



HAL
open science

Expectation-Maximization Based Direction of Arrival Estimation Under a Mixture of Noise

Yassine Mhiri, Mohammed Nabil El Korso, Lucien Bacharach, Arnaud Breloy,
Pascal Larzabal

► **To cite this version:**

Yassine Mhiri, Mohammed Nabil El Korso, Lucien Bacharach, Arnaud Breloy, Pascal Larzabal. Expectation-Maximization Based Direction of Arrival Estimation Under a Mixture of Noise. EU-SIPCO, Aug 2021, Dublin (virtual), Ireland. 10.23919/EUSIPCO54536.2021.9615971 . hal-03156738

HAL Id: hal-03156738

<https://hal.science/hal-03156738>

Submitted on 2 Mar 2021

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.

Expectation-Maximization Based Direction of Arrival Estimation Under a Mixture of Noise

Yassine Mhiri*, Mohammed Nabil El Korso[†], Lucien Bacharach*, Arnaud Breloy[†] and Pascal Larzabal*

* SATIE, Université Paris Saclay, 91190 Gif-Sur-Yvette, France

[†] Paris-Nanterre University/LEME, 92410 Ville d'Avray, France

Abstract—In this paper, we propose a novel scheme for direction of arrival estimation in the presence of a noise which is a combination of white Gaussian distributed noise and spherically invariant random distributed noise. Such combination arises in practical scenario, in which, the Gaussian component represents the thermal noise (i.e., the internal noise), whereas, the spherically invariant random process represents the possible presence of outliers and/or non-homogeneities of the environment (i.e., interference, clutter, jammers). The classical direction of arrival estimation using the maximum likelihood is computationally intractable. In order to overcome this drawback while maintaining a fine accuracy and taking into account the presence of the two noise components, we design an expectation-maximization algorithm. Finally, numerical simulations show that the proposed algorithm outperforms the state-of-the-art.

Index Terms—Direction of arrival, robust parameter estimation, array processing, expectation-maximization.

I. INTRODUCTION

Direction of arrival (DoA) estimation is an important topic with a large panel of applications, such as seismology, digital communications, sonar, radar, etc. [1]–[3]. A plethora of algorithms have been proposed to deal with conditional (i.e., when the signal source is assumed to be deterministic and unknown) and unconditional (i.e., when the signal source are assumed to follow a Gaussian distribution) DoA estimation [4], [5]. In particular, the majority of the proposed algorithms assume that the noise follows a Gaussian distribution. In this context, the conventional maximum likelihood (ML) estimator was largely studied [6]. In the same vein, sub-optimal algorithms, based on the second order statistics, aiming to overcome the computational drawback of the ML estimator have been proposed in the last three decades [7]–[10]. The latter reveals to be adequate when the noise is Gaussian distributed. Such assumption, which is in general a good approximation, relies on the central limit theorem. Nevertheless, in certain scenario, as high resolution sensing systems, low-grazing-angle radar context, etc., the Gaussian assumption is not valid any more [11]–[15]. Such mismatch leads to a dramatic performance loss when applying the conventional ML estimator or the aforementioned sub-optimal algorithms under a non-Gaussian scenario [16]. To account for such possible non-Gaussian distributed noise, the spherically invariant random process (SIRP) has been introduced in the literature [17]. The SIRP has become popular since it owns a great flexibility allowing to gather several distributions, e.g., Gaussian, K, Student's, Cauchy distributions, etc. [16] [27]. Specifically, the SIRP can

be described as the product of a Gaussian component, the so-called speckle, and a positive random variable, named the texture.

In this paper, we aim at proposing a novel DoA estimation in the context of a noise which is a combination of a Gaussian component and a SIRP component. This modeling is of interest in practical scenario, since, recent sensing systems are faced to two kinds of noise. The first one, is an internal noise which is generally a thermal noise modeled as a Gaussian distributed noise thanks to the central limit theorem. The second component is external and reflects the possible presence of outliers and/or non-homogeneities of the environment. Consequently, the presence of such two components lead to a noise mixture of Gaussian and SIRP distributed noise. Solving such inference problem by a brute force ML estimator is unfeasible due to its high computational cost. In our work, we propose the use of the expectation-maximization (EM) algorithm. The EM algorithm is an iterative algorithm that aims to achieve the ML estimate in a computable way [18]. In order to design the EM, we select in a proper way the complete data which is constituted of the unknown realization of the external noise and the observations. Interestingly, such choice leads to closed-form expressions regarding the E-step, mostly closed-form expressions in the M-step (apart from the optimization part regarding the DoA).

