) to zero while fixing the other terms leads us to solve the following linear system of equations

$$[\mathbf{Z}_1]_{i,i} = [\mathbf{Z}_2\mathbf{G}^H\boldsymbol{\Omega}^{-1}]_{i,i}, \quad i = 1, \ldots, M. \tag{17}$$

Furthermore, let us define the matrix $\mathbf{Z}_3$ such that $[\mathbf{Z}_3]_{l,i} = [\mathbf{Z}_2]_{l,i}^*[\boldsymbol{\Omega}^{-1}]_{i,l}^*$ for $l, i = 1, \ldots, M$. In an equivalent way, we can rewrite (17) as

$$[\mathbf{Z}_1]_{l,l} = \sum_{i=1}^{M}[\mathbf{Z}_3]_{l,i}^*g_i^*, \quad l = 1, \ldots, M. \tag{18}$$

Solving this linear system, we obtain for the IML algorithm

$$\hat{\mathbf{g}}_{\text{IML}} = \mathbf{Z}_3^{\dagger}\left[[\mathbf{Z}_1]_{1,1}, \ldots, [\mathbf{Z}_1]_{M,M}\right]^H. \tag{19}$$

Consequently, $\hat{\mathbf{G}}_{\text{IML}} = \text{diag}\{\hat{\mathbf{g}}_{\text{IML}}\}$.

3) Estimation of $\mathbf{S}_U$: Let us denote $\bar{\mathbf{A}}(\boldsymbol{\theta}_K) = \boldsymbol{\Omega}^{-\frac{1}{2}}\mathbf{G}\mathbf{A}(\boldsymbol{\theta}_K)$, $\bar{\mathbf{A}}(\boldsymbol{\theta}_U) = \boldsymbol{\Omega}^{-\frac{1}{2}}\mathbf{G}\mathbf{A}(\boldsymbol{\theta}_U)$, $\tilde{\mathbf{Y}} = \boldsymbol{\Omega}^{-\frac{1}{2}}\mathbf{Y}$, $\bar{\mathbf{Y}} = \tilde{\mathbf{Y}} - \bar{\mathbf{A}}(\boldsymbol{\theta}_K)\mathbf{S}_K$ and $\hat{\mathbf{R}} = \frac{1}{N}\bar{\mathbf{Y}}\bar{\mathbf{Y}}^H$. The second part of the r.h.s. of (11) can be written as

$$\text{tr}\{\mathbf{V}^H\boldsymbol{\Omega}^{-1}\mathbf{V}\} = \text{tr}\{\bar{\mathbf{Y}}^H\bar{\mathbf{Y}} - \bar{\mathbf{Y}}^H\bar{\mathbf{A}}(\boldsymbol{\theta}_U)\mathbf{S}_U - \mathbf{S}_U^H\bar{\mathbf{A}}(\boldsymbol{\theta}_U)^H\bar{\mathbf{Y}} + \mathbf{S}_U^H\bar{\mathbf{A}}(\boldsymbol{\theta}_U)^H\bar{\mathbf{A}}(\boldsymbol{\theta}_U)\mathbf{S}_U\}. \tag{20}$$

We take the derivative of $L(\boldsymbol{\eta})$ with respect to $[\mathbf{S}_U]_{h,l}$, for $h = 1, \ldots, (D - P)$ and $l = 1, \ldots, N$ and obtain the estimate in the least squares sense

$$\hat{\mathbf{S}}_U = \left(\bar{\mathbf{A}}(\boldsymbol{\theta}_U)^H\bar{\mathbf{A}}(\boldsymbol{\theta}_U)\right)^{-1}\bar{\mathbf{A}}(\boldsymbol{\theta}_U)^H\bar{\mathbf{Y}}. \tag{21}$$

4) Estimation of $\boldsymbol{\theta}_U$: Plugging (21) into (12), we obtain

$$\hat{\mathbf{V}} = \boldsymbol{\Omega}^{\frac{1}{2}}\mathbf{P}_{\bar{\mathbf{A}}(\boldsymbol{\theta}_U)}^{\perp}\bar{\mathbf{Y}} \tag{22}$$

