electronics

MDPI

Communication

Stochastic Maximum Likelihood Direction Finding in the Presence of Nonuniform Noise Fields

Ming-Yan Gong [1,*] and Bin Lyu [2]

[1] School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China
[2] Key Laboratory of Ministry of Education in Broadband Wireless Communication and Sensor Network Technology, Nanjing University of Posts and Telecommunications, Nanjing 210003, China; blyu@njupt.edu.cn
* Correspondence: jinyan0_o@outlook.com

Abstract: The maximum likelihood (ML) technique plays an important role in direction-of-arrival (DOA) estimation. In this paper, we employ and design the expectation–conditional maximization either (ECME) algorithm, a generalization of the expectation–maximization algorithm, for solving the ML direction finding problem of stochastic sources, which may be correlated, in unknown nonuniform noise. Unlike alternating maximization, the ECME algorithm updates both the source and noise covariance matrix estimates by explicit formulas, and can guarantee that both estimates are positive semi-definite and definite, respectively. Thus, the ECME algorithm is computationally efficient and operationally stable. Simulation results confirm that the ECME algorithm can efficiently obtain the ML based DOA estimate of each stochastic source.

Keywords: array signal processing; DOA estimation; EM algorithm; maximum likelihood estimation; nonuniform Gaussian noise; statistical signal processing; stochastic signal model

Citation: Gong, M.-Y.; Lyu, B. Stochastic Maximum Likelihood Direction Finding in the Presence of Nonuniform Noise Fields. *Electronics* **2023**, *12*, 2191. https://doi.org/10.3390/electronics12102191

Academic Editors: Manuel Arrebola, Fangqing Wen, Yangyang Dong and Hua Chen

Received: 19 March 2023
Revised: 2 May 2023
Accepted: 9 May 2023
Published: 11 May 2023

1. Introduction

Two source signal models are widely used in Cramer–Rao lower bound (CRLB) and maximum likelihood (ML) direction finding, i.e., the deterministic signal model where signals are deterministic and unknown, and the stochastic signal model where signals are Gaussian. For example, various CRLBs using both models have been derived [1–7]. However, ML direction finding generally involves high-dimensional search algorithms for both models, which causes a significant increase in computational complexity.

In order to reduce the computational complexity, two classic methods have been developed, i.e., alternating maximization (AM)-type [8] and expectation–maximization (EM)-type [9–12] algorithms. These two methods are first applied under uniform Gaussian noise, which decreases the number of parameters and simplifies the problem. However, the uniform noise model is unrealistic in many situations, and nonuniform noise has been considered in numerous papers [4,13–19]. In nonuniform noise, the covariance matrix still keeps a diagonal structure, but the diagonal elements are no longer identical, which hinders direction-of-arrival (DOA) estimation. To tackle the problem of direction finding in unknown nonuniform noise, diverse subspace separation approaches based on the subspace technique have been proposed in the literature [13–18].

For obtaining ML based solutions, AM- and EM-type algorithms have also been applied to this problem. However, the AM-type algorithms usually require high-dimensional numerical searches due to the noise nonuniformity at each iteration [4,19], which leads to heavy computational burdens. Moreover, when considering Gaussian source signals, the AM algorithm presented in [19] has one severe shortcoming: the source and noise covariance matrix estimates cannot be guaranteed to be positive semi-definite and definite, respectively. To this end, we have designed several computationally efficient EM-type algorithms in [20] that only need low-dimensional (one or two-dimensional) numerical searches at every iteration. In these EM-type algorithms using the stochastic signal model, however, the sources must be uncorrelated. This restricts the use of stochastic ML direction finding in some situations, e.g., multipath conditions. As a consequence, efficient algorithms are urgently needed to address this issue.

In this paper, we employ and design the expectation–conditional maximization either (ECME) algorithm [21], a generalization of the EM algorithm, for solving the ML direction finding problem of stochastic sources, which may be correlated, in unknown nonuniform noise. Unlike the AM algorithm in [19], the ECME algorithm updates both the source and noise covariance matrix estimates by explicit formulas and can guarantee that both estimates are positive semi-definite and definite, respectively. Thus, the ECME algorithm is computationally efficient and operationally stable. Simulation results confirm the effectiveness of the algorithm.

The rest of this paper is outlined as follows: In Section 2, we formulate the stochastic ML direction finding problem in unknown nonuniform noise. In Sections 3 and 4, we design the ECME algorithm and provide simulation results to show its effectiveness, respectively. Lastly, we conclude this paper in Section 5.

2. Problem Statement

For simplicity, let a uniformly spaced linear array of W sensors receive the plane waves impinging from V ($V < W$) narrow-band sources of wavelength ι. The distance between any adjacent sensors is $\iota/2$. We denote the direction associated with the vth source by $\beta_v \in (0, \pi)$ (radian), and write the received signal as

$$\mathbf{r}(t) = \sum_{v=1}^{V} \mathbf{a}(\beta_v) k_v(t) + \mathbf{j}(t) = \mathbf{A}(\boldsymbol{\beta})\mathbf{k}(t) + \mathbf{j}(t), \tag{1}$$

where $\mathbf{a}(\beta_v) = [1 \; a_v \; \cdots \; a_v^{W-1}]^T$, $a_v = \exp\big(-\jmath\pi\cos(\beta_v)\big)$, $[\cdot]^T$ denotes transposition, $\jmath = \sqrt{-1}$, $k_v(t)$ is the signal with respect to the vth source, and $\mathbf{j}(t)$ means nonuniform complex Gaussian noise of zero mean and covariance $\mathbf{Q}$, i.e., $\mathbf{j}(t) \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q})$. Here, $\mathbf{Q}$ is diagonal and expressed as $\mathbf{Q} = \mathrm{diag}\{\boldsymbol{\delta}\}$, where $\boldsymbol{\delta} = [\delta_1 \; \cdots \; \delta_W]^T > \mathbf{0}$ and $\mathbf{Q}$ is positive definite, i.e., $\mathbf{Q} \succ \mathbf{0}_W$ ($\mathbf{0}_W$ is the $W \times W$ zero matrix). Furthermore, if $\delta_1 = \cdots = \delta_W = \delta > 0$, $\mathbf{Q} = \delta\mathbf{I}_W$ ($\mathbf{I}_W$ is the $W \times W$ identity matrix), which makes the noise uniform. In (1), $\mathbf{A}(\boldsymbol{\beta}) = [\mathbf{a}(\beta_1) \; \cdots \; \mathbf{a}(\beta_V)]$ is the array manifold matrix, $\boldsymbol{\beta} = [\beta_1 \; \cdots \; \beta_V]^T \in \boldsymbol{\Gamma}$ with $\boldsymbol{\Gamma} = (0, \pi)^V$, and $\mathbf{k}(t) = [k_1(t) \; \cdots \; k_V(t)]^T$. For notational convenience, we use $\mathbf{A}$ instead of $\mathbf{A}(\boldsymbol{\beta})$ hereafter.