Notation: The following notations are used through the paper. Matrices and vectors are represented by bold uppercase and bold lowercase letters, respectively. Vectors are by default in column orientation. $(\cdot)^T$, $(\cdot)^H$, $Tr\{\cdot\}$ and $|\cdot|$ denote the transpose, the conjugate transpose, the trace and the determinant of a matrix, respectively. $E\{\cdot\}$ represents the expectation and the weighted norm is defined as $\|\mathbf{w}\|_{\mathbf{W}}^2 = \mathbf{w}^H \mathbf{W}^{-1} \mathbf{w}$. Finally, $\mathcal{CN}(\mathbf{m}, \mathbf{C})$ denotes a complex circular Gaussian distribution with mean \mathbf{m} and covariance matrix \mathbf{C} and $\text{invGamma}(a, b)$ represents an inverse Gamma distribution parameterized by a scale parameter, a , and shape parameter b .

II. MODEL SETUP

Consider an array (possibly non uniform) of n sensors that receives a deterministic and unknown signal denoted $s(t)$. The output array signal reads

$$\mathbf{y}(t) = \mathbf{a}(\boldsymbol{\theta})s(t) + \mathbf{n}(t) \quad t = 1, \dots, T$$

in which $\mathbf{a}(\boldsymbol{\theta})$ is the steering vector indexed by the unknown parameter of interest $\boldsymbol{\theta}$ ($\boldsymbol{\theta}$ can denote the bearing/DoA of a

far field source observed by a linear array, the concatenation of the azimuth and the elevation of a planar array, or the bearing and the range in a near-field scenario, ect. [2]). In addition, $\mathbf{n}(t)$ denotes the noise component and T the total number of snapshots. As stated before, the noise is assumed to be a mixture of a thermal noise, which follows a Gaussian distribution denoted $\mathbf{t}(t)$, and an additive noise which represents the possible presence of outliers $\mathbf{o}(t)$. The latter is assumed to follow a spherically invariant random process [17]. Consequently, the mixture of noise reads

$$\mathbf{n}(t) = \mathbf{t}(t) + \mathbf{o}(t)$$

in which the i.i.d. thermal noise $\mathbf{t}(t) \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I})$, whereas the SIRV term $\mathbf{o}(t) = \sqrt{\tau(t)} \mathbf{x}(t)$ where the texture parameter $\tau(t) \sim p_\tau$ and the speckle parameter reads $\mathbf{x}(t) \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma})$. It is worth mentioning that the use of the spherically invariant random process is motivated by its capability to model non-Gaussian heavy tailed distributed noise, but also to adaptively consider Gaussian noise in the extreme case when there are no outliers, thus gathering a wide range of distribution (the Student's t, K, Gaussian and Cauchy distributions just to cite a few [16].) In the following, we consider the texture parameter as deterministic but unknown. By doing this, we do not specify the distribution of the texture parameter (which is commonly unknown in practice) and thus, we ensure flexibility and robustness of the proposed algorithm.

Based on the hypothesis above, the maximum likelihood estimator reads as the following minimization

$$\hat{\phi}_{\text{ML}} = \arg \max_{\phi} \mathcal{L}_{\text{ML}}(\mathbf{y}(1), \dots, \mathbf{y}(T) | \phi) \quad (1)$$

in which the negative log likelihood is

$$\begin{aligned} \mathcal{L}_{\text{ML}}(\mathbf{y}(1), \dots, \mathbf{y}(T) | \phi) &= \sum_1^{t=T} \log |\tau(t) \mathbf{\Sigma} + \sigma^2 \mathbf{I}| \\ &+ (\mathbf{y}(t) - \mathbf{a}(\boldsymbol{\theta})s(t))^H (\tau(t) \mathbf{\Sigma} + \sigma^2 \mathbf{I})^{-1} (\mathbf{y}(t) - \mathbf{a}(\boldsymbol{\theta})s(t)), \end{aligned}$$

the unknown parameter being

$$\phi = \left\{ \boldsymbol{\theta}, \{s(t)\}_{t=1, \dots, T}, \zeta, \sigma^2, \{\tau(t)\}_{t=1, \dots, T} \right\}$$

where ζ is the concatenation of the non redundant elements in $\mathbf{\Sigma}$.

