

Article

# EM and SAGE Algorithms for DOA Estimation in the Presence of Unknown Uniform Noise

Ming-Yan Gong <sup>1,\*</sup>  and Bin Lyu <sup>2</sup> <sup>1</sup> School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China<sup>2</sup> 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 existing expectation maximization (EM) and space-alternating generalized EM (SAGE) algorithms are only applied to direction of arrival (DOA) estimation in known noise. In this paper, the two algorithms are designed for DOA estimation in unknown uniform noise. Both the deterministic and random signal models are considered. In addition, a new modified EM (MEM) algorithm applicable to the noise assumption is also proposed. Next, these EM-type algorithms are improved to ensure the stability when the powers of sources are not equal. After being improved, simulation results illustrate that the EM algorithm has similar convergence with the MEM algorithm, the SAGE algorithm outperforms the EM and MEM algorithms for the deterministic signal model, and the SAGE algorithm cannot always outperform the EM and MEM algorithms for the random signal model. Furthermore, simulation results show that processing the same snapshots from the random signal model, the SAGE algorithm for the deterministic signal model can require the fewest computations.

**Keywords:** array signal processing; DOA estimation; EM algorithm; maximum likelihood estimation; statistical signal processing



**Citation:** Gong, M.-Y.; Lyu, B. EM and SAGE Algorithms for DOA Estimation in the Presence of Unknown Uniform Noise. *Sensors* **2023**, *23*, 4811. <https://doi.org/10.3390/s23104811>

Academic Editors: Chee Kiat Seow, Henrik Hesse, Kai Wen, Soon Yim Tan and Yunjia Wang

Received: 27 April 2023

Revised: 9 May 2023

Accepted: 15 May 2023

Published: 16 May 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Direction of arrival (DOA) estimation is an important part of array signal processing and some high-resolution estimation techniques have been developed in the literature [1,2]. In particular, the maximum likelihood (ML) technique plays a critical role since it can offer the highest advantage in terms of both accuracy and spatial resolution. However, ML direction finding problems are non-convex and difficult to obtain their solutions in closed form.

One computationally efficient method to solve ML estimation problems is the classic expectation maximization (EM) algorithm [3,4], which has been employed for ML direction finding [5,6]. Each iteration of the EM algorithm is composed of an expectation step (E-step) and a maximization step (M-step). At the M-step, however, the EM algorithm updates all of the parameter estimates simultaneously, which causes slow convergence. In order to speed up the convergence of the EM algorithm, the space-alternating generalized EM (SAGE) algorithm has been proposed in [7]. References [8,9] show that the SAGE algorithm does yield faster convergence in terms of DOA estimation.

The existing EM and SAGE algorithms are usually derived under known noise [5,6,8,9]. The known noise is without unknown parameters, which may be unrealistic in certain applications. In fact, many seminal works in ML direction finding consider the so-called unknown uniform noise model [1,2,10–12], i.e., the noise covariance matrix can be expressed as  $\tau \mathbf{I}_K$ , where  $\tau$  is the only unknown noise parameter and  $\mathbf{I}_K$  is the  $K \times K$  identity matrix. Under this noise assumption, a computationally attractive alternating projection algorithm is presented for computing the deterministic ML based DOA estimator in [10]. The authors in [11] investigate the statistical performance of this ML estimator and derive the Cramer–Rao lower bound. Moreover, some statistical properties of both the deterministic

and random ML based DOA estimators under unknown uniform noise are compared in [12]. In addition to uniform noise, non-uniform noise has also attracted increasing attention [13–16]. The non-uniform noise has an arbitrary diagonal covariance matrix and, thus, makes DOA estimation complex. For efficiently computing both the deterministic and random ML based DOA estimators in unknown non-uniform noise, some feasible algorithms have been proposed in [17–20].

In this paper, we apply and design the EM and SAGE algorithms for DOA estimation in unknown uniform noise. Theoretical analyses indicate that for the deterministic signal model,  $\tau$  has little effect on the two algorithms. However, the problem in the M-step of the EM algorithm for the random signal model can be no longer decomposed into parallel subproblems easily when  $\tau$  is unknown. To proceed, we divide the M-step into two conditional M-steps (CM-steps) based on the expectation CM (ECM) algorithm [21]. In addition, we propose a new modified EM (MEM) algorithm applicable to the unknown uniform noise assumption. Note that although the EM algorithm in [22] is similar to the MEM algorithm, it is incorrectly derived. In brief, the proposed EM-type algorithms only need low-dimensional numerical searches at each iteration and are easy to perform. However, the proposed EM-type algorithms require accurate initial points, which is generally a computationally expensive task.

Existing simulations using EM-type algorithms always adopt sources of equal power [5,6,8,9,22]. We find, however, that when the powers of sources are unequal, the DOA estimates of multiple sources obtained by these EM-type algorithms tend to be consistent with the true DOA of the source with the largest power. To this end, we improve the proposed EM-type algorithms. After being improved, simulation results illustrate that (1) the EM algorithm has similar convergence with the MEM algorithm, (2) the SAGE algorithm outperforms the EM and MEM algorithms for the deterministic signal model, i.e., the SAGE algorithm converges faster and can avoid the convergence to an undesirable stationary point of the log-likelihood function (LLF) more efficiently, and (3) the SAGE algorithm cannot always outperform the EM and MEM algorithms for the random signal model.

The proposed EM-type algorithms for the deterministic signal model can process snapshots from the random signal model, so we, via simulation, compare these algorithms for both signal models. Simulations show that, under the same snapshots, initial DOA estimates, and stopping criterion, the SAGE algorithm for the deterministic signal model can require the fewest iterations and computations.

The contributions of this paper can be enumerated as follows:

- We apply and design the EM and SAGE algorithms for DOA estimation in unknown uniform noise. In particular, we derive the SAGE algorithm for random ML direction finding, which is not discussed in [7,8,22].
- We propose a new MEM algorithm applicable to the unknown uniform noise assumption.
- We improve these EM-type algorithms to ensure the stability when the powers of sources are not equal.
- Via simulation we show that the EM algorithm has similar convergence with the MEM algorithm and the SAGE algorithm outperforms the EM and MEM algorithms for the deterministic signal model. However, the SAGE algorithm cannot always outperform the EM and MEM algorithms for the random signal model.
- Via simulation we show that processing the same snapshots from the random signal model, the SAGE algorithm for the deterministic signal model can require the fewest iterations and computations.