We consider Gaussian source signals, which may be correlated, and have $\mathbf{k}(t) \sim \mathcal{CN}(\mathbf{0}, \mathbf{O})$, where $\mathbf{O}$ is the source covariance matrix and positive semi-definite, i.e., $\mathbf{O} \succeq \mathbf{0}_V$. Let the sources be uncorrelated with the noise, such that

$$\mathbf{r}(t) \sim \mathcal{CN}(\mathbf{0}, \mathbf{G}), \mathbf{G} = \mathbf{A}\mathbf{O}\mathbf{A}^H + \mathbf{Q} \succ \mathbf{0}_W,$$

where $[\cdot]^H$ is conjugate transposition. On this foundation, the log-likelihood function (LLF) of L statistically independent snapshots can be formulated as

$$\mathcal{J}(\boldsymbol{\beta}, \boldsymbol{\rho}, \boldsymbol{\delta}) = \sum_{t=1}^{L} \log p\big(\mathbf{r}(t); \boldsymbol{\beta}, \boldsymbol{\rho}, \boldsymbol{\delta}\big) = f - L\big(\log|\mathbf{G}| + \mathrm{trace}\big[\mathbf{G}^{-1}\hat{\mathbf{R}}\big]\big), \tag{2}$$

where $|\cdot|$, $\mathrm{trace}[\cdot]$, and $(\cdot)^{-1}$ denote determinant, trace, and inversion, respectively. In (2), f is a constant, $\hat{\mathbf{R}} = (1/L)\sum_{t=1}^{L} \mathbf{r}(t)\mathbf{r}^H(t)$ means the covariance matrix of snapshots. Moreover,

$$\boldsymbol{\rho} = \big([\mathbf{O}]_{1,1}, \ldots, [\mathbf{O}]_{V,V}, \mathrm{Re}\big\{[\mathbf{O}]_{1,2}\big\}, \mathrm{Im}\big\{[\mathbf{O}]_{1,2}\big\}, \ldots, \mathrm{Re}\big\{[\mathbf{O}]_{V-1,V}\big\}, \mathrm{Im}\big\{[\mathbf{O}]_{V-1,V}\big\}\big),$$

where $[\mathbf{O}]_{p,q}$ is the (p,q)th element of $\mathbf{O}$, $\mathrm{Re}\{a\}$ and $\mathrm{Im}\{a\}$ represent the real part and imaginary part of a, respectively. Consequently, the ML based DOA estimation problem is

$$\max_{\boldsymbol{\beta}\in\boldsymbol{\Gamma}, \mathbf{O}\succeq\mathbf{0}_V, \boldsymbol{\delta}>\mathbf{0}} \mathcal{J}(\boldsymbol{\beta}, \boldsymbol{\rho}, \boldsymbol{\delta}). \tag{3}$$

We assume $\mathcal{R}[\mathbf{A}] = V$, where $\mathcal{R}[\mathbf{A}]$ is the rank of $\mathbf{A}$, and can thus eliminate $\mathbf{O}$ in (3) by [22]

$$\begin{aligned}\hat{\mathbf{O}}(\boldsymbol{\beta},\boldsymbol{\delta}) &= (\tilde{\mathbf{A}}^H\tilde{\mathbf{A}})^{-1}\tilde{\mathbf{A}}^H(\widetilde{\mathbf{R}}-\mathbf{I}_W)\tilde{\mathbf{A}}(\tilde{\mathbf{A}}^H\tilde{\mathbf{A}})^{-1} \\ &= (\tilde{\mathbf{A}}^H\tilde{\mathbf{A}})^{-1}\tilde{\mathbf{A}}^H\widetilde{\mathbf{R}}\tilde{\mathbf{A}}(\tilde{\mathbf{A}}^H\tilde{\mathbf{A}})^{-1}-(\tilde{\mathbf{A}}^H\tilde{\mathbf{A}})^{-1},\end{aligned} \tag{4}$$

where $\mathbf{Q}^{-1/2} = \mathrm{diag}\{1/\sqrt{\delta_1},\ldots,1/\sqrt{\delta_W}\}$, $\tilde{\mathbf{A}} = \mathbf{Q}^{-1/2}\mathbf{A}$, and $\widetilde{\mathbf{R}} = \mathbf{Q}^{-1/2}\hat{\mathbf{R}}\mathbf{Q}^{-1/2}$. In other words, $\boldsymbol{\rho}$ can be estimated using the estimates of $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$. On the basis of (4), $\mathbf{G}$ is rewritten as

$$\mathbf{G} = \mathbf{A}\hat{\mathbf{O}}(\boldsymbol{\beta},\boldsymbol{\delta})\mathbf{A}^H + \mathbf{Q} = \mathbf{Q}^{1/2}(\boldsymbol{\Pi}_{\tilde{\mathbf{A}}}\widetilde{\mathbf{R}}\boldsymbol{\Pi}_{\tilde{\mathbf{A}}} + \boldsymbol{\Pi}^{\perp}_{\tilde{\mathbf{A}}})\mathbf{Q}^{1/2},$$

where $\boldsymbol{\Pi}_{\tilde{\mathbf{A}}} = \tilde{\mathbf{A}}(\tilde{\mathbf{A}}^H\tilde{\mathbf{A}})^{-1}\tilde{\mathbf{A}}^H$ and $\boldsymbol{\Pi}^{\perp}_{\tilde{\mathbf{A}}} = \mathbf{I}_W - \boldsymbol{\Pi}_{\tilde{\mathbf{A}}}$. Then, Problem (3) is reduced to [22]

$$\min_{\boldsymbol{\beta}\in\boldsymbol{\Gamma},\boldsymbol{\delta}>\mathbf{0}} \mathcal{H}(\boldsymbol{\beta},\boldsymbol{\delta}) = \log\left|\mathbf{Q}^{1/2}(\boldsymbol{\Pi}_{\tilde{\mathbf{A}}}\widetilde{\mathbf{R}}\boldsymbol{\Pi}_{\tilde{\mathbf{A}}} + \boldsymbol{\Pi}^{\perp}_{\tilde{\mathbf{A}}})\mathbf{Q}^{1/2}\right| + \mathrm{trace}\left[(\boldsymbol{\Pi}_{\tilde{\mathbf{A}}}\widetilde{\mathbf{R}}\boldsymbol{\Pi}_{\tilde{\mathbf{A}}} + \boldsymbol{\Pi}^{\perp}_{\tilde{\mathbf{A}}})^{-1}\widetilde{\mathbf{R}}\right]. \tag{5}$$