From above, it is clear that a brute force maximization of (1) is unfeasible. In the following, we make use of the expectation maximization algorithm in order to approach (asymptotically) the maximum likelihood estimate $\hat{\phi}_{\text{ML}}$.

III. THE PROPOSED EM BASED SCHEME

The EM algorithm is an iterative scheme that attempts to find the maximum likelihood in an elegant manner [18]. Specifically, it consists of two steps. In the first step (i.e., the E-step (expectation)), we compute the expectation of the log likelihood of the complete data conditioned to the observation (i.e., the incomplete data) and the estimate from the previous m -th iteration denoted by $\hat{\phi}^{(m)}$. In the second step (i.e., the

M-step (maximization)), we maximize the previously derived expectation w.r.t. to ϕ for a fixed $\hat{\phi}^{(m)}$ in order to obtain a new refinement $\hat{\phi}^{(m+1)}$. Then, we iterate these two steps until convergence.

The complete data should be carefully chosen in order to make the M-step as simple as possible (and if possible in closed-form expressions) but also in a way that we would be able to derive, analytically, the expectation of the log likelihood of the complete data. In our case, it is of interest to consider the complete data as $\mathbf{z}(t) = \{\mathbf{y}(t), \mathbf{o}(t)\}$ for $t = 1, \dots, T$. Consequently, the complete likelihood reads

$$\begin{aligned} p(\mathbf{z}(t) | \phi) &= p(\mathbf{y}(t) | \mathbf{o}(t), \boldsymbol{\theta}, s(t), \sigma^2) p(\mathbf{o}(t) | \zeta, \tau(t)) \\ &= \frac{1}{\sigma^{2n} \tau(t)^n |\mathbf{\Sigma}|} \exp^{-\|\mathbf{y}(t) - \mathbf{a}(\boldsymbol{\theta})s(t) - \mathbf{o}(t)\|_{\sigma^2}^2 - \|\mathbf{o}(t)\|_{\tau(t)\mathbf{\Sigma}}^2}. \end{aligned}$$

Thus,

$$\begin{aligned} \log p(\mathbf{z}(t) | \phi) &= -n \log |\sigma^2| - n \log |\tau(t)| - \log |\mathbf{\Sigma}| \\ &- \|\mathbf{y}(t) - \mathbf{a}(\boldsymbol{\theta})s(t) - \mathbf{o}(t)\|_{\sigma^2}^2 - \|\mathbf{o}(t)\|_{\tau(t)\mathbf{\Sigma}}^2 \end{aligned}$$

In the following, we compute the expectation of the latter complete log likelihood.

A. E-step

First, let us note that the observations and the outliers term are jointly Gaussian, i.e.,

$$\begin{pmatrix} \mathbf{y}(t) - \mathbf{a}(\boldsymbol{\theta})s(t) \\ \mathbf{o}(t) \end{pmatrix} | \phi \sim \mathcal{CN}(\mathbf{0}, \bar{\mathbf{\Sigma}}) \quad (2)$$

with

$$\bar{\mathbf{\Sigma}} = \begin{pmatrix} \tau(t)\mathbf{\Sigma} + \sigma^2 \mathbf{I} & \tau(t)\mathbf{\Sigma} \\ \tau(t)\mathbf{\Sigma} & \tau(t)\mathbf{\Sigma} \end{pmatrix}$$

Consequently, it is easy to see that (e.g., [19])

$$\mathbf{o}(t) | \mathbf{y}(t), \phi \sim \mathcal{CN}(\mathbf{m}(t)_{\mathbf{o}(t) | \mathbf{y}(t)}, \text{Cov}_{\mathbf{o}(t) | \mathbf{y}(t)})$$

in which

$$\mathbf{m}(t)_{\mathbf{o}(t) | \mathbf{y}(t)} = \tau(t) \mathbf{\Sigma}^H (\tau(t) \mathbf{\Sigma} + \sigma^2 \mathbf{I})^{-1} (\mathbf{y}(t) - \mathbf{a}(\boldsymbol{\theta})s(t))$$

and

$$\text{Cov}_{\mathbf{o}(t) | \mathbf{y}(t)} = \tau(t) \mathbf{\Sigma} - \tau(t)^2 \mathbf{\Sigma}^H (\tau(t) \mathbf{\Sigma} + \sigma^2 \mathbf{I})^{-1} \mathbf{\Sigma}.$$