The rest of this paper is outlined as follows: in Section 2, we formulate both the deterministic and random ML direction finding problems in unknown uniform noise. In Sections 3–5, we design the EM, MEM, and SAGE algorithms, respectively. We analyze some convergence properties of these EM-type algorithms in Section 6 and provide simulation results to compare the convergence of these EM-type algorithms in Section 7. Finally, we conclude this paper in Section 8.

## 2. Signal Model and Problem Statement

An array of  $K$  sensors is assumed to receive the plane waves emitted from  $G$  narrow-band sources, which share the same known center wavelength  $\chi$ . We use the Cartesian coordinate  $\zeta_k = [x_k \ y_k \ z_k]^T$  and the Spherical coordinate  $(1, \mu_g, \eta_g)$  to locate the  $k$ th sensor and the direction of the  $g$ th source, respectively. Here,  $[\cdot]^T$  denotes transpose,  $\mu_g$  and  $\eta_g$  denote the elevation and azimuth angles of the  $g$ th source, respectively. For convenience, we transform  $(1, \mu_g, \eta_g)$  into the corresponding Cartesian coordinate  $\gamma_g = [\sin(\mu_g) \cos(\eta_g) \ \sin(\mu_g) \sin(\eta_g) \ \cos(\mu_g)]^T$ . Let the origin be the reference point such that the signal received at the array is written as

$$\mathbf{w}(t) = \sum_{g=1}^G \mathbf{b}(\zeta_g) m_g(t) + \mathbf{v}(t) = \mathbf{B}(\boldsymbol{\zeta}) \mathbf{m}(t) + \mathbf{v}(t), \tag{1}$$

where  $\mathbf{b}(\zeta_g) = [e^{j\phi_{1,g}} \ \dots \ e^{j\phi_{K,g}}]^T$  with DOA  $\zeta_g = (\mu_g, \eta_g) \in \Omega$ ,  $\phi_{k,g} = -\frac{2\pi}{\chi} \zeta_k^T \gamma_g$ , and  $j = \sqrt{-1}$ ,  $m_g(t)$  is the signal of the  $g$ th source, and  $\mathbf{v}(t)$  is complex Gaussian noise of zero mean and covariance  $\tau \mathbf{I}_K$ , i.e.,  $\mathbf{v}(t) \sim \mathcal{CN}(\mathbf{0}, \tau \mathbf{I}_K)$  with  $\mathbf{0} = [0 \ \dots \ 0]^T$ . In (1),  $\mathbf{B}(\boldsymbol{\zeta}) = [\mathbf{b}(\zeta_1) \ \dots \ \mathbf{b}(\zeta_G)]$ ,  $\boldsymbol{\zeta} = (\zeta_1, \dots, \zeta_G) \in \Omega$  with  $\Omega = \Omega^G$ , and  $\mathbf{m}(t) = [m_1(t) \ \dots \ m_G(t)]^T$ .

EM-type algorithms need to define the unavailable complete data. According to the classic EM paradigm for superimposed signals [5,6], we construct  $L$  independent snapshots, the incomplete data of the EM algorithm, by

$$\mathbf{w}(t) = \sum_{g=1}^G [\mathbf{b}(\zeta_g) m_g(t) + \mathbf{v}_g(t)] = \sum_{g=1}^G \mathbf{h}_g(t), t = 1, 2, \dots, L, \tag{2}$$

where the  $\mathbf{h}_g(t)$ 's are the complete data. Moreover, the  $\mathbf{v}_g(t)$ 's are mutually uncorrelated and  $\mathbf{v}_g(t) \sim \mathcal{CN}(\mathbf{0}, \beta_g \tau \mathbf{I}_K)$ , where  $\boldsymbol{\beta} = [\beta_1 \ \dots \ \beta_G]^T > \mathbf{0}$  and  $\mathbf{1}^T \boldsymbol{\beta} = 1$  with  $\mathbf{1} = [1 \ \dots \ 1]^T$ . Note that the incomplete- and complete-data LLFs require the distributions of the  $m_g(t)$ 's, we adopt the following two statistical models separately.

### 2.1. Deterministic Signal Model

We let the  $m_g(t)$ 's be deterministic and unknown [5,6,10,11], which leads to  $\mathbf{h}_g(t) \sim \mathcal{CN}(\mathbf{b}(\zeta_g) m_g(t), \beta_g \tau \mathbf{I}_K)$  and  $\mathbf{w}(t) \sim \mathcal{CN}(\mathbf{B}(\boldsymbol{\zeta}) \mathbf{m}(t), \tau \mathbf{I}_K)$ . Then, the incomplete- and complete-data LLFs are formulated by

$$\begin{aligned} \mathcal{J}(\boldsymbol{\Psi}, \tau) &= \sum_{t=1}^L \log p(\mathbf{w}(t); \boldsymbol{\zeta}, \mathbf{m}(t), \tau) \\ &= -LK \log(\pi\tau) - \frac{1}{\tau} \sum_{t=1}^L \|\mathbf{w}(t) - \mathbf{B}(\boldsymbol{\zeta}) \mathbf{m}(t)\|^2, \end{aligned} \tag{3}$$

$$\begin{aligned} \mathcal{Z}(\boldsymbol{\Psi}, \tau) &= \sum_{t=1}^L \sum_{g=1}^G \log p(\mathbf{h}_g(t); \zeta_g, m_g(t), \tau) \\ &= -LKG \log(\pi\tau) - LK \sum_{g=1}^G \log(\beta_g) - \frac{1}{\tau} \sum_{t=1}^L \sum_{g=1}^G \frac{1}{\beta_g} \|\mathbf{h}_g(t) - \mathbf{b}(\zeta_g) m_g(t)\|^2, \end{aligned} \tag{4}$$

where  $\|\cdot\|$  denotes Euclidean norm and  $\mathbf{M} = [\mathbf{m}(1) \ \dots \ \mathbf{m}(L)]$ . Note that  $\boldsymbol{\Psi} = (\boldsymbol{\zeta}, \mathbf{M})$  denotes the signal parameters while  $\tau$  is the only noise parameter. Finally, the ML estimation problem is

$$\max_{\boldsymbol{\zeta} \in \Omega, \mathbf{M}, \tau > 0} \mathcal{J}(\boldsymbol{\Psi}, \tau). \tag{5}$$