in which $\mathbf{P}_{\bar{\mathbf{A}}(\boldsymbol{\theta}_U)}^{\perp} = \mathbf{I} - \bar{\mathbf{A}}(\boldsymbol{\theta}_U)\bar{\mathbf{A}}(\boldsymbol{\theta}_U)^{\dagger}$ is the projector orthogonal to the space spanned by the column vectors of $\bar{\mathbf{A}}(\boldsymbol{\theta}_U)$. Using (14) into (11), we can prove that $\text{tr}\{\mathbf{V}^H\hat{\boldsymbol{\Omega}}^{-1}\mathbf{V}\} = NM$. Omitting this constant term and considering the Hermitian symmetry of $\boldsymbol{\Omega}^{\frac{1}{2}}$ and $\mathbf{P}_{\bar{\mathbf{A}}(\boldsymbol{\theta}_U)}^{\perp}$, one can rewrite

$$L(\boldsymbol{\theta}, \hat{\mathbf{S}}_U, \hat{\boldsymbol{\Omega}}, \mathbf{G}) = -N\log(\det\{\mathbf{Z}\}) \tag{23}$$

where we note $\mathbf{Z} = (\boldsymbol{\Omega}^{\frac{1}{2}}\mathbf{P}_{\bar{\mathbf{A}}(\boldsymbol{\theta}_U)}^{\perp}\hat{\mathbf{R}}\mathbf{P}_{\bar{\mathbf{A}}(\boldsymbol{\theta}_U)}^{\perp}\boldsymbol{\Omega}^{\frac{1}{2}}) \odot \mathbf{E}$. The optimization process is thus

$$\hat{\boldsymbol{\theta}}_U = \arg\min_{\boldsymbol{\theta}_U}\left(\log(\det\{\mathbf{Z}\})\right). \tag{24}$$

Remark: To perform the optimization step of the cost function $F(\boldsymbol{\theta}_U) = \log(\det\{\mathbf{Z}\})$, we use a Newton-type algorithm [18], characterized by a quadratic convergence. For $l = 1, \ldots, (D - P)$, the gradient and the hessian are given by

$$\frac{\partial F}{[\boldsymbol{\theta}_U]_l} = \text{tr}\left\{\mathbf{Z}^{-1}\frac{\partial \mathbf{Z}}{[\boldsymbol{\theta}_U]_l}\right\} \text{ with}$$

$$\frac{\partial \mathbf{Z}}{[\boldsymbol{\theta}_U]_l} = \left(\boldsymbol{\Omega}^{\frac{1}{2}}\left(\frac{\partial \mathbf{P}_{\bar{\mathbf{A}}(\boldsymbol{\theta}_U)}^{\perp}}{\partial[\boldsymbol{\theta}_U]_l}\hat{\mathbf{R}}\mathbf{P}_{\bar{\mathbf{A}}(\boldsymbol{\theta}_U)}^{\perp} + \mathbf{P}_{\bar{\mathbf{A}}(\boldsymbol{\theta}_U)}^{\perp}\hat{\mathbf{R}}\frac{\partial \mathbf{P}_{\bar{\mathbf{A}}(\boldsymbol{\theta}_U)}^{\perp}}{\partial[\boldsymbol{\theta}_U]_l}\right)\boldsymbol{\Omega}^{\frac{1}{2}}\right) \odot \mathbf{E}, \text{ and}$$

$$\frac{\partial^2 F}{\partial[\boldsymbol{\theta}_U]_l^2} = \text{tr}\left\{-\mathbf{Z}^{-1}\frac{\partial \mathbf{Z}}{[\boldsymbol{\theta}_u]_l}\mathbf{Z}^{-1}\frac{\partial \mathbf{Z}}{[\boldsymbol{\theta}_u]_l} + \mathbf{Z}^{-1}\frac{\partial^2 \mathbf{Z}}{\partial[\boldsymbol{\theta}_u]_l^2}\right\} \text{ with}$$

$$\frac{\partial^2 \mathbf{Z}}{\partial[\boldsymbol{\theta}_U]_l^2} = \left(\boldsymbol{\Omega}^{\frac{1}{2}}\left(\frac{\partial^2 \mathbf{P}_{\bar{\mathbf{A}}(\boldsymbol{\theta}_U)}^{\perp}}{\partial[\boldsymbol{\theta}_U]_l^2}\hat{\mathbf{R}}\mathbf{P}_{\bar{\mathbf{A}}(\boldsymbol{\theta}_U)}^{\perp} + 2\frac{\partial \mathbf{P}_{\bar{\mathbf{A}}(\boldsymbol{\theta}_U)}^{\perp}}{\partial[\boldsymbol{\theta}_U]_l}\hat{\mathbf{R}}\frac{\partial \mathbf{P}_{\bar{\mathbf{A}}(\boldsymbol{\theta}_U)}^{\perp}}{\partial[\boldsymbol{\theta}_U]_l} + \mathbf{P}_{\bar{\mathbf{A}}(\boldsymbol{\theta}_U)}^{\perp}\hat{\mathbf{R}}\frac{\partial^2 \mathbf{P}_{\bar{\mathbf{A}}(\boldsymbol{\theta}_U)}^{\perp}}{\partial[\boldsymbol{\theta}_U]_l^2}\right)\boldsymbol{\Omega}^{\frac{1}{2}}\right) \odot \mathbf{E}.$$