Meaning that the expectation of the sufficient statistics conditioned to the observation $\mathbf{y}(t)$ and the previous estimate $\hat{\phi}^{(m)}$ reads

$$\hat{\mathbf{o}}(t) = E_{\mathbf{o}(t) | \mathbf{y}(t), \hat{\phi}^{(m)}} \{\mathbf{o}(t)\} = \quad (3)$$

$$\begin{aligned} &\hat{\tau}(t)^{(m)} \left(\hat{\mathbf{\Sigma}}^{(m)} \right)^H \left(\hat{\tau}(t)^{(m)} \hat{\mathbf{\Sigma}}^{(m)} + \widehat{\sigma^2}^{(m)} \mathbf{I} \right)^{-1} \\ &\left(\mathbf{y}(t) - \mathbf{a}(\hat{\boldsymbol{\theta}}^{(m)})s(t) \right) \end{aligned} \quad (4)$$

and

$$\begin{aligned} \widehat{\mathbf{o}(t)\mathbf{o}(t)^H} &= E_{\mathbf{o}(t) | \mathbf{y}(t), \hat{\phi}^{(m)}} \{\mathbf{o}(t)\mathbf{o}(t)^H\} = \hat{\tau}^{(m)}(t) \hat{\mathbf{\Sigma}}^{(m)} \\ &- \left(\hat{\tau}(t)^{(m)} \right)^2 \left(\hat{\mathbf{\Sigma}}^{(m)} \right)^H \left(\hat{\tau}(t)^{(m)} \hat{\mathbf{\Sigma}}^{(m)} + \widehat{\sigma^2}^{(m)} \mathbf{I} \right)^{-1} \hat{\mathbf{\Sigma}}^{(m)} \\ &+ \hat{\mathbf{o}}(t) \hat{\mathbf{o}}(t)^H. \end{aligned} \quad (5)$$

Consequently, the so-called surrogate function reduces to

$$\begin{aligned} Q_t(\phi|\hat{\phi}^{(m)}) &\triangleq E_{\mathbf{o}(t)|\mathbf{y}(t),\hat{\phi}^{(m)}} \{ \log p(\mathbf{z}(t)|\phi) \} \\ &= -n \log |\sigma^2| - n \log |\tau(t)| - \log |\Sigma| \\ &\quad - \|\mathbf{y}(t) - \mathbf{a}(\boldsymbol{\theta})s(t) - \hat{\mathbf{o}}(t)\|_{\sigma^2}^2 - \frac{1}{\sigma^2} \text{Tr} \left\{ \widehat{\mathbf{o}(t)\mathbf{o}(t)^H} \right\} \\ &\quad + \frac{1}{\sigma^2} \hat{\mathbf{o}}(t)^H \hat{\mathbf{o}}(t) - \text{Tr} \left\{ \widehat{\mathbf{o}(t)\mathbf{o}(t)^H} (\tau(t)\Sigma)^{-1} \right\}, \end{aligned}$$

and due to the i.i.d. assumption stated above

$$\begin{aligned} Q(\phi|\hat{\phi}^{(m)}) &= \sum_{t=1}^T Q_t(\phi|\hat{\phi}^{(m)}) = -nT \log |\sigma^2| - T \log |\Sigma| \\ &\quad - \sum_{t=1}^T \left(\|\mathbf{y}(t) - \mathbf{a}(\boldsymbol{\theta})s(t) - \hat{\mathbf{o}}(t)\|_{\sigma^2}^2 + n \log |\tau(t)| \right) \\ &\quad + \frac{1}{\sigma^2} \text{Tr} \left\{ \widehat{\mathbf{o}(t)\mathbf{o}(t)^H} \right\} - \frac{1}{\sigma^2} \hat{\mathbf{o}}(t)^H \hat{\mathbf{o}}(t) \\ &\quad + \text{Tr} \left\{ \widehat{\mathbf{o}(t)\mathbf{o}(t)^H} (\tau(t)\Sigma)^{-1} \right\}. \end{aligned}$$