## 2.2. Random Signal Model

We assume  $m_g(t) \sim \mathcal{CN}(0, \rho_g)$  where  $\rho_g$  is the power of the  $g$ th source. For simplicity, all of the  $m_g(t)$ 's and  $v_g(t)$ 's are assumed to be mutually uncorrelated [5,6]. Next, we have  $\mathbf{h}_g(t) \sim \mathcal{CN}(\mathbf{0}, \mathbf{N}_g)$  with  $\mathbf{N}_g = \rho_g \mathbf{b}(\zeta_g) \mathbf{b}^H(\zeta_g) + \beta_g \tau \mathbf{I}_K$ , where  $[\cdot]^H$  denotes conjugate transpose, and  $\mathbf{w}(t) \sim \mathcal{CN}(\mathbf{0}, \mathbf{N}_w)$  with  $\mathbf{N}_w = \sum_{g=1}^G \mathbf{N}_g$ . The incomplete- and complete-data LLFs are formulated by

$$\begin{aligned} \mathcal{J}(\Psi, \tau) &= \sum_{t=1}^L \log p(\mathbf{w}(t); \zeta, \rho, \tau) \\ &= -L [K \log(\tau) + \log(\text{Det}(\mathbf{N}_w)) + \text{Tr}(\mathbf{N}_w^{-1} \hat{\mathbf{P}}_w)], \end{aligned} \quad (6)$$

$$\begin{aligned} \mathcal{Z}(\Psi, \tau) &= \sum_{t=1}^L \sum_{g=1}^G \log p(\mathbf{h}_g(t); \zeta_g, \rho_g, \tau) \\ &= -L \sum_{g=1}^G [K \log(\tau) + \log(\text{Det}(\mathbf{N}_g)) + \text{Tr}(\mathbf{N}_g^{-1} \hat{\mathbf{P}}_g)], \end{aligned} \quad (7)$$

where  $[\cdot]^{-1}$ ,  $\text{Det}(\cdot)$ , and  $\text{Tr}(\cdot)$  denote inversion, determinant, and trace, respectively. Moreover,  $\Psi = (\zeta, \rho)$ ,  $\hat{\mathbf{P}}_w = (1/L) \sum_{t=1}^L \mathbf{w}(t) \mathbf{w}^H(t)$ ,  $\hat{\mathbf{P}}_g = (1/L) \sum_{t=1}^L \mathbf{h}_g(t) \mathbf{h}_g^H(t)$ , and  $\rho = [\rho_1 \cdots \rho_G]^T$ . Finally, the ML estimation problem is

$$\max_{\zeta \in \Omega, \rho \geq 0, \tau > 0} \mathcal{J}(\Psi, \tau). \quad (8)$$

## 3. EM Algorithm

In this section, we design and derive the EM algorithm for solving problems (5) and (8). The E- and M-steps at the  $r$ th iteration are introduced below. Let  $[\cdot]^{(r)}$ ,  $\mathcal{E}\{\cdot\}$ , and  $\mathcal{D}\{\cdot\}$  denote an iterative value at the  $r$ th iteration, expectation, and covariance, respectively.  $[\cdot]^{(0)}$  is an initial estimate.

### 3.1. Deterministic Signal Model

#### 3.1.1. E-Step

Calculate the conditional expectation of the complete-data LLF in (4)

$$\begin{aligned} \mathcal{E}\{\mathcal{Z}(\Delta) | \mathbf{W}; \Delta^{(r-1)}\} &= C - L \left\{ KG \log(\tau) + K \sum_{g=1}^G \log(\beta_g) + \frac{1}{\tau} [u^{(r)} + \right. \\ &\quad \left. \frac{1}{L} \sum_{t=1}^L \sum_{g=1}^G \frac{1}{\beta_g} \|\mathbf{h}_g^{(r)}(t) - \mathbf{b}(\zeta_g) m_g(t)\|^2] \right\}, \end{aligned} \quad (9)$$

where  $\Delta = (\Psi, \tau) = (\zeta, \mathbf{M}, \tau)$ ,  $\mathbf{W} = [\mathbf{w}(1) \cdots \mathbf{w}(L)]$ , and  $C = -LKG \log(\tau)$ . In (9), the conditional distribution of  $\mathbf{h}_g(t)$  can be derived from [23] and

$$\begin{aligned} \mathbf{h}_g^{(r)}(t) &= \mathcal{E}\{\mathbf{h}_g(t) | \mathbf{W}; \Delta^{(r-1)}\} \\ &= \mathbf{b}(\zeta_g^{(r-1)}) m_g^{(r-1)}(t) + \beta_g [\mathbf{w}(t) - \mathbf{B}(\zeta^{(r-1)}) \mathbf{m}^{(r-1)}(t)], \end{aligned} \quad (10)$$

$$\begin{aligned} u^{(r)} &= \frac{1}{L} \sum_{t=1}^L \sum_{g=1}^G \frac{1}{\beta_g} \text{Tr}(\mathcal{D}\{\mathbf{h}_g(t) | \mathbf{W}; \Delta^{(r-1)}\}) = \frac{1}{L} \sum_{t=1}^L \sum_{g=1}^G \frac{1}{\beta_g} \text{Tr}(\beta_g (1 - \beta_g) \tau^{(r-1)} \mathbf{I}_K) \\ &= K(G - 1) \tau^{(r-1)}. \end{aligned} \quad (11)$$

### 3.1.2. M-Step

Update the estimates of  $\Psi$  and  $\tau$  by solving

$$\min_{\xi \in \Omega, \mathbf{M}, \tau > 0} KG \log(\tau) + \frac{1}{\tau} [u^{(r)} + \frac{1}{L} \sum_{t=1}^L \sum_{g=1}^G \frac{1}{\beta_g} \|\mathbf{h}_g^{(r)}(t) - \mathbf{b}(\xi_g)m_g(t)\|^2]. \tag{12}$$