IML algorithm:

input : $\mathbf{Y}$, $\mathbf{E}$, $\mathbf{S}_K$, $\mathbf{A}(\boldsymbol{\theta}_K)$, N, M, L, D, P
output : estimates of $\boldsymbol{\theta}_U$, $\mathbf{S}_U$, $\boldsymbol{\Omega}_{\text{IML}}$ and $\mathbf{G}_{\text{IML}}$
initialize: $\boldsymbol{\Omega}_{\text{IML}} = \mathbf{I}$, $\mathbf{G}_{\text{IML}} = \mathbf{I}$
while *stop criterion unreached* **do**

1   Estimation of $\boldsymbol{\theta}_U$ by (24)
2   Estimation of $\mathbf{S}_U$ by (21)
3   Estimation of $\boldsymbol{\Omega}_{\text{IML}}$ by (14)
4   Estimation of $\mathbf{G}_{\text{IML}}$ by (19)

end

3.2. Modified iterative ML (MIML) algorithm for joint array calibration and DOA estimation

In practical scenario, calibration is performed with respect to powerful radiating sources. The remaining $(D - P)$ sources, see the partitioning model in (4) and (5), have a negligible power in comparison with these calibration sources. Consequently, the distribution of the observations at each snapshot can be approximated by

$$\mathbf{y}(t) \sim CN\left(\mathbf{G}\mathbf{A}(\boldsymbol{\theta}_K)\mathbf{s}_K(t), \boldsymbol{\Omega}\right). \tag{25}$$

The key idea of this alternative method is to estimate the calibration parameters and the noise covariance matrix based on the calibration sources at the first step. Once these parameters are estimated, the second step consists in estimating the unknown DOAs and signal sources. For the MIML algorithm, we only present the results but the methodology is the same as in section 3.1. Taking into account (25) and the previous calculus performed to obtain (19), we can estimate $\mathbf{G}$ by solving the following system

$$\hat{\mathbf{g}}_{\text{MIML}} = \tilde{\mathbf{Z}}_3^{\dagger}\left[[\tilde{\mathbf{Z}}_1]_{1,1}, \ldots, [\tilde{\mathbf{Z}}_1]_{M,M}\right]^H \tag{26}$$

where $[\tilde{\mathbf{Z}}_3]_{l,i} = [\tilde{\mathbf{Z}}_2]_{l,i}^*[\boldsymbol{\Omega}^{-1}]_{i,l}^*$ for $l, i = 1, \ldots, M$. Here, we have $\tilde{\mathbf{Z}}_1 = \mathbf{A}(\boldsymbol{\theta}_K)\mathbf{S}_K\mathbf{Y}^H\boldsymbol{\Omega}^{-1}$ and $\tilde{\mathbf{Z}}_2 = \mathbf{A}(\boldsymbol{\theta}_K)\mathbf{S}_K\mathbf{S}_K^H\mathbf{A}(\boldsymbol{\theta}_K)^H$. Consequently, $\hat{\mathbf{G}}_{\text{MIML}} = \text{diag}\{\hat{\mathbf{g}}_{\text{MIML}}\}$. Following the same methodology to obtain (14), the estimate of $\boldsymbol{\Omega}$ is given by

$$\hat{\boldsymbol{\Omega}}_{\text{MIML}} = \frac{1}{N}(\mathbf{V}_K\mathbf{V}_K^H) \odot \mathbf{E} \tag{27}$$

in which $\mathbf{V}_K = \mathbf{Y} - \mathbf{G}\mathbf{A}(\boldsymbol{\theta}_K)\mathbf{S}_K$.