In particular, if the noise is uniform Gaussian noise, Problem (5) can be further reduced to [23]

$$\min_{\boldsymbol{\beta}\in\boldsymbol{\Gamma}} \mathcal{G}(\boldsymbol{\beta}) = \left|\mathbf{A}\hat{\mathbf{O}}(\boldsymbol{\beta})\mathbf{A}^H + \hat{\delta}(\boldsymbol{\beta})\mathbf{I}_W\right|, \tag{6}$$

where

$$\begin{aligned}\hat{\delta}(\boldsymbol{\beta}) &= \mathrm{trace}\left[(\mathbf{I}_W - \mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H)\hat{\mathbf{R}}\right]/(W-V), \\ \hat{\mathbf{O}}(\boldsymbol{\beta}) &= (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H(\hat{\mathbf{R}} - \hat{\delta}(\boldsymbol{\beta})\mathbf{I}_W)\mathbf{A}(\mathbf{A}^H\mathbf{A})^{-1}.\end{aligned}$$

Unfortunately, it is very difficult to reduce Problem (5) to some problems with fewer parameters under nonuniform Gaussian noise. Of course, applying gradient-type algorithms to search the solution of Problem (5) is computationally intensive due to the search space of dimension $W+V$ and the complexity of $\mathcal{H}(\boldsymbol{\beta},\boldsymbol{\delta})$.

In fact, when direct maximization over all parameters is intractable, AM can always be utilized. As stated before, the authors in [19] have presented an AM algorithm consisting of two steps at every iteration for Problem (3). Specifically, the first step obtains $\boldsymbol{\delta}^{(d)}$, the estimate of $\boldsymbol{\delta}$ at the dth iteration, by a gradient based algorithm, which is called the "modified inverse iteration algorithm" and satisfies

$$\mathcal{J}(\boldsymbol{\beta}^{(d-1)},\boldsymbol{\rho}^{(d-1)},\boldsymbol{\delta}^{(d)}) \geq \mathcal{J}(\boldsymbol{\beta}^{(d-1)},\boldsymbol{\rho}^{(d-1)},\boldsymbol{\delta}^{(d-1)}), \tag{7}$$

where $[\cdot]^{(0)}$ means an initial estimate. Then, the second step simultaneously obtains $\boldsymbol{\beta}^{(d)}$ and $\boldsymbol{\rho}^{(d)}$ by

$$(\boldsymbol{\beta}^{(d)},\boldsymbol{\rho}^{(d)}) = \arg\max_{\boldsymbol{\beta}\in\boldsymbol{\Gamma},\mathbf{O}\succeq\mathbf{0}_V} \mathcal{J}(\boldsymbol{\beta},\boldsymbol{\rho},\boldsymbol{\delta}^{(d)}), \tag{8}$$

which is solved in a separable manner, i.e.,

$$\boldsymbol{\beta}^{(d)} = \arg\min_{\boldsymbol{\beta}\in\boldsymbol{\Gamma}} \mathcal{H}(\boldsymbol{\beta},\boldsymbol{\delta}^{(d)}), \tag{9}$$

$$\mathbf{O}^{(d)} = \hat{\mathbf{O}}(\boldsymbol{\beta}^{(d)},\boldsymbol{\delta}^{(d)}). \tag{10}$$

However, the AM algorithm has two drawbacks: (1) obtaining $\boldsymbol{\delta}^{(d)}$ and $\boldsymbol{\beta}^{(d)}$ is computationally expensive; (2) $\mathbf{Q}^{(d)} \succ \mathbf{0}_W$ (or $\boldsymbol{\delta}^{(d)} > \mathbf{0}$) and $\mathbf{O}^{(d)} \succeq \mathbf{0}_V$ cannot be guaranteed [24,25]. To efficiently obtain the ML estimate of $\boldsymbol{\beta}$ in (3), we employ and design the ECME algorithm in the next section.

3. ECME Algorithm

Existing EM-type algorithms for stochastic ML direction finding are only applicable to uncorrelated sources [9,12,20], i.e., $\mathbf{O}$ is diagonal. In this section, we employ and design the ECME algorithm [21], a generalization of the EM algorithm, to solve Problem (3) associated with correlated sources.

3.1. Procedure

The sources in (1) may be correlated, so we choose $\mathbf{K} = [\mathbf{k}(1) \ \cdots \ \mathbf{k}(L)]$ and $\mathbf{J} = [\mathbf{j}(1) \ \cdots \ \mathbf{j}(L)]$ as augmented data. We express the augmented-data LLF as

$$\begin{aligned}\mathcal{M}(\mathbf{K},\mathbf{J};\boldsymbol{\rho},\delta) = &\ \textstyle\sum_{t=1}^{L}\left[\log p(\mathbf{k}(t);\boldsymbol{\rho}) + \log p(\mathbf{j}(t);\delta)\right]\\ = &\ h - L\big(\log|\mathbf{O}| + \operatorname{trace}[\mathbf{O}^{-1}\hat{\mathbf{N}}_k]\big) +\\ &\ f - L\big(\log|\mathbf{Q}| + \operatorname{trace}[\mathbf{Q}^{-1}\hat{\mathbf{N}}_j]\big),\end{aligned} \tag{11}$$

where h is a constant, $\hat{\mathbf{N}}_k = (1/L)\sum_{t=1}^{L}\mathbf{k}(t)\mathbf{k}^H(t)$, and $\hat{\mathbf{N}}_j = (1/L)\sum_{t=1}^{L}\mathbf{j}(t)\mathbf{j}^H(t)$. With (11), we first construct the EM algorithm [26], whose expectation and maximization steps at the dth iteration are derived below. Let $\mathcal{E}\{\cdot\}$ and $\mathcal{D}\{\cdot\}$ represent expectation and covariance, respectively.

3.1.1. Expectation Step

Compute the conditional expectation of the augmented-data LLF, i.e.,

$$\begin{aligned}\mathcal{M}\big(\boldsymbol{\rho},\delta;\boldsymbol{\Omega}^{(d-1)}\big) = &\ \mathcal{E}\{\mathcal{M}(\mathbf{K},\mathbf{J};\boldsymbol{\rho},\delta) \mid \mathbf{F};\boldsymbol{\Omega}^{(d-1)}\}\\ = &\ h - L\big(\log|\mathbf{O}| + \operatorname{trace}[\mathbf{O}^{-1}\hat{\mathbf{N}}_k^{(d)}]\big) +\\ &\ f - L\big(\log|\mathbf{Q}| + \operatorname{trace}[\mathbf{Q}^{-1}\hat{\mathbf{N}}_j^{(d)}]\big)\end{aligned} \tag{12}$$

with $\boldsymbol{\Omega}^{(d-1)} = (\boldsymbol{\beta}^{(d-1)},\boldsymbol{\rho}^{(d-1)},\delta^{(d-1)})$, $\boldsymbol{\beta}^{(d-1)} \in \boldsymbol{\Gamma}$, $\mathbf{O}^{(d-1)} \succeq \mathbf{0}_V$, $\delta^{(d-1)} > \mathbf{0}$, and $\mathbf{F} = [\mathbf{r}(1) \ \cdots \ \mathbf{r}(L)]$. Moreover,