$\Psi^{(r)} = (\xi^{(r)}, \mathbf{M}^{(r)})$  and  $\tau^{(r)}$  are obtained by [6]

$$\xi_g^{(r)} = \arg \max_{\xi_g \in \Omega} \text{Tr}(\Gamma_g \hat{\mathbf{P}}_g^{(r)}), \forall g, \tag{13}$$

$$m_g^{(r)}(t) = \mathbf{b}^H(\xi_g^{(r)})\mathbf{h}_g^{(r)}(t)/K, \forall g, t, \tag{14}$$

$$\tau^{(r)} = (1 - 1/G)\tau^{(r-1)} + (1/K/G)\sum_{g=1}^G d_g^{(r)}/\beta_g, \tag{15}$$

where  $\Gamma_g = \mathbf{b}(\xi_g)\mathbf{b}^H(\xi_g)/K$ ,  $\hat{\mathbf{P}}_g^{(r)} = (1/L)\sum_{t=1}^L \mathbf{h}_g^{(r)}(t)[\mathbf{h}_g^{(r)}(t)]^H$ , and  $d_g^{(r)} = \text{Tr}((\mathbf{I}_K - \Gamma_g^{(r)})\hat{\mathbf{P}}_g^{(r)}) \geq 0$ . In (15),  $\tau^{(r)} > 0$  if  $\tau^{(r-1)} > 0$ .

**Remark 1.** Note that the  $\mathbf{h}_g^{(r)}(t)$ 's in (10), the  $\xi_g^{(r)}$ 's in (13), and the  $m_g^{(r)}(t)$ 's in (14) are unrelated to  $\tau^{(r-1)}$ , we can omit (15) due to the nuisance parameter  $\tau$ .

### 3.2. Random Signal Model

#### 3.2.1. E-Step

Calculate the conditional expectation of the complete-data LLF in (7)

$$\mathcal{E}\{\mathcal{Z}(\Delta)|\mathbf{W}; \Delta^{(r-1)}\} = C - L \sum_{g=1}^G [\log(\text{Det}(\mathbf{N}_g)) + \text{Tr}(\mathbf{N}_g^{-1}\hat{\mathbf{P}}_g^{(r)})], \tag{16}$$

where  $\Delta = (\Psi, \tau) = (\xi, \rho, \tau)$  and

$$\begin{aligned} \hat{\mathbf{P}}_g^{(r)} &= \mathcal{E}\{\hat{\mathbf{P}}_g|\mathbf{W}; \Delta^{(r-1)}\} \\ &= \mathbf{N}_g^{(r-1)} - \mathbf{N}_g^{(r-1)}[\mathbf{N}_w^{(r-1)}]^{-1}\mathbf{N}_g^{(r-1)} + \mathbf{N}_g^{(r-1)}[\mathbf{N}_w^{(r-1)}]^{-1}\hat{\mathbf{P}}_w[\mathbf{N}_w^{(r-1)}]^{-1}\mathbf{N}_g^{(r-1)}. \end{aligned} \tag{17}$$

#### 3.2.2. M-Step

Update the estimates of  $\Psi$  and  $\tau$  by solving

$$\min_{\xi \in \Omega, \rho \geq 0, \tau > 0} \sum_{g=1}^G [\log(\text{Det}(\mathbf{N}_g)) + \text{Tr}(\mathbf{N}_g^{-1}\hat{\mathbf{P}}_g^{(r)})], \tag{18}$$

which is difficult to be decomposed into parallel subproblems due to  $\tau$ . To obtain  $\Psi^{(r)}$  and  $\tau^{(r)}$  easily, we rewrite (18) as

$$\min_{\xi \in \Omega, \sigma \geq 0, \tau > 0} KG \log(\tau) + \sum_{g=1}^G [\log(\text{Det}(\mathbf{Q}_g)) + \frac{1}{\tau} \text{Tr}(\mathbf{Q}_g^{-1}\hat{\mathbf{P}}_g^{(r)})], \tag{19}$$

where  $\mathbf{N}_g = \tau\mathbf{Q}_g$  with  $\mathbf{Q}_g = \sigma_g\mathbf{b}(\xi_g)\mathbf{b}^H(\xi_g) + \beta_g\mathbf{I}_K$  and  $\sigma = [\sigma_1 \cdots \sigma_G]^T$  with  $\sigma_g = \rho_g/\tau$ . We now divide the M-step into the following two CM-steps based on the ECM algorithm [21], i.e., the algorithm becomes the ECM algorithm. For convenience, we still call it the EM algorithm.

- *First CM-step:* Estimate  $\Psi$  but hold  $\tau = \tau^{(r-1)}$  fixed. Then, problem (19) can be decomposed into the  $G$  parallel subproblems

$$\min_{\xi_g \in \Omega, \sigma_g \geq 0} \log(\text{Det}(\mathbf{Q}_g)) + \frac{1}{\tau^{(r-1)}} \text{Tr}(\mathbf{Q}_g^{-1} \hat{\mathbf{P}}_g^{(r)}), \forall g. \quad (20)$$

$\Psi^{(r)} = (\xi^{(r)}, \rho^{(r)})$  is obtained by [6]

$$\xi_g^{(r)} = \arg \max_{\xi_g \in \Omega} \text{Tr}(\Gamma_g \hat{\mathbf{P}}_g^{(r)}), \forall g, \quad (21)$$

$$\rho_g^{(r)} = \sigma_g^{(r)} \tau^{(r-1)} = \max\{(e_g^{(r)} - \beta_g \tau^{(r-1)})/K, 0\}, \forall g, \quad (22)$$

where  $e_g^{(r)} = \text{Tr}(\Gamma_g^{(r)} \hat{\mathbf{P}}_g^{(r)})$  and  $\xi_g^{(r)}$  is indeterminate if  $\rho_g^{(r)} = 0$ .