B. M-step

The M-step will be carried out by an alternating scheme. This leads to closed form expressions for all the unknown parameters except, obviously, regarding to $\boldsymbol{\theta}$ due to the non convexity of the cost function for fixed $\Sigma, \sigma^2, \{s(t)\}, \{\tau(t)\}$. First, let us consider the derivative of $Q(\phi|\hat{\phi}^{(m)})$ w.r.t. $\tau(t)$. After, equating the latter to zero, we obtain

$$\hat{\tau}(t)^{(m+1)} = \frac{1}{n} \text{Tr} \left(\widehat{\mathbf{o}(t)\mathbf{o}(t)^H} (\hat{\Sigma}^{(m)})^{-1} \right) \quad (6)$$

Regarding the update of $\hat{\Sigma}^{(m+1)}$, we obtain

$$\hat{\Sigma}^{(m+1)} = \frac{n}{T} \sum_{t=1}^T \frac{\widehat{\mathbf{o}(t)\mathbf{o}(t)^H}}{\text{Tr} \left(\widehat{\mathbf{o}(t)\mathbf{o}(t)^H} (\hat{\Sigma}^{(m)})^{-1} \right)} \quad (7)$$

which is a fixed point estimate, known to converge whatever the initial estimate $\hat{\Sigma}^{(0)}$ [20]. In order to remove the ambiguity between the speckle and texture parameter, we normalize the estimated $\hat{\Sigma}^{(m+1)}$ [21] as (with some abuse of notation)

$$\hat{\Sigma}^{(m+1)} = \frac{\hat{\Sigma}^{(m+1)}}{\text{Tr} \left\{ \hat{\Sigma}^{(m+1)} \right\}}.$$

Now, we focus on the estimation of $\boldsymbol{\theta}, \sigma^2, \{s(t)\}_{t=1,\dots,T}$, for which

$$Q(\phi|\hat{\phi}^{(m)}) \propto -nT \log |\sigma^2| - \sum_{t=1}^T \|\mathbf{y}(t) - \mathbf{a}(\boldsymbol{\theta})s(t) - \hat{\mathbf{o}}(t)\|_{\sigma^2}^2$$

Classical calculus (based on [22], [23]) lead to following concentrated procedure

$$\widehat{\sigma^2}^{(m+1)} = \frac{1}{n} \text{Tr} \left(\mathbf{\Pi}_{\hat{\boldsymbol{\theta}}^{(m)}}^{\perp} E_{\mathbf{o}(t)|\mathbf{y}(t),\hat{\phi}^{(m)}} \{ \mathbf{R}_{\mathbf{o}} \} \right) \quad (8)$$

in which

$$\mathbf{R}_{\mathbf{o}} = \frac{1}{T} \sum_{t=1}^T (\mathbf{y}(t) - \mathbf{o}(t)) (\mathbf{y}(t) - \mathbf{o}(t))^H$$

leading to

$$E_{\mathbf{o}(t)|\mathbf{y}(t),\hat{\phi}^{(m)}} \{ \mathbf{R}_{\mathbf{o}} \} = \hat{\mathbf{R}}_{\mathbf{o}} + \text{Tr} \{ \widehat{\mathbf{o}(t)\mathbf{o}(t)^H} - \hat{\mathbf{o}}(t)\hat{\mathbf{o}}(t)^H \}.$$

The adjusted sample covariance matrix is given by

$$\hat{\mathbf{R}}_{\mathbf{o}} = \frac{1}{T} \sum_{t=1}^T (\mathbf{y}(t) - \hat{\mathbf{o}}(t)) (\mathbf{y}(t) - \hat{\mathbf{o}}(t))^H$$

and the orthogonal projection reads

$$\mathbf{\Pi}_{\mathbf{a}(\boldsymbol{\theta})}^{\perp} = \mathbf{I} - \mathbf{\Pi}_{\mathbf{a}(\boldsymbol{\theta})}$$

with

$$\mathbf{\Pi}_{\mathbf{a}(\boldsymbol{\theta})} = \frac{\mathbf{a}(\boldsymbol{\theta})\mathbf{a}(\boldsymbol{\theta})^H}{\mathbf{a}(\boldsymbol{\theta})^H \mathbf{a}(\boldsymbol{\theta})},$$

$$\hat{s}(t)^{(m+1)} = \frac{\mathbf{a}(\hat{\boldsymbol{\theta}}^{(m)})^H}{\mathbf{a}(\hat{\boldsymbol{\theta}}^{(m)})^H \mathbf{a}(\hat{\boldsymbol{\theta}}^{(m)})} (\mathbf{y}(t) - \hat{\mathbf{o}}(t)) \quad (9)$$

and

$$\hat{\boldsymbol{\theta}}^{(m+1)} = \arg \min_{\boldsymbol{\theta}} \text{Tr} \left\{ \mathbf{\Pi}_{\mathbf{a}(\boldsymbol{\theta})}^{\perp} \hat{\mathbf{R}}_{\mathbf{o}} \right\}. \quad (10)$$