$$\begin{aligned}\hat{\mathbf{N}}_k^{(d)} = &\ \mathcal{E}\{\hat{\mathbf{N}}_k \mid \mathbf{F};\boldsymbol{\Omega}^{(d-1)}\}\\ = &\ [\mathbf{H}^{(d-1)}]^H\hat{\mathbf{R}}\mathbf{H}^{(d-1)} + \mathbf{O}^{(d-1)} - [\mathbf{H}^{(d-1)}]^H\mathbf{G}^{(d-1)}\mathbf{H}^{(d-1)} \succeq \mathbf{0}_V,\end{aligned} \tag{13}$$

$$\begin{aligned}\hat{\mathbf{N}}_j^{(d)} = &\ \mathcal{E}\{\hat{\mathbf{N}}_j \mid \mathbf{F};\boldsymbol{\Omega}^{(d-1)}\}\\ = &\ \mathbf{Q}^{(d-1)}[\mathbf{G}^{(d-1)}]^{-1}\hat{\mathbf{R}}[\mathbf{G}^{(d-1)}]^{-1}\mathbf{Q}^{(d-1)} +\\ &\ \mathbf{Q}^{(d-1)} - \mathbf{Q}^{(d-1)}[\mathbf{G}^{(d-1)}]^{-1}\mathbf{Q}^{(d-1)} \succeq \mathbf{0}_W,\end{aligned} \tag{14}$$

where $\mathbf{H}^{(d-1)} = [\mathbf{G}^{(d-1)}]^{-1}\mathbf{A}^{(d-1)}\mathbf{O}^{(d-1)}$, the conditional distributions of $\mathbf{k}(t)$ and $\mathbf{j}(t)$ can be obtained in [27], and

$$\begin{aligned}\mathcal{E}\{\mathbf{k}(t) \mid \mathbf{F};\boldsymbol{\Omega}^{(d-1)}\} &= [\mathbf{H}^{(d-1)}]^H\mathbf{r}(t),\\ \mathcal{D}\{\mathbf{k}(t) \mid \mathbf{F};\boldsymbol{\Omega}^{(d-1)}\} &= \mathbf{O}^{(d-1)} - [\mathbf{H}^{(d-1)}]^H\mathbf{G}^{(d-1)}\mathbf{H}^{(d-1)} \succeq \mathbf{0}_V,\\ \mathcal{E}\{\mathbf{j}(t) \mid \mathbf{F};\boldsymbol{\Omega}^{(d-1)}\} &= \mathbf{Q}^{(d-1)}[\mathbf{G}^{(d-1)}]^{-1}\mathbf{r}(t),\\ \mathcal{D}\{\mathbf{j}(t) \mid \mathbf{F};\boldsymbol{\Omega}^{(d-1)}\} &= \mathbf{Q}^{(d-1)} - \mathbf{Q}^{(d-1)}[\mathbf{G}^{(d-1)}]^{-1}\mathbf{Q}^{(d-1)} \succeq \mathbf{0}_W.\end{aligned}$$

3.1.2. Maximization Step

Obtain $\boldsymbol{\rho}^{(d)}$ and $\boldsymbol{\delta}^{(d)}$ by maximizing $\mathcal{M}(\boldsymbol{\rho}, \boldsymbol{\delta}; \boldsymbol{\Omega}^{(d-1)})$ with respect to $\boldsymbol{\rho}$ and $\boldsymbol{\delta}$, which leads to the two parallel subproblems

$$\min_{\mathbf{O}\succeq\mathbf{0}_V} \log|\mathbf{O}| + \operatorname{trace}\left[\mathbf{O}^{-1}\hat{\mathbf{N}}_k^{(d)}\right], \tag{15}$$

$$\min_{\mathbf{Q}\succ\mathbf{0}_W} \log|\mathbf{Q}| + \operatorname{trace}\left[\mathbf{Q}^{-1}\hat{\mathbf{N}}_j^{(d)}\right]. \tag{16}$$

$\boldsymbol{\rho}^{(d)}$ and $\boldsymbol{\delta}^{(d)}$ are simultaneously obtained by

$$\mathbf{O}^{(d)} = \hat{\mathbf{N}}_k^{(d)} \succeq \mathbf{0}_V, \tag{17}$$

$$\delta_w^{(d)} = \begin{cases} [\hat{\mathbf{N}}_j^{(d)}]_{w,w}, & [\hat{\mathbf{N}}_j^{(d)}]_{w,w} > 0, \\ \delta_w^{(d-1)}/2, & [\hat{\mathbf{N}}_j^{(d)}]_{w,w} = 0, \end{cases} \quad \forall w. \tag{18}$$

From (17) and (18), we have the monotonicity of generalized EM algorithms [26], i.e.,

$$\mathcal{J}\left(\boldsymbol{\beta}^{(d-1)}, \boldsymbol{\rho}^{(d)}, \boldsymbol{\delta}^{(d)}\right) \geq \mathcal{J}\left(\boldsymbol{\beta}^{(d-1)}, \boldsymbol{\rho}^{(d-1)}, \boldsymbol{\delta}^{(d-1)}\right). \tag{19}$$

Obviously, $\boldsymbol{\beta}^{(d)}$ is not obtained at the dth iteration of the EM algorithm.

3.1.3. Conditional Maximization Step

In order to obtain $\boldsymbol{\beta}^{(d)}$, we now add a conditional maximization step at this iteration. Considering the following monotonicity:

$$\mathcal{J}\left(\boldsymbol{\beta}^{(d)}, \boldsymbol{\rho}^{(d)}, \boldsymbol{\delta}^{(d)}\right) \geq \mathcal{J}\left(\boldsymbol{\beta}^{(d-1)}, \boldsymbol{\rho}^{(d)}, \boldsymbol{\delta}^{(d)}\right), \boldsymbol{\beta}^{(d)} \in \boldsymbol{\Gamma}, \tag{20}$$

We can design this step as

$$\boldsymbol{\beta}^{(d)} = \arg\max_{\boldsymbol{\beta}\in\boldsymbol{\Gamma}} \mathcal{J}\left(\boldsymbol{\beta}, \boldsymbol{\rho}^{(d)}, \boldsymbol{\delta}^{(d)}\right), \tag{21}$$

or use a gradient-type algorithm to obtain $\boldsymbol{\beta}^{(d)}$ based on (20), e.g., Algorithm 1 in the next section. Due to the additional step unrelated to augmented data, the above EM algorithm becomes the ECME algorithm [21].