- *Second CM-step:* Estimate  $\tau$  but hold  $\Psi = \Psi^{(r)}$  fixed. Then, problem (19) is simplified to

$$\min_{\tau > 0} KG \log(\tau) + \frac{1}{\tau} \sum_{g=1}^G \text{Tr}([\mathbf{Q}_g^{(r)}]^{-1} \hat{\mathbf{P}}_g^{(r)}). \quad (23)$$

$\tau^{(r)}$  is obtained by

$$\tau^{(r)} = \frac{1}{KG} \sum_{g=1}^G \text{Tr}([\mathbf{Q}_g^{(r)}]^{-1} \hat{\mathbf{P}}_g^{(r)}) = \frac{1}{K} \tau^{(r-1)} + \frac{1}{KG} \sum_{g=1}^G d_g^{(r)} / \beta_g, \quad (24)$$

where  $d_g^{(r)} = \text{Tr}((\mathbf{I}_K - \Gamma_g^{(r)}) \hat{\mathbf{P}}_g^{(r)}) \geq 0$  and  $\tau^{(r)} > 0$  if  $\tau^{(r-1)} > 0$ .

#### 4. MEM Algorithm

In the previous section,  $\beta$  is fixed and known. In this section, we regard  $\beta$  as a parameter to be estimated and, thus, propose an MEM algorithm applicable to the unknown uniform noise assumption.

To estimate  $\tau$  and  $\beta$  easily, we introduce  $\tau_g = \beta_g \tau$  as the common noise variance of the  $g$ th source and have

$$\mathbf{v}_g(t) \sim \mathcal{CN}(\mathbf{0}, \tau_g \mathbf{I}_K). \quad (25)$$

Clearly,  $\tau = \sum_{g=1}^G \tau_g$  and  $\beta_g = \tau_g / \tau$ . The E- and M-steps at the  $r$ th iteration are introduced below. Let  $\boldsymbol{\tau} = [\tau_1 \cdots \tau_G]^T$ .

##### 4.1. Deterministic Signal Model

Based on (25), the complete-data LLF in (4) is rewritten as

$$\mathcal{Z}(\boldsymbol{\Delta}) = C - LK \sum_{g=1}^G \log(\tau_g) - \sum_{t=1}^L \sum_{g=1}^G \frac{1}{\tau_g} \|\mathbf{h}_g(t) - \mathbf{b}(\xi_g) m_g(t)\|^2, \quad (26)$$

where  $\boldsymbol{\Delta} = (\Psi, \boldsymbol{\tau}) = (\xi, \mathbf{M}, \boldsymbol{\tau})$  and  $\mathbf{h}_g(t) \sim \mathcal{CN}(\mathbf{b}(\xi_g) m_g(t), \tau_g \mathbf{I}_K)$ .

##### 4.1.1. E-Step

Calculate the conditional expectation of the complete-data LLF in (26)

$$\begin{aligned} \mathcal{E}\{\mathcal{Z}(\boldsymbol{\Delta}) | \mathbf{W}; \boldsymbol{\Delta}^{(r-1)}\} &= C - L \sum_{g=1}^G \left\{ K \log(\tau_g) + \frac{1}{\tau_g} [u_g^{(r)} + \right. \\ &\quad \left. \frac{1}{L} \sum_{t=1}^L \|\mathbf{h}_g^{(r)}(t) - \mathbf{b}(\xi_g) m_g(t)\|^2] \right\}, \end{aligned} \quad (27)$$

where

$$\begin{aligned} \mathbf{h}_g^{(r)}(t) &= \mathcal{E}\{\mathbf{h}_g(t)|\mathbf{W};\Delta^{(r-1)}\} \\ &= \mathbf{b}(\zeta_g^{(r-1)})m_g^{(r-1)}(t) + (\tau_g^{(r-1)}/\tau^{(r-1)})[\mathbf{w}(t) - \mathbf{B}(\zeta_g^{(r-1)})\mathbf{m}^{(r-1)}(t)], \end{aligned} \quad (28)$$

$$u_g^{(r)} = \frac{1}{L} \sum_{t=1}^L \text{Tr}(\mathcal{D}\{\mathbf{h}_g(t)|\mathbf{W};\Delta^{(r-1)}\}) = K\tau_g^{(r-1)}(1 - \tau_g^{(r-1)}/\tau^{(r-1)}). \quad (29)$$

#### 4.1.2. M-Step

Update the estimates of  $\Psi$  and  $\tau$  by solving the  $G$  parallel subproblems

$$\min_{\zeta_g \in \Omega, \mathbf{m}_g, \tau_g > 0} K \log(\tau_g) + \frac{1}{\tau_g} [u_g^{(r)} + \frac{1}{L} \sum_{t=1}^L \|\mathbf{h}_g^{(r)}(t) - \mathbf{b}(\zeta_g)m_g(t)\|^2], \forall g, \quad (30)$$

where  $\mathbf{m}_g = [m_g(1) \cdots m_g(L)]$ .  $\Psi^{(r)} = (\zeta^{(r)}, \mathbf{M}^{(r)})$  and  $\tau^{(r)}$  are obtained by

$$\zeta_g^{(r)} = \arg \max_{\zeta_g \in \Omega} \text{Tr}(\Gamma_g \hat{\mathbf{P}}_g^{(r)}), \forall g, \quad (31)$$

$$m_g^{(r)}(t) = \mathbf{b}^H(\zeta_g^{(r)})\mathbf{h}_g^{(r)}(t)/K, \forall g, t, \quad (32)$$

$$\tau_g^{(r)} = \tau_g^{(r-1)}(1 - \tau_g^{(r-1)}/\tau^{(r-1)}) + d_g^{(r)}/K, \forall g, \quad (33)$$

where  $\hat{\mathbf{P}}_g^{(r)} = (1/L) \sum_{t=1}^L \mathbf{h}_g^{(r)}(t)[\mathbf{h}_g^{(r)}(t)]^H$  and  $d_g^{(r)} = \text{Tr}((\mathbf{I}_K - \Gamma_g^{(r)})\hat{\mathbf{P}}_g^{(r)}) \geq 0$ . In (33), if  $\tau^{(r-1)} > 0$ , we have  $1 - \tau_g^{(r-1)}/\tau^{(r-1)} > 0, \forall g$ , and then  $\tau_g^{(r)} > 0, \forall g$ , i.e.,  $\tau^{(r)} > 0$ .