Remark: *Extension to the multi-sources scenario*

The above analysis is dedicated to the one source case. In the multi-sources scenario, a straightforward extension is to replace the update of $\{s(t)\}_{t=1,\dots,T}$, $\boldsymbol{\theta}$ and σ^2 , respectively by,

$$(\widehat{\sigma^2})^{(m+1)} = \frac{1}{n} \text{Tr} \left\{ \mathbf{\Pi}_{\mathbf{A}(\hat{\boldsymbol{\theta}}^{(m)})}^{\perp} \hat{\mathbf{R}}_{\mathbf{o}} \right\} \quad (11)$$

$$\hat{\mathbf{S}}^{(m+1)} = \mathbf{A}(\hat{\boldsymbol{\theta}}^{(m)})^{\sharp} (\mathbf{Y} - \hat{\mathbf{O}}) \quad (12)$$

$$\hat{\boldsymbol{\theta}}^{(m+1)} = \arg \min_{\boldsymbol{\theta}} \text{Tr} \left\{ \mathbf{\Pi}_{\mathbf{A}(\boldsymbol{\theta})}^{\perp} \hat{\mathbf{R}}_{\mathbf{o}} \right\} \quad (13)$$

in which $\mathbf{A}(\boldsymbol{\theta})^{\sharp}$ denotes the pseudo inverse of $\mathbf{A}(\boldsymbol{\theta})$, \mathbf{Y} , $\hat{\mathbf{O}}$ and \mathbf{S} represents the concatenation of the incomplete data, the guess and the emitted signal sources, respectively.

IV. NUMERICAL EVALUATIONS

The aim of this section is to evaluate the performance of the proposed EM based algorithm in the case of a mixture noise constituted of a Gaussian and spherically invariant random process. Specifically, we consider a non-uniform linear array of $n = 9$ sensors which observes a far field source emitting a signal from $\theta = 25$ deg during $T = 25$ snapshots. The interspacing elements of the array is less than half a wavelength. Regarding the noise mixture, the spherically invariant random component is Student't distributed. Namely, the texture parameter follows an inverse-Gamma distribution $\tau \sim \text{invGamma}(a, b)$, in which the scale and the shape parameters are given by $a = 1.1$ and $b = 2$, respectively. The signal-to-noise ratio is defined as $\text{SNR} = \frac{\|\mathbf{s}\|_2^2}{\sigma^2}$. The entries of the covariance matrix Σ are generated by $[\Sigma]_{m,n} = \sigma^2 \times 0.9^{|m-n|} \exp(j2\pi(m-n))$. Finally, we recall that the EM algorithm is an iterative algorithm that attempts to reach the maximum likelihood estimate which is given as the maximization of a non-convex and a highly non-linear

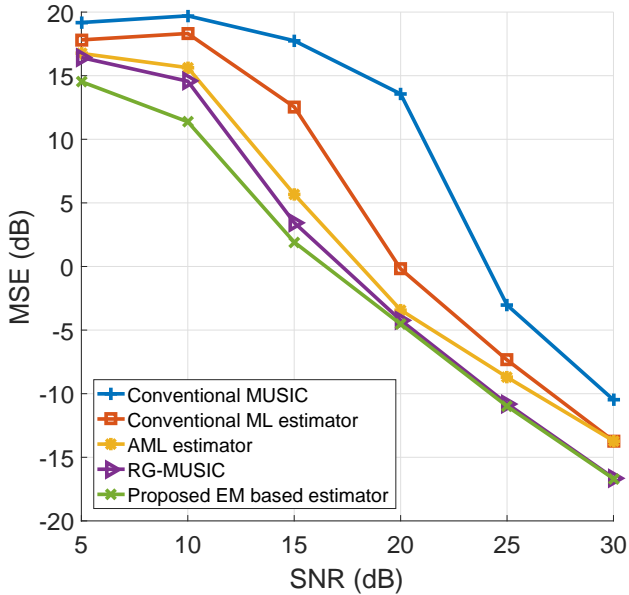


Fig. 1. Mean square error vs. signal to noise ratio w.r.t. θ .

cost function (1). Consequently, the initialization step is of importance to avoid the numerous local minimum. One way is to use some competitive algorithms, as the RG-MUSIC based estimator [24], in order to initialize the EM algorithm.