Algorithm 1 Steepest Descent-Based DOA Estimation

$f(\boldsymbol{\beta}) = -\mathcal{J}\left(\boldsymbol{\beta}, \boldsymbol{\rho}^{(d)}, \boldsymbol{\delta}^{(d)}\right)/L$, initialize $\boldsymbol{\beta} = \boldsymbol{\beta}^{(d-1)} \in \boldsymbol{\Gamma}$.
while $\|\nabla f(\boldsymbol{\beta})\|_2 > 0.001$ **do**
$$t_v = \begin{cases} -(\pi - \beta_v)/f_v'(\boldsymbol{\beta}), & f_v'(\boldsymbol{\beta}) < 0, \\ \beta_v/f_v'(\boldsymbol{\beta}), & f_v'(\boldsymbol{\beta}) > 0, \quad \forall v. \\ \infty, & f_v'(\boldsymbol{\beta}) = 0, \end{cases}$$
  $t = 0.1 \times \min\{t_1, \ldots, t_V\}$.
  while $f\left(\boldsymbol{\beta} - t\nabla f(\boldsymbol{\beta})\right) > f(\boldsymbol{\beta}) - 0.3t\|\nabla f(\boldsymbol{\beta})\|_2^2$ **do**
    $t = 0.5t$.
  end while
  $\boldsymbol{\beta} = \boldsymbol{\beta} - t\nabla f(\boldsymbol{\beta}) \in \boldsymbol{\Gamma}$.
end while
$\boldsymbol{\beta}^{(d)} = \boldsymbol{\beta}$.

3.2. Stability and Complexity

The stable operation of the ECME algorithm requires $\mathbf{Q}^{(d)} \succ \mathbf{0}_W$ (or $\boldsymbol{\delta}^{(d)} > \mathbf{0}$) and $\mathbf{O}^{(d)} \succeq \mathbf{0}_V$ for $d \geq 0$, so we give the following proposition.

Proposition 1. *In the ECME algorithm, $\mathbf{Q}^{(d)} \succ \mathbf{0}_W$ (or $\delta^{(d)} > \mathbf{0}$) and $\mathbf{O}^{(d)} \succeq \mathbf{0}_V$ for $d \geq 1$ if $\mathbf{Q}^{(0)} \succ \mathbf{0}_W$ (or $\delta^{(0)} > \mathbf{0}$) and $\mathbf{O}^{(0)} \succeq \mathbf{0}_V$.*

Proof. We utilize the mathematical induction method. If $\mathbf{Q}^{(u)} \succ \mathbf{0}_W$ (or $\delta^{(u)} > \mathbf{0}$) and $\mathbf{O}^{(u)} \succeq \mathbf{0}_V$, we have $\mathbf{G}^{(u)} = \mathbf{A}^{(u)}\mathbf{O}^{(u)}[\mathbf{A}^{(u)}]^H + \mathbf{Q}^{(u)} \succ \mathbf{0}_W$, which leads to $\hat{\mathbf{N}}_j^{(u+1)} \succeq \mathbf{0}_W$ in (14) and then in (18) $\delta_w^{(u+1)} > 0, \forall w$, i.e., $\mathbf{Q}^{(u+1)} \succ \mathbf{0}_W$ (or $\delta^{(u+1)} > \mathbf{0}$). Furthermore, $\mathbf{O}^{(u+1)} = \hat{\mathbf{N}}_k^{(u+1)} \succeq \mathbf{0}_V$ is straightforward in (13). The proof is completed. □

Proposition 1 indicates that, when $\mathbf{Q}^{(0)} \succ \mathbf{0}_W$ (or $\delta^{(0)} > \mathbf{0}$) and $\mathbf{O}^{(0)} \succeq \mathbf{0}_V$ in the ECME algorithm, $\boldsymbol{\rho}^{(d)}$ and $\delta^{(d)}$ obtained at the dth iteration are in the parameter spaces, respectively. Hence, the ECME algorithm is operationally stable.

Since $\boldsymbol{\rho}^{(d)}$ and $\delta^{(d)}$ are obtained via the explicit formulas in (17) and (18), the computational complexity of the ECME algorithm is dominated by obtaining $\boldsymbol{\beta}^{(d)}$ in (20). Compared with the AM algorithm in [19], the ECME algorithm is, thus, computationally efficient.

3.3. Limit Point

According to [21,28], we know that the ECME algorithm satisfies certain regularity conditions and always converges to a stationary point of $\mathcal{J}(\boldsymbol{\beta}, \boldsymbol{\rho}, \delta)$. Unfortunately, $\mathcal{J}(\boldsymbol{\beta}, \boldsymbol{\rho}, \delta)$ tends to have multiple stationary points, and the limit point of the ECME algorithm may be an undesirable stationary point. To deal with this issue, we need to provide an accurate initial point. Following the method in [19], we can assume that the noise is uniform and then evaluate $\mathcal{G}(\boldsymbol{\beta})$ in (6) on a coarse V-dimensional grid to find a grid point, close to the global minimum of $\mathcal{G}(\boldsymbol{\beta})$, as $\boldsymbol{\beta}^{(0)}$ of the ECME algorithm. We can also use the estimate of $\boldsymbol{\beta}$, obtained by a subspace [29] or a sparse representation based [30] algorithm, as $\boldsymbol{\beta}^{(0)}$ due to the higher accuracy of the stochastic ML estimate of $\boldsymbol{\beta}$ [2].

On the boundary of the positive semi-definite region of $\boldsymbol{\rho}$, i.e., the set $\eth = \{\boldsymbol{\rho} \mid \mathbf{O} \succeq \mathbf{0}_V$ and $\mathcal{R}[\mathbf{O}] < V\}$, we give the following proposition. Let $\mathcal{N}[\mathbf{O}]$ denote the null space of $\mathbf{O}$.

Proposition 2. *In the ECME algorithm, $\mathcal{N}[\mathbf{O}^{(d)}] = \mathcal{N}[\mathbf{O}^{(0)}]$ for $d \geq 1$ if $\mathbf{Q}^{(0)} \succ \mathbf{0}_W$ (or $\delta^{(0)} > \mathbf{0}$) and $\mathbf{O}^{(0)} \succeq \mathbf{0}_V$.*

Proof. From Proposition 1, we first know that $\mathbf{Q}^{(d)} \succ \mathbf{0}_W$ (or $\delta^{(d)} > \mathbf{0}$), $\mathbf{O}^{(d)} \succeq \mathbf{0}_V$, and $\mathbf{G}^{(d)} \succ \mathbf{0}_W$ for $d \geq 0$ due to $\mathbf{Q}^{(0)} \succ \mathbf{0}_W$ (or $\delta^{(0)} > \mathbf{0}$) and $\mathbf{O}^{(0)} \succeq \mathbf{0}_V$. Then, a proof by the mathematical induction method is given.