#### 4.2. Random Signal Model

Based on (25), the complete-data LLF in (7) is rewritten as

$$\mathcal{Z}(\Delta) = C - L \sum_{g=1}^G [\log(\text{Det}(\mathbf{N}_g)) + \text{Tr}(\mathbf{N}_g^{-1}\hat{\mathbf{P}}_g)], \quad (34)$$

where  $\Delta = (\Psi, \tau) = (\zeta, \rho, \tau)$  and  $\mathbf{N}_g = \rho_g \mathbf{b}(\zeta_g)\mathbf{b}^H(\zeta_g) + \tau_g \mathbf{I}_K$ .

##### 4.2.1. E-Step

Calculate the conditional expectation of the complete-data LLF in (34)

$$\mathcal{E}\{\mathcal{Z}(\Delta)|\mathbf{W};\Delta^{(r-1)}\} = C - L \sum_{g=1}^G [\log(\text{Det}(\mathbf{N}_g)) + \text{Tr}(\mathbf{N}_g^{-1}\hat{\mathbf{P}}_g^{(r)})], \quad (35)$$

where

$$\begin{aligned} \hat{\mathbf{P}}_g^{(r)} &= \mathcal{E}\{\hat{\mathbf{P}}_g|\mathbf{W};\Delta^{(r-1)}\} \\ &= \mathbf{N}_g^{(r-1)} - \mathbf{N}_g^{(r-1)}[\mathbf{N}_w^{(r-1)}]^{-1}\mathbf{N}_g^{(r-1)} + \mathbf{N}_g^{(r-1)}[\mathbf{N}_w^{(r-1)}]^{-1}\hat{\mathbf{P}}_w[\mathbf{N}_w^{(r-1)}]^{-1}\mathbf{N}_g^{(r-1)}. \end{aligned} \quad (36)$$

##### 4.2.2. M-Step

Update the estimates of  $\Psi$  and  $\tau$  by solving the  $G$  parallel subproblems

$$\min_{\zeta_g \in \Omega, \sigma_g \geq 0, \tau_g > 0} K \log(\tau_g) + \log(\text{Det}(\mathbf{Q}_g)) + \frac{1}{\tau_g} \text{Tr}(\mathbf{Q}_g^{-1}\hat{\mathbf{P}}_g^{(r)}), \forall g, \quad (37)$$

where  $\mathbf{N}_g = \tau_g \mathbf{Q}_g$  with  $\mathbf{Q}_g = \sigma_g \mathbf{b}(\zeta_g) \mathbf{b}^H(\zeta_g) + \mathbf{I}_K$  and  $\sigma_g = \rho_g / \tau_g$ . Since  $\text{Det}(\mathbf{Q}_g) = K\sigma_g + 1$  and  $\mathbf{Q}_g^{-1} = \mathbf{I}_K - \frac{K\sigma_g}{K\sigma_g + 1} \mathbf{\Gamma}_g$ , subproblems (37) are rewritten as

$$\min_{\zeta_g \in \Omega, \sigma_g \geq 0, \tau_g > 0} K \log(\tau_g) + \log(K\sigma_g + 1) + \frac{1}{\tau_g} \text{Tr}(\hat{\mathbf{P}}_g^{(r)}) - \frac{K\sigma_g}{\tau_g(K\sigma_g + 1)} \text{Tr}(\mathbf{\Gamma}_g \hat{\mathbf{P}}_g^{(r)}), \forall g. \quad (38)$$

To proceed, we first eliminate  $\sigma = [\sigma_1 \cdots \sigma_G]^T$  in (38) [24,25]. Thus, when obtaining  $\zeta^{(r)}$  and  $\tau^{(r)}$ ,  $\sigma^{(r)}$  and  $\rho^{(r)}$  are obtained by [6]

$$\sigma_g^{(r)} = \max \{ (e_g^{(r)} / \tau_g^{(r)} - 1) / K, 0 \}, \forall g, \quad (39)$$

$$\rho_g^{(r)} = \sigma_g^{(r)} \tau_g^{(r)} = \max \{ (e_g^{(r)} - \tau_g^{(r)}) / K, 0 \}, \forall g, \quad (40)$$

where  $e_g^{(r)} = \text{Tr}(\mathbf{\Gamma}_g \hat{\mathbf{P}}_g^{(r)})$ . Note that if  $\sigma_g^{(r)} = 0$ ,  $\zeta_g^{(r)}$  is indeterminate and  $\tau_g^{(r)} = \text{Tr}(\hat{\mathbf{P}}_g^{(r)}) / K$  by (38). To obtain  $\zeta^{(r)}$  and  $\tau^{(r)}$ , we assume  $\sigma^{(r)} > \mathbf{0}$ . After eliminating  $\sigma$ , subproblems (38) are simplified to

$$\min_{\zeta_g \in \Omega, \tau_g > 0} (K-1) \log(\tau_g) + \log(\text{Tr}(\mathbf{\Gamma}_g \hat{\mathbf{P}}_g^{(r)})) + \frac{1}{\tau_g} \text{Tr}((\mathbf{I}_K - \mathbf{\Gamma}_g) \hat{\mathbf{P}}_g^{(r)}), \forall g. \quad (41)$$

Next, we eliminate  $\tau$  in (41). Thus, when obtaining  $\zeta^{(r)}$ ,  $\tau^{(r)}$  is obtained by  $\tau_g^{(r)} = d_g^{(r)} / (K-1)$ ,  $\forall g$ , where  $d_g^{(r)} = \text{Tr}((\mathbf{I}_K - \mathbf{\Gamma}_g) \hat{\mathbf{P}}_g^{(r)})$ . After eliminating  $\tau$ , subproblems (41) are simplified to

$$\min_{\zeta_g \in \Omega} (K-1) \log(\text{Tr}(\hat{\mathbf{P}}_g^{(r)}) - \text{Tr}(\mathbf{\Gamma}_g \hat{\mathbf{P}}_g^{(r)})) + \log(\text{Tr}(\mathbf{\Gamma}_g \hat{\mathbf{P}}_g^{(r)})), \forall g, \quad (42)$$

where  $\zeta_g \in \delta_g^{(r)} = \{ \zeta_g \in \Omega \mid \text{Tr}(\mathbf{\Gamma}_g \hat{\mathbf{P}}_g^{(r)}) > \text{Tr}(\hat{\mathbf{P}}_g^{(r)}) / K \}$  due to the fact that when  $\zeta_g^{(r)} \in \delta_g^{(r)}$ ,  $e_g^{(r)} = \text{Tr}(\mathbf{\Gamma}_g \hat{\mathbf{P}}_g^{(r)}) > \text{Tr}(\hat{\mathbf{P}}_g^{(r)}) / K$  and

$$\sigma_g^{(r)} = \max \{ (e_g^{(r)} / \tau_g^{(r)} - 1) / K, 0 \} = (e_g^{(r)} - \text{Tr}(\hat{\mathbf{P}}_g^{(r)}) / K) / d_g^{(r)} > 0, \forall g. \quad (43)$$