In the following simulation, we consider some classical algorithms as the MUSIC algorithm [25] and the conventional ML estimator (i.e., the Gaussian ML estimator), but also, some recent competitive algorithms as the RG-MUSIC algorithm [24] and the so-called robust AMLE (Approximated ML estimator) [26]. Fig. 1 represents the mean square error of the aforementioned estimators from which we can notice that the proposed estimator outperforms the MUSIC, conventional ML estimator, RG-MUSIC and the AMLE. This fact is not surprising since the proposed estimator relies on the assumption of mixture of Gaussian and spherically invariant random distributed noise.

Finally, we represent in Fig. 2 the convergence of the direction of arrival estimates. We can notice that few iterations are needed in order to obtain a final estimate (less than 10 in our case). This means that the complexity cost of the proposed algorithm remains acceptable. Specifically, the cost of each iteration is determined by the costly iteration, i.e., minimization of (10). The latter is comparable, in terms of complexity, to a minimization of a MUSIC-based scheme's objective function.

V. CONCLUSION

In this paper we presented a novel direction of arrival estimation in the context of a noise mixture constituted of a Gaussian noise, which represents the thermal noise, and a spherically invariant random process, which reflects in practice the possible presence of outliers. An EM-based algorithm is proposed in order to maintain a reasonable computational cost

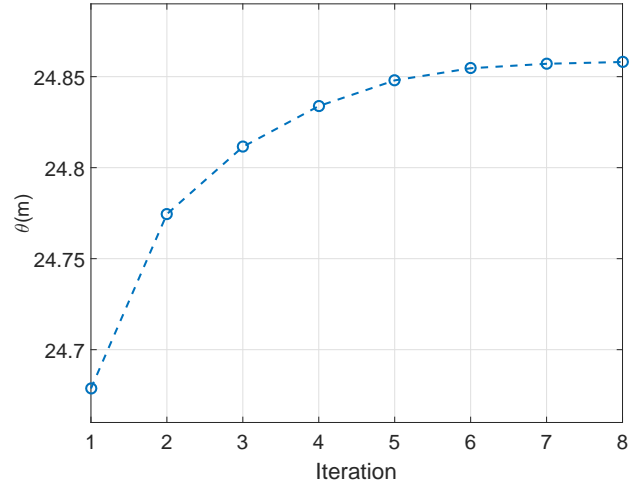


Fig. 2. Convergence rate of the parameter of interest estimates under SNR=10dB.

while solving the ML criteria and taking into account the two mixture components. Numerical simulations reveal that the proposed algorithm outperforms the state-of-the-art with only a few number of iterations.