If $\mathbf{O}^{(u)}\mathbf{v} = \mathbf{0}$, we have $\mathbf{O}^{(u+1)}\mathbf{v} = \hat{\mathbf{N}}_k^{(u+1)}\mathbf{v} = \mathbf{0}$ in (13) and thus $\mathcal{N}[\mathbf{O}^{(u)}] \subseteq \mathcal{N}[\mathbf{O}^{(u+1)}]$. Furthermore, if $\mathbf{O}^{(u+1)}\mathbf{v} = \hat{\mathbf{N}}_k^{(u+1)}\mathbf{v} = \mathbf{0}$, we have $\mathbf{v}^H\hat{\mathbf{N}}_k^{(u+1)}\mathbf{v} = 0$ and in (13)

$$\mathbf{v}^H[\mathbf{H}^{(u)}]^H\hat{\mathbf{R}}\mathbf{H}^{(u)}\mathbf{v} = 0 \quad \Rightarrow \quad [\mathbf{H}^{(u)}]^H\hat{\mathbf{R}}\mathbf{H}^{(u)}\mathbf{v} = \mathbf{0}, \tag{22}$$

$$\mathbf{v}^H(\mathbf{O}^{(u)} - [\mathbf{H}^{(u)}]^H\mathbf{G}^{(u)}\mathbf{H}^{(u)})\mathbf{v} = 0 \quad \Rightarrow \quad (\mathbf{O}^{(u)} - [\mathbf{H}^{(u)}]^H\mathbf{G}^{(u)}\mathbf{H}^{(u)})\mathbf{v} = \mathbf{0}. \tag{23}$$

In order to proceed, we use the matrix inversion formula [22]

$$\mathbf{G}^{-1} = \mathbf{Q}^{-1/2}\left[\mathbf{I}_W - \tilde{\mathbf{A}}(\mathbf{O}\tilde{\mathbf{A}}^H\tilde{\mathbf{A}} + \mathbf{I}_V)^{-1}\mathbf{O}\tilde{\mathbf{A}}^H\right]\mathbf{Q}^{-1/2} \tag{24}$$

and obtain

$$\mathbf{O} - \mathbf{H}^H\mathbf{G}\mathbf{H} = (\mathbf{O}\tilde{\mathbf{A}}^H\tilde{\mathbf{A}} + \mathbf{I}_V)^{-1}\mathbf{O}, \tag{25}$$

which suggests $\mathcal{N}[\mathbf{O}] = \mathcal{N}[\mathbf{O} - \mathbf{H}^H\mathbf{G}\mathbf{H}]$ and $\mathcal{N}[\mathbf{O}^{(u)}] = \mathcal{N}[\mathbf{O}^{(u)} - [\mathbf{H}^{(u)}]^H\mathbf{G}^{(u)}\mathbf{H}^{(u)}]$. Accordingly, $\mathbf{O}^{(u)}\mathbf{v} = \mathbf{0}$ in (22) and (23), leading to $\mathcal{N}[\mathbf{O}^{(u)}] \supseteq \mathcal{N}[\mathbf{O}^{(u+1)}]$. Finally, by combining $\mathcal{N}[\mathbf{O}^{(u)}] \subseteq \mathcal{N}[\mathbf{O}^{(u+1)}]$ and $\mathcal{N}[\mathbf{O}^{(u)}] \supseteq \mathcal{N}[\mathbf{O}^{(u+1)}]$, we obtain $\mathcal{N}[\mathbf{O}^{(u)}] = \mathcal{N}[\mathbf{O}^{(u+1)}]$. The proof is completed. □

Proposition 2 indicates that if $\boldsymbol{\rho}^{(0)}$ in the ECME algorithm is on the boundary, i.e., $\boldsymbol{\rho}^{(0)} \in \eth$ and $\mathcal{N}[\mathbf{O}^{(0)}]$ is nonempty, the limit point of $\boldsymbol{\rho}$ is also on the boundary. Hence, let $(\boldsymbol{\beta}^*, \boldsymbol{\rho}^*, \delta^*)$ denote the solution of Problem (3) and if $\boldsymbol{\rho}^* \in \eth$, we may need to estimate $\mathcal{N}[\mathbf{O}^*]$ before implementing the ECME algorithm. Fortunately, $(\boldsymbol{\beta}^*, \boldsymbol{\rho}^*, \delta^*)$ is always an interior point of the parameter space (i.e., $\boldsymbol{\rho}^* \notin \eth$, $\mathbf{O}^* \succ \mathbf{0}_V$, and $\mathcal{N}[\mathbf{O}^*]$ is empty) in practice even if the true value of $\boldsymbol{\rho}$ is on the boundary. As a result, we can always adopt $\mathbf{O}^{(0)} \succ \mathbf{0}_V$ in the ECME algorithm, e.g., the simulation results in Figure 1 related to coherent sources.

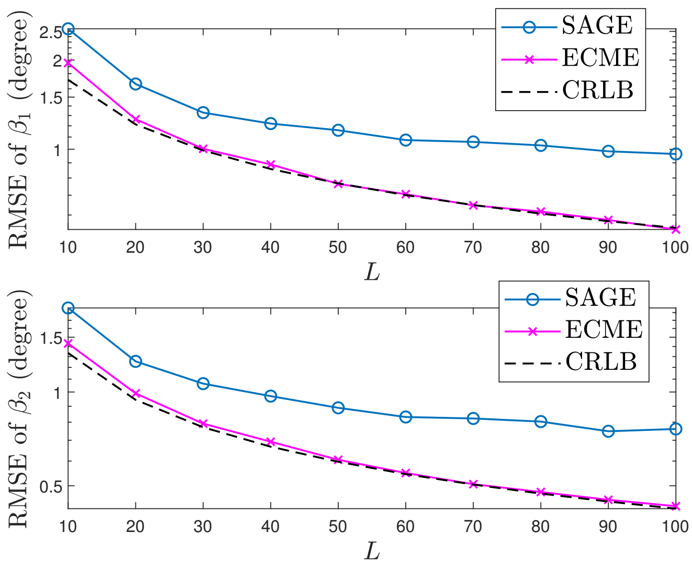

Figure 1. Relationship between the RMSE performance of the ECME algorithm and the CRLB. $\gamma = 2$, $\beta_1^{(0)} = 45°$, $\mathbf{O}^{(0)} = \mathbf{I}_V$, $\beta_2^{(0)} = 95°$, and $\mathbf{Q}^{(0)} = \mathbf{I}_W$.