Since  $(K-1) \log(\text{Tr}(\hat{\mathbf{P}}_g^{(r)}) - x) + \log(x)$  is a monotonically decreasing function of  $x$  for  $x \geq \text{Tr}(\hat{\mathbf{P}}_g^{(r)}) / K$ , subproblems (42) are equivalent to

$$\max_{\zeta_g \in \delta_g^{(r)}} \text{Tr}(\mathbf{\Gamma}_g \hat{\mathbf{P}}_g^{(r)}), \forall g. \quad (44)$$

Based on the above analysis,  $\Psi^{(r)} = (\zeta^{(r)}, \rho^{(r)})$  and  $\tau^{(r)}$  are obtained by

$$\zeta_g^{(r)} = \arg \max_{\zeta_g \in \Omega} \text{Tr}(\mathbf{\Gamma}_g \hat{\mathbf{P}}_g^{(r)}), \forall g, \quad (45)$$

$$\tau_g^{(r)} = \begin{cases} d_g^{(r)} / (K-1) \geq 0, & e_g^{(r)} > \text{Tr}(\hat{\mathbf{P}}_g^{(r)}) / K, \\ \text{Tr}(\hat{\mathbf{P}}_g^{(r)}) / K \geq 0, & e_g^{(r)} \leq \text{Tr}(\hat{\mathbf{P}}_g^{(r)}) / K, \end{cases} \forall g, \quad (46)$$

$$\rho_g^{(r)} = \max \{ (e_g^{(r)} - \text{Tr}(\hat{\mathbf{P}}_g^{(r)}) / K) / (K-1), 0 \}, \forall g, \quad (47)$$

where we can use a proof by contradiction to verify that  $\tau^{(r)} > \mathbf{0}$  if  $\tau^{(r-1)} > \mathbf{0}$ .

## 5. SAGE Algorithm

In the SAGE algorithm, each iteration consists of  $G$  cycles and  $\zeta_q^{(r)}$  is obtained at the  $q$ th cycle of the  $r$ th iteration. Let  $[\cdot]^{(r,q)}$  mean an iterative value at the  $q$ th cycle of the  $r$ th iteration,  $[\cdot]^{(r-1)} = [\cdot]^{(r-1,G)} = [\cdot]^{(r,0)}$ .

At the  $q$ th cycle of the  $r$ th iteration, the SAGE algorithm first constructs the complete data by [7,8]

$$\mathbf{h}_g(t) = \begin{cases} \mathbf{b}(\tilde{\zeta}_g)m_g(t) + \mathbf{v}(t) & g = q, \\ \mathbf{b}(\tilde{\zeta}_g)m_g(t) & g \neq q. \end{cases} \quad (48)$$

Then, the E- and M-steps at the  $q$ th cycle of the  $r$ th iteration are introduced below.

### 5.1. Deterministic Signal Model

Based on (48), we have  $\mathbf{h}_q(t) \sim \mathcal{CN}(\mathbf{b}(\tilde{\zeta}_q)m_q(t), \tau\mathbf{I}_K)$  and the  $\mathbf{h}_g(t)$ 's with  $g \neq q$  are deterministic. The complete-data LLF is expressed as

$$\begin{aligned} \mathcal{Z}(\tilde{\zeta}_q, \mathbf{m}_q, \tau) &= \sum_{t=1}^L \log p(\mathbf{h}_q(t); \tilde{\zeta}_q, m_q(t), \tau) \\ &= -LK \log(\pi\tau) - \frac{1}{\tau} \sum_{t=1}^L \|\mathbf{h}_q(t) - \mathbf{b}(\tilde{\zeta}_q)m_q(t)\|^2. \end{aligned} \quad (49)$$

#### 5.1.1. E-Step

Calculate the conditional expectation of the complete-data LLF in (49)

$$\mathcal{E}\{\mathcal{Z}(\tilde{\zeta}_q, \mathbf{m}_q, \tau) | \mathbf{W}; \Delta^{(r,q-1)}\} = -LK \log(\pi\tau) - \frac{1}{\tau} \sum_{t=1}^L \|\mathbf{h}_q^{(r)}(t) - \mathbf{b}(\tilde{\zeta}_q)m_q(t)\|^2, \quad (50)$$

where  $\Delta = (\tilde{\zeta}, \mathbf{M}, \tau)$ ,  $\mathbf{0}_K$  is the  $K \times K$  zero matrix, and

$$\begin{aligned} \mathbf{h}_q^{(r)}(t) &= \mathbf{h}_q^{(r,q)}(t) = \mathcal{E}\{\mathbf{h}_q(t) | \mathbf{W}; \Delta^{(r,q-1)}\} \\ &= \mathbf{b}(\tilde{\zeta}_q^{(r,q-1)})m_q^{(r,q-1)}(t) + [\mathbf{w}(t) - \mathbf{B}(\tilde{\zeta}_q^{(r,q-1)})\mathbf{m}^{(r,q-1)}(t)], \end{aligned} \quad (51)$$

$$\mathcal{D}\{\mathbf{h}_q(t) | \mathbf{W}; \Delta^{(r,q-1)}\} = \mathbf{0}_K. \quad (52)$$