REFERENCES

- [1] J. Shi, G. Hu, and T. Lei, "Doa estimation algorithms for low-angle targets with mimo radar," *Electronics Letters*, vol. 52, no. 8, pp. 652–654, 2016.
- [2] M. Haardt, M. Pesavento, F. Roemer, and M. N. El Korso, "Subspace methods and exploitation of special array structures," *Electronic Reference in Signal Processing: Array and Statistical Signal Processing* (M. Viberg, ed.), vol. 3, pp. 651–717, Academic Press Library in Signal Processing, Elsevier Ltd., Chapter 2.15, ISBN 978-0-12-411597-2, 2014.
- [3] J. Shi, Q. Zhang, and W. Shi, "Wideband doa estimation with deficient snapshots using low rank toeplitz structure," *Electronics Letters*, vol. 55, no. 17, pp. 961–963, 2019.
- [4] F. Yan, T. Jin, M. Jin, and Y. Shen, "Subspace-based direction-of-arrival estimation using centro-symmetrical arrays," *Electronics Letters*, vol. 52, no. 22, pp. 1895–1896, 2016.
- [5] R. Mallik, K. Rangarao, and U. Murthy, "Doa estimation by least squares approach," *Electronics Letters*, vol. 34, no. 12, pp. 1187–1189, 1998.
- [6] H. L. VanTrees, *Detection, Estimation and Modulation Theory: Radar-Sonar Signal Processing and Gaussian Signals in Noise*. New York: Wiley, 2001, vol. 3.
- [7] J. Li, Y. Li, and X. Zhang, "Direction of arrival estimation using combined coprime and nested array," *Electronics Letters*, vol. 55, no. 8, pp. 487–489, 2019.
- [8] A. Belouchrani and M. G. Amin, "Time-frequency MUSIC," *IEEE Signal Processing Letters*, vol. 5, pp. 109–110, Jun. 1999.
- [9] B. Rao and K. Hari, "Performance analysis of root-MUSIC," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 12, pp. 1939–1949, Dec. 1989.
- [10] E. Grosicki, K. Abed-Meraim, and Y. Hua, "A weighted linear prediction method for near-field source localization," *IEEE Trans. Signal Processing*, vol. 53, pp. 3651–3660, 2005.
- [11] F. Gini and M. V. Greco, "Covariance matrix estimation for CFAR detection in correlated heavy tailed clutter," *Signal Process.*, vol. 82, pp. 1847–1859, Dec. 2002.
- [12] O. Besson and Y. Abramovich, "Sensitivity Analysis of Likelihood Ratio Test in K Distributed and/or Gaussian Noise," *IEEE Signal Processing Letters*, vol. 22, pp. 2329–2333, Dec. 2015.
- [13] W. Zhang, S. Wu, and J. Wang, "Robust capon beamforming in presence of large doa mismatch," *Electronics Letters*, vol. 49, no. 1, pp. 75–76, 2013.

- [14] L. Zhang, J. Mei, A. Zielinski, and P. Cai, "Direction-of-arrival estimation for far-field acoustic signal in presence of near-field interferences," *Electronics Letters*, vol. 51, no. 1, pp. 101–103, 2015.
- [15] Z. Luo, R. Guo, X. Zhang, and Y. Nie, "Optimal and efficient designs of gaussian-tailed non-linearity in symmetric alpha-stable noise," *Electronics Letters*, vol. 55, no. 6, pp. 353–355, 2019.
- [16] E. Ollila, D. Tyler, V. Koivunen, and H. Poor, "Complex elliptically symmetric distributions: Survey, new results and applications," *IEEE Transactions on Signal Processing*, vol. 60, no. 11, pp. 5597–5625, 2012.
- [17] K. Yao, "A representation theorem and its applications to spherically invariant random processes," *IEEE Trans. Inf. Theory*, vol. 19, pp. 600–608, Sep. 1973.
- [18] A. Dempster, M. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 39, pp. 1–22, Jan. 1977.
- [19] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*. New York: Wiley, 1984.
- [20] F. Pascal, Y. Chitour, J.-P. Ovarlez, P. Forster, and P. Larzabal, "Covariance structure maximum likelihood estimates in compound gaussian noise: Existence and algorithm analysis," *IEEE Trans. Signal Process.*, vol. 56, pp. 34–48, Jan. 2008.
- [21] M. Mahot, F. Pascal, P. Forster, and J. P. Ovarlez, "Asymptotic properties of robust complex covariance matrix estimates," *IEEE Transactions on Signal Processing*, vol. 61, no. 13, pp. 3348–3356, July 2013.
- [22] 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.
- [23] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood and the Cramér Rao bound," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 720–741, May 1989.
- [24] R. Couillet, "Robust spiked random matrices and a robust G-MUSIC estimator," *Journal of Multivariate Analysis*, vol. 140, no. 9, pp. 139–161, 2015.
- [25] R. O. Schmidt, "A signal subspace approach to multiple emitter location and spectral estimation," Ph.D. dissertation, Stanford University, Stanford, CA, Nov. 1981.
- [26] O. Besson, Y. Abramovich, and B. Johnson, "Direction-of-arrival estimation in a mixture of K-distributed and Gaussian noise," *Signal Process.*, vol. 128, pp. 512–520, Nov. 2016.
- [27] R. Ben Abdallah, A. Breloy, M. N. El Korso and D. Lautru, "Bayesian Signal Subspace Estimation with Compound Gaussian Sources," *Signal Process.*, vol. 167, pp. 107310, 2020.