The estimation of the other parameters $\boldsymbol{\theta}_U$ and $\mathbf{S}_U$ is then performed with the same expressions as in the first proposed scheme and taking into account the estimations $\hat{\mathbf{G}}_{\text{MIML}}$ and $\hat{\boldsymbol{\Omega}}_{\text{MIML}}$ obtained with (26) and (27).

As our simulations will show, the MIML algorithm reaches convergence faster than the IML algorithm. Furthermore, the latter requires greater computational complexity, due to the presence of more estimation steps in the loop.
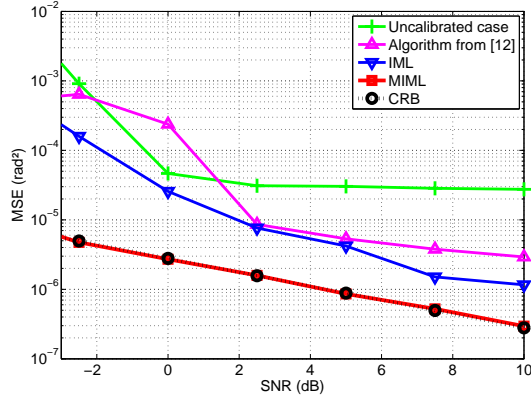
MIML algorithm:

    input   : $\mathbf{Y}, \mathbf{E}, \mathbf{S}_K, \mathbf{A}(\boldsymbol{\theta}_K), N, M, L, D, P$
    output : estimates of $\boldsymbol{\theta}_U, \mathbf{S}_U, \boldsymbol{\Omega}_{\text{MIML}}$ and $\mathbf{G}_{\text{MIML}}$
    initialize: $\boldsymbol{\Omega}_{\text{MIML}} = \mathbf{I}$
    while *stop criterion unreached* **do**
1 | Estimation of $\mathbf{G}_{\text{MIML}}$ by (26)
2 | Estimation of $\boldsymbol{\Omega}_{\text{MIML}}$ by (27)
    end
3 Estimation of $\boldsymbol{\theta}_U$ by (24)
**4** Estimation of $\mathbf{S}_U$ by (21)

Fig. 1. Comparison between the IML and the MIML algorithms, optimizing with Newton.

4. NUMERICAL SIMULATIONS

In the following simulations, we consider two sources, a calibration source at $\theta_1 = 7°$ and an unknown source at $\theta_2 = 15°$, as well as 300 Monte-Carlo and $N = 160$ snapshots. The full array is composed of 3 linear subarrays with 4, 3 and 2 sensors in each. The inter-element spacing is $\frac{\lambda}{2}$ in each subarray, 3λ and $\frac{7\lambda}{2}$ between the three successive subarrays. We consider a noise covariance matrix $\boldsymbol{\Omega}$ with an identical noise power for each sensor of the same subarray. The amplitude gains and the phases are generated respectively uniformly on $[0, 1]$ and $[0, 2\pi]$. The signal-to-noise ratio (SNR) is denoted by:

$$\text{SNR} = \frac{\sum_{t=1}^{N} \|\mathbf{s}_u(t)\|^2}{NM} \sum_{i=1}^{M} \frac{1}{[\boldsymbol{\Omega}]_{i,i}}. \tag{28}$$

In Fig. 1, we plot the mean square error (MSE) vs. SNR, for both schemes, as well as for the uncalibrated case, meaning that the observations are given by (7) but estimation of matrix $\mathbf{G}$ is not performed in the estimation process, it is maintained equal to $\mathbf{I}$. In this case, we notice the degradation of performances, moreover the MSE is no longer decreasing from a certain value of the SNR. We also plot the MSE of the algorithm proposed in [10], for which the presence of calibration sources is not taken into account. As expected, the presence of a calibration source enables to achieve better performances, particularly with the MIML algorithm. The different computation times for the two exposed methods are 141.473 seconds for the IML algorithm (4 iterations) and 42.841 seconds for the MIML algorithm
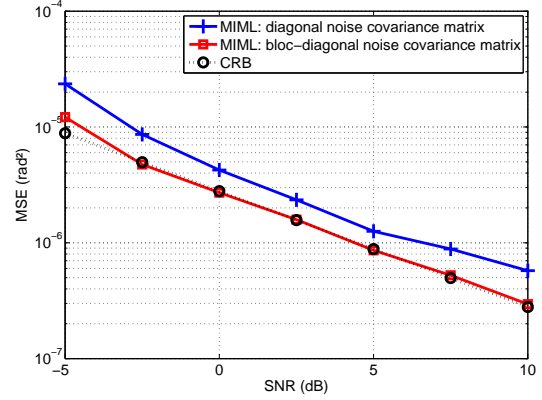
Fig. 2. Effect of a priori knowledge about the structure of the noise covariance matrix, for the MIML algorithm.