4. Simulation Results

Simulation results are provided to confirm the effectiveness of the ECME algorithm, i.e., the ECME algorithm is able to obtain the ML estimate of $\boldsymbol{\beta}$ in (3). We set $V = 2$, $\beta_1 = 50°$, $W = 6$, $\beta_2 = 100°$, and $\boldsymbol{\delta} = [1\ 2\ 3\ 4\ 2\ 10]^T$. Algorithm 1 is used to obtain $\boldsymbol{\beta}^{(d)}$ in (20) and $\|\boldsymbol{\beta}^{(u+1)} - \boldsymbol{\beta}^{(u)}\|_2 \leq 0.001°$ is adopted as the stopping criterion. The ECME algorithm is given an accurate initial point for obtaining the ML estimate of $\boldsymbol{\beta}$. In Figures 1 and 2, we consider the coherent (or fully correlated) source model with [25]

$$\mathbf{O} = \begin{bmatrix} \gamma & \gamma \\ \gamma & \gamma \end{bmatrix} = \begin{bmatrix} \sqrt{\gamma} & 0 \\ 0 & \sqrt{\gamma} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{\gamma} & 0 \\ 0 & \sqrt{\gamma} \end{bmatrix}.$$

In Figure 3, we consider the partly correlate source model with

$$\mathbf{O} = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix} = \begin{bmatrix} \sqrt{5} & 0 \\ 0 & \sqrt{5} \end{bmatrix} \begin{bmatrix} 1 & 4/5 \\ 4/5 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{5} & 0 \\ 0 & \sqrt{5} \end{bmatrix}.$$

In Figures 1 and 2, we compare the root mean square error (RMSE) performance of the ECME algorithm with the CRLB [4,5]. In addition, we simulate the second space-alternating generalized EM (SAGE) algorithm for uncorrelated sources in [20], and this SAGE algorithm adopts the same simulation settings in [20]. Each RMSE is based on 2000 independent trials and the two algorithms share the same initial point. As expected, the ECME algorithm obtains smaller RMSEs than the SAGE algorithm. More importantly, the ECME algorithm attains the CRLB of $\boldsymbol{\beta}$ when the number of snapshots L or γ is large, which coincides with the well-known conclusion that *the stochastic CRLB of $\boldsymbol{\beta}$ can be achieved asymptotically by the stochastic ML estimator of $\boldsymbol{\beta}$* [2]. Hence, the ECME algorithm is able to obtain the stochastic ML estimate of $\boldsymbol{\beta}$ in (3) given an accurate initial point.

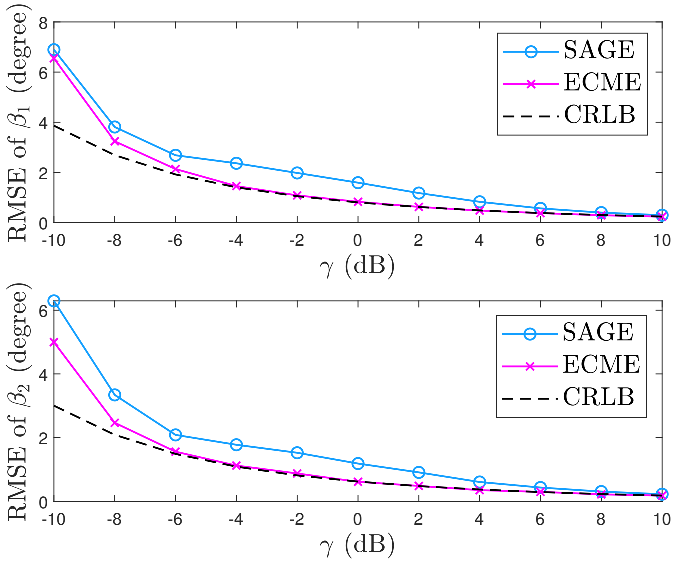

Figure 2. Relationship between the RMSE performance of the ECME algorithm and the CRLB. $L = 100$, $\beta_1^{(0)} = 45°$, $\mathbf{O}^{(0)} = \mathbf{I}_V$, $\beta_2^{(0)} = 95°$, and $\mathbf{Q}^{(0)} = \mathbf{I}_W$.

In Figure 3, we compare the ECME algorithm with two subspace-based algorithms, which utilize the state-of-the-art subspace separation approaches in [17,18] and are called "Approach 1+Root-MUSIC" and "Approach 2+Root-MUSIC", respectively. The three algorithms process the same snapshots of each trial. As expected, the ECME algorithm yields more closely spaced estimates of (β_1, β_2) centered on $(50°, 100°)$ since in DOA estimation, the ML technique offers the highest advantage in terms of accuracy.

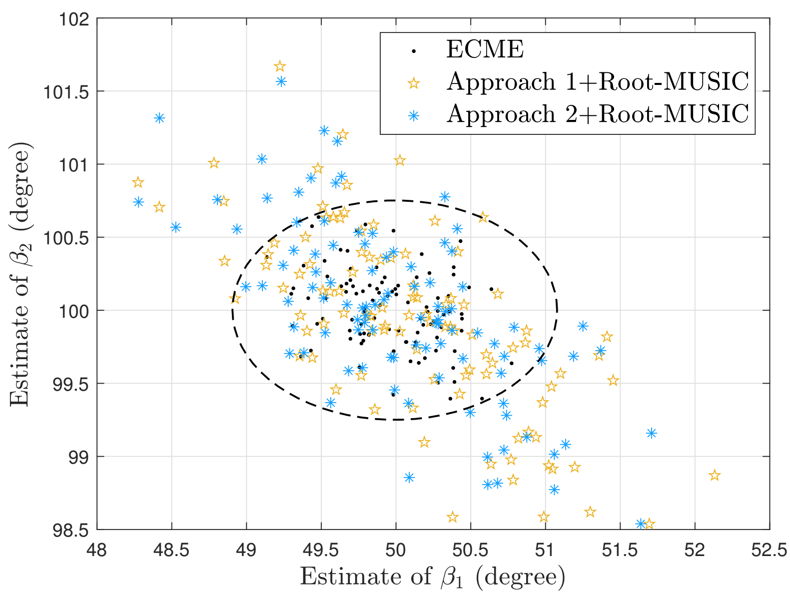

Figure 3. Estimates of (β_1, β_2) obtained from the ECME and two subspace-based algorithms under 100 independent trials. $\beta_1^{(0)} = 45°$, $\beta_2^{(0)} = 95°$, $\mathbf{O}^{(0)} = \mathbf{I}_V$, $L = 100$, and $\mathbf{Q}^{(0)} = \mathbf{I}_W$.

5. Conclusions

In this paper, we employed and designed the ECME algorithm for stochastic ML direction finding, where sources may be correlated, in unknown nonuniform noise. Theoretical analysis indicated that the ECME algorithm is computationally efficient and operationally stable. Simulation results confirmed that the ECME algorithm can efficiently obtain the ML based DOA estimate of each stochastic source.