(2 iterations).

Finally, the Cramér-Rao bound (CRB) [10,30,31] was plotted. It is noticed that the best compromise between computation time and accuracy of estimation is achieved with the MIML algorithm. Indeed, we observe that numerically the MSE asymptotically reaches the CRB. Such performances are due to the estimation of the calibration matrix $\mathbf{G}$ which is performed separately from the estimation of $\boldsymbol{\theta}_U$ and mainly depends on the calibration sources ($\mathbf{A}(\boldsymbol{\theta}_K)$ and $\mathbf{S}_K$), contrary to the first algorithm where it depends on the unknown sources as well ($\mathbf{A}(\boldsymbol{\theta})$ and $\mathbf{S}$). The IML algorithm requires more iterations to have better accuracy in the estimation.

Finally, Fig. 2 represents the MSE of the MIML algorithm for the two following cases: i) taking into account the true structure of the noise covariance matrix, and ii) assuming that the noise covariance matrix is diagonal. In the two cases, the observations are generated using the true noise covariance matrix which is structured as described by (8). As expected, such misspecification leads to a higher MSE (case ii) which shows the importance of taking into account the spatial correlation due to the array geometry.

5. CONCLUSION

In this paper, we proposed two iterative algorithms for joint calibration and DOA estimation. They are based on the ML estimation method and are applied in a particular context: some calibration sources are present, the sensors are characterized by unknown DOA independent complex gains and the noise covariance matrix has a bloc-diagonal structure. The MIML algorithm outperforms the IML algorithm and numerically attains the CRB for a low number of iterations. The proposed algorithm is general and can be adapted, for example, in the context of radio astronomy.

6. REFERENCES

[1] M. Haardt, M. Pesavento, F. Roemer, and 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.)*. Academic Press Library in Signal Processing, Elsevier Ltd., 2014, vol. 3.

[2] P. Stoica and K. C. Sharman, "Maximum likelihood methods for direction-of-arrival estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, no. 7, pp. 1132–1143, 1990.

[3] L. C. Godara, "Application of antenna arrays to mobile communications, part II: Beam-forming and direction-of-arrival considerations," *Proceedings of the IEEE*, vol. 85, no. 8, pp. 1195–1245, 1997.

[4] K. T. Wong and M. D. Zoltowski, "Root-MUSIC-based azimuth-elevation angle-of-arrival estimation with uniformly spaced but arbitrarily oriented velocity hydrophones," *IEEE Trans. Signal Processing*, vol. 47, no. 12, pp. 3250–3260, 1999.

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

[6] H. Ye and R. D. DeGroat, "Maximum likelihood DOA estimation and asymptotic Cramér-Rao bounds for additive unknown colored noise," *IEEE Trans. Signal Processing*, vol. 43, no. 4, pp. 938–949, 1995.

[7] B. Friedlander and A. J. Weiss, "Direction finding using noise covariance modeling," *IEEE Trans. Signal Processing*, vol. 43, no. 7, pp. 1557–1567, 1995.

[8] P. Stoica, M. Viberg, K. M. Wong, and Q. Wu, "Maximum-likelihood bearing estimation with partly calibrated arrays in spatially correlated noise fields," *IEEE Trans. Signal Processing*, vol. 44, no. 4, pp. 888–899, 1996.

[9] M. Li and Y. Lu, "Maximum likelihood DOA estimation in unknown colored noise fields," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 44, no. 3, pp. 1079–1090, 2008.

[10] S. Vorobyov, A. B. Gershman, and K. M. Wong, "Maximum likelihood direction-of-arrival estimation in unknown noise fields using sparse sensor arrays," *IEEE Trans. Signal Processing*, vol. 53, no. 1, pp. 34–43, 2005.

[11] P. Stoica and A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, no. 10, pp. 1783–1795, 1990.

[12] M. N. El Korso, G. Bouleux, R. Boyer, and S. Marcos, "Sequential estimation of the range and the bearing using the zero-forcing MUSIC approach," in *17th European Signal Processing Conf.*, Glasgow, Scotland, 2009.