Author Contributions: This paper was co-authored by M.-Y.G. and B.L. Conceptualization, M.-Y.G.; methodology, M.-Y.G.; software, M.-Y.G.; validation, M.-Y.G. and B.L.; formal analysis, M.-Y.G.; investigation, M.-Y.G.; resources, B.L.; data curation, M.-Y.G.; writing—original draft preparation, M.-Y.G.; writing—review and editing, B.L.; supervision, B.L.; funding acquisition, B.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Open Research Project of Jiangsu Provincial Key Laboratory of Photonic and Electronic Materials Sciences and Technology, grant number NJUZDS 2022-008.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available, due to the data in this paper not being from publicly available datasets but obtained from the simulation of the signal models listed in the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Stoica, P.; Nehorai, A. MUSIC, maximum likelihood, and Cramer-Rao bound. *IEEE Trans. Acoust. Speech Signal Process.* **1989**, *37*, 720–741. [CrossRef]
2. Stoica, P.; Nehorai, A. Performance study of conditional and unconditional direction-of-arrival estimation. *IEEE Trans. Acoust. Speech Signal Process.* **1990**, *38*, 1783–1795. [CrossRef]
3. Stoica, P.; Larsson, E.G.; Gershman, A.B. The stochastic CRB for array processing: A textbook derivation. *IEEE Signal Process. Lett.* **2001**, *8*, 148–150. [CrossRef]
4. Pesavento, M.; Gershman, A.B. Maximum-likelihood direction-of-arrival estimation in the presence of unknown nonuniform noise. *IEEE Trans. Signal Process.* **2001**, *37*, 1310–1324. [CrossRef]
5. Gershman, A.B.; Pesavento, M.; Stoica, P.; Larsson, E.G. The stochastic CRB for array processing in unknown noise fields. In Proceedings of the 2001 IEEE International Conference on Acoustics, Speech, and Signal Processing, Proceedings (Cat. No. 01CH37221), Salt Lake City, UT, USA, 22–27 May 2001.
6. Delmas, J.; Abeida, H. Stochastic Cramer-Rao bound for noncircular signals with application to DOA estimation. *IEEE Trans. Signal Process.* **2004**, *52*, 3192–3199. [CrossRef]
7. Abeida, H.; Delmas, J. Gaussian Cramer-Rao bound for direction estimation of noncircular signals in unknown noise fields. *IEEE Trans. Signal Process.* **2005**, *53*, 4610–4618. [CrossRef]
8. Ziskind, I.; Wax, M. Maximum likelihood localization of multiple sources by alternating projection. *IEEE Trans. Acoust. Speech Signal Process.* **1988**, *36*, 1553–1560. [CrossRef]
9. Miller, M.I.; Fuhrmann, D.R. Maximum-likelihood narrow-band direction finding and the EM algorithm. *IEEE Trans. Acoust. Speech Signal Process.* **1990**, *38*, 1560–1577. [CrossRef]
10. Chung, P.; Bohme, J.F. Comparative convergence analysis of EM and SAGE algorithms in DOA estimation. *IEEE Trans. Signal Process.* **2001**, *49*, 2940–2949. [CrossRef]
11. Gong, M.; Lyu, B. Alternating maximization and the EM algorithm in maximum-likelihood direction finding. *IEEE Trans. Veh. Technol.* **2021**, *70*, 9634–9645. [CrossRef]
12. EM and SAGE Algorithms for DOA Estimation in the Presence of Unknown Uniform Noise. Available online: https://arxiv.org/abs/2208.07510 (accessed on 16 August 2022).
13. Zoubir, A.M.; Aouada, S. High resolution estimation of directions of arrival in nonuniform noise. In Proceedings of the 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing, Montreal, QC, Canada, 17–21 May 2004.
14. Madurasinghe, D. A new DOA estimator in nonuniform noise. *IEEE Signal Process. Lett.* **2005**, *12*, 337–339. [CrossRef]
15. Liao, B.; Chan, S.; Huang, L.; Guo, C. Iterative methods for subspace and DOA estimation in nonuniform noise. *IEEE Trans. Signal Process.* **2016**, *64*, 3008–3020. [CrossRef]
16. Liao, B.; Huang, L.; Guo, C.; So, H.C. New approaches to direction-of-arrival estimation with sensor arrays in unknown nonuniform noise. *IEEE Sens. J.* **2016**, *16*, 8982–8989. [CrossRef]
17. Esfandiari, M.; Vorobyov, S.A.; Alibani, S.; Karimi, M. Non-iterative subspace-based DOA estimation in the presence of nonuniform noise. *IEEE Signal Process. Lett.* **2019**, *26*, 848–852. [CrossRef]
18. Esfandiari, M.; Vorobyov, S.A. A novel angular estimation method in the presence of nonuniform noise. In Proceedings of the ICASSP 2022—2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Singapore, 23–27 May 2022.
19. Chen, C.E.; Lorenzelli, F.; Hudson, R.E.; Yao, K. Stochastic maximum-likelihood DOA estimation in the presence of unknown nonuniform noise. *IEEE Trans. Signal Process.* **2008**, *56*, 3038–3044. [CrossRef]
20. EM-Type Algorithms for DOA Estimation in Unknown Nonuniform Noise. Available online: https://arxiv.org/abs/2211.02458 (accessed on 4 November 2022).

21. Liu, C.; Rubin, D.B. The ECME algorithm: A simple extension of EM and ECM with faster monotone convergence. *Biometrika* **1994**, *81*, 633–648. [CrossRef]
22. Jaffer, A.G. Maximum likelihood direction finding of stochastic sources: A separable solution. In Proceedings of the ICASSP-88, International Conference on Acoustics, Speech, and Signal Processing, New York, NY, USA, 11–14 April 1988.
23. Stoica, P.; Nehorai, A. On the concentrated stochastic likelihood function in array signal processing. *Circuits Syst. Signal Process.* **1995**, *14*, 669–674. [CrossRef]
24. Bresler, Y. Maximum likelihood estimation of linearly stmctured covariance with application to antenna array processing. In Proceedings of the 4th ASSP Workshop Spectrum Estimation Modeling, Minneapolis, MN, USA, 3–5 August 1988.
25. Stoica, P.; Ottersten, B.; Viberg, M.; Moses, R.L. Maximum likelihood array processing for stochastic coherent sources. *IEEE Trans. Signal Process.* **1996**, *44*, 96–105. [CrossRef]
26. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Statist. Soc. Ser. B (Methodol.)* **1977**, *39*, 1–38.
27. Rhodes, I.B. A tutorial introduction to estimation and filtering. *IEEE Trans. Autom. Control* **1971**, *16*, 688–706. [CrossRef]
28. Jeff Wu, C.F. On the convergence properties of the EM algorithm. *Ann. Statist.* **1983**, *11*, 95–103.
29. Stoica, P.; Gershman, A.B. Maximum-likelihood DOA estimation by data-supported grid search. *IEEE Signal Process. Lett.* **1999**, *6*, 273–275. [CrossRef]
30. Malioutov, D.; Cetin, M.; Willsky, A.S. A sparse signal reconstruction perspective for source localization with sensor arrays. *IEEE Trans. Signal Process.* **2005**, *53*, 3010–3022. [CrossRef]