[13] M. Pesavento and A. B. Gershman, "Maximum-likelihood direction-of-arrival estimation in the presence of unknown nonuniform noise," *IEEE Trans. Signal Processing*, vol. 49, no. 7, pp. 1310–1324, 2001.

[14] A. J. Weiss and B. Friedlander, "Eigenstructure methods for direction finding with sensor gain and phase uncertainties," *Circuits, Syst., Signal Process.*, vol. 9, no. 3, pp. 271–300, 1990.

[15] J. Pierre and M. Kaveh, "Experimental performance of calibration and direction-finding algorithms," in *Proc. of IEEE Int. Conf. Acoust., Speech, Signal Processing*, Toronto, Ontario, Canada, 1991, pp. 1365–1368.

[16] B. C. Ng and C. M. S. See, "Sensor-array calibration using a maximum-likelihood approach," *IEEE Trans. Antennas Propag.*, vol. 44, no. 6, pp. 827–835, 1996.

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

[18] B. Ottersten, M. Viberg, P. Stoica, and A. Nehorai, "Exact and large sample maximum likelihood techniques for parameter estimation and detection in array processing," in *Radar Array Processing*, S. Haykin, J. Litva, and T. J. Shepherd, Eds. Berlin: Springer-Verlag, 1993, ch. 4, pp. 99–151.

[19] P. Stoica and A. Nehorai, "On the concentrated stochastic likelihood function in array signal processing," *Circuits, Syst., Signal Process.*, vol. 14, no. 5, pp. 669–674, 1995.

[20] A. Gershman, P. Stoica, M. Pesavento, and E. G. Larsson, "Stochastic Cramér-Rao bound for direction estimation in unknown noise fields," *IEE Proceedings-Radar, Sonar and Navigation*, vol. 149, no. 1, pp. 2–8, 2002.

[21] C. E. Chen, F. Lorenzelli, R. E. Hudson, and K. Yao, "Stochastic maximum-likelihood DOA estimation in the presence of unknown nonuniform noise," *IEEE Trans. Signal Processing*, vol. 56, no. 7, pp. 3038–3044, 2008.

[22] M. N. El Korso, R. Boyer, A. Renaux, and S. Marcos, "Conditional and unconditional Cramér-Rao bounds for near-field source localization," *IEEE Trans. Signal Processing*, vol. 58, no. 5, pp. 2901–2907, 2010.

[23] H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," *IEEE Signal Processing Mag.*, vol. 13, no. 4, pp. 67–94, 1996.

[24] R. Boyer and G. Bouleux, "Oblique projections for direction-of-arrival estimation with prior knowledge," *IEEE Trans. Signal Processing*, vol. 56, no. 4, pp. 1374–1387, 2008.

[25] G. Bouleux, P. Stoica, and R. Boyer, "An optimal prior-knowledge-based DOA estimation method," in *17th European Signal Processing Conf.*, Glasgow, United Kingdom, 2009, pp. 869–873.

[26] M. N. El Korso, R. Boyer, and S. Marcos, "Fast sequential source localization using the projected companion matrix approach," in *Proceedings of the 3rd IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, Aruba, Dutch Antilles, The Netherlands, 2009, pp. 245–248.

[27] M. D. Zoltowski and K. T. Wong, "Closed-form eigenstructure-based direction finding using arbitrary but identical subarrays on a sparse uniform Cartesian array grid," *IEEE Trans. Signal Processing*, vol. 48, no. 8, pp. 2205–2210, 2000.

[28] M. Pesavento, A. B. Gershman, and K. M. Wong, "Direction of arrival estimation in partly calibrated time-varying sensor arrays," in *Proc. of IEEE Int. Conf. Acoust., Speech, Signal Processing*, Salt Lake City, UT, 2001, pp. 3005–3008.

[29] X. Zhang, M. N. El Korso, and M. Pesavento, "Maximum likelihood and maximum a posteriori direction-of-arrival estimation in the presence of SIRP noise," in *Proc. of IEEE Int. Conf. Acoust., Speech, Signal Processing*, Shanghai, China, 2016.

[30] P. Stoica and R. L. Moses, *Spectral analysis of signals*. Pearson/Prentice Hall Upper Saddle River, NJ, 2005.

[31] A. L. Matveyev, A. B. Gershman, and J. F. Böhme, "On the direction estimation Cramér-Rao bounds in the presence of uncorrelated unknown noise," *Circuits, Syst., Signal Process.*, vol. 18, no. 5, pp. 479–487, 1999.