#### 5.1.2. M-Step

Update the estimates of  $\tilde{\zeta}_q$ ,  $\mathbf{m}_q$ , and  $\tau$  by solving

$$\min_{\tilde{\zeta}_q \in \Omega, \mathbf{m}_q, \tau > 0} K \log(\tau) + \frac{1}{\tau L} \sum_{t=1}^L \|\mathbf{h}_q^{(r)}(t) - \mathbf{b}(\tilde{\zeta}_q)m_q(t)\|^2. \quad (53)$$

$\tilde{\zeta}_q^{(r)}$ ,  $\mathbf{m}_q^{(r)}$ , and  $\tau^{(r,q)}$  are obtained by

$$\tilde{\zeta}_q^{(r)} = \tilde{\zeta}_q^{(r,q)} = \arg \max_{\tilde{\zeta}_q \in \Omega} \text{Tr}\{\Gamma_q \hat{\mathbf{P}}_q^{(r)}\}, \quad (54)$$

$$m_q^{(r)}(t) = m_q^{(r,q)}(t) = \mathbf{b}^H(\tilde{\zeta}_q^{(r)})\mathbf{h}_q^{(r)}(t)/K, \forall t, \quad (55)$$

$$\tau^{(r,q)} = d_q^{(r)}/K, \quad (56)$$

where  $\hat{\mathbf{P}}_q^{(r)} = (1/L) \sum_{t=1}^L \mathbf{h}_q^{(r)}(t) [\mathbf{h}_q^{(r)}(t)]^H$  and  $d_q^{(r)} = \text{Tr}((\mathbf{I}_K - \Gamma_q^{(r)})\hat{\mathbf{P}}_q^{(r)})$ .

Moreover, the other parameter estimates are not updated at this cycle and their iterative values are

$$\tilde{\zeta}_g^{(r,q)} = \tilde{\zeta}_g^{(r,q-1)}, \forall g \neq q, \quad (57)$$

$$m_g^{(r,q)}(t) = m_g^{(r,q-1)}(t), \forall g \neq q, t. \quad (58)$$

**Remark 2.** Since the  $\mathbf{h}_q^{(r)}(t)$ 's in (51),  $\tilde{\zeta}_q^{(r)}$  in (54), and the  $m_q^{(r)}(t)$ 's in (55) are unrelated to  $\tau^{(r,q-1)}$ , we can omit (56) due to the nuisance parameter  $\tau$ .

### 5.2. Random Signal Model

Based on (48), we have  $\mathbf{h}_q(t) \sim \mathcal{CN}(\mathbf{0}, \mathbf{N}_q)$  with  $\mathbf{N}_q = \rho_q \mathbf{b}(\zeta_q) \mathbf{b}^H(\zeta_q) + \tau \mathbf{I}_K$ . However, the distribution of  $\mathbf{h}_g(t)$  with  $g \neq q$  is only associated with  $m_g(t)$ . The complete-data LLF is written as

$$\begin{aligned} \mathcal{Z}(\zeta_q, \rho, \tau) &= \sum_{t=1}^L \sum_{g \neq q} \log p(m_g(t); \rho_g) + \sum_{t=1}^L \log p(\mathbf{h}_q(t); \zeta_q, \rho_q, \tau) \\ &= -L(G-1) \log(\pi) - L \sum_{g \neq q} [\log(\rho_g) + \hat{P}_g / \rho_g] \\ &\quad -LK \log(\pi) - L[\log(\text{Det}(\mathbf{N}_q)) + \text{Tr}(\mathbf{N}_q^{-1} \hat{\mathbf{P}}_q)], \end{aligned} \quad (59)$$

where  $\hat{P}_g = (1/L) \sum_{t=1}^L |m_g(t)|^2$  and  $|\cdot|$  denotes modulus.

#### 5.2.1. E-Step

Calculate the conditional expectation of the complete-data LLF in (59)

$$\begin{aligned} \mathcal{E}\{\mathcal{Z}(\zeta_q, \rho, \tau) | \mathbf{W}; \Delta^{(r,q-1)}\} &= V - L \sum_{g \neq q} [\log(\rho_g) + \hat{P}_g^{(r,q)} / \rho_g] \\ &\quad -L[\log(\text{Det}(\mathbf{N}_q)) + \text{Tr}(\mathbf{N}_q^{-1} \hat{\mathbf{P}}_q^{(r)})], \end{aligned} \quad (60)$$

where  $\Delta = (\zeta, \rho, \tau)$  and  $V = -L(K+G-1) \log(\pi)$ . In (60),

$$\begin{aligned} \hat{P}_g^{(r,q)} &= \mathcal{E}\{\hat{P}_g | \mathbf{W}; \Delta^{(r,q-1)}\} \\ &= \rho_g^{(r,q-1)} [1 - \mathbf{b}^H(\zeta_g^{(r,q-1)}) \mathbf{d}_g^{(r,q-1)}] + [\mathbf{d}_g^{(r,q-1)}]^H \hat{\mathbf{P}}_w \mathbf{d}_g^{(r,q-1)} \geq 0 \end{aligned} \quad (61)$$

with  $\mathbf{d}_g^{(r,q-1)} = [\mathbf{N}_w^{(r,q-1)}]^{-1} \mathbf{b}(\zeta_g^{(r,q-1)}) \rho_g^{(r,q-1)}$  and

$$\begin{aligned} \hat{\mathbf{P}}_q^{(r)} &= \hat{\mathbf{P}}_q^{(r,q)} = \mathcal{E}\{\hat{\mathbf{P}}_q | \mathbf{W}; \Delta^{(r,q-1)}\} \\ &= \mathbf{N}_q^{(r,q-1)} - \mathbf{N}_q^{(r,q-1)} [\mathbf{N}_w^{(r,q-1)}]^{-1} \mathbf{N}_q^{(r,q-1)} \\ &\quad + \mathbf{N}_q^{(r,q-1)} [\mathbf{N}_w^{(r,q-1)}]^{-1} \hat{\mathbf{P}}_w [\mathbf{N}_w^{(r,q-1)}]^{-1} \mathbf{N}_q^{(r,q-1)}. \end{aligned} \quad (62)$$

#### 5.2.2. M-Step

Update the estimates of  $\zeta_q$ ,  $\rho$ , and  $\tau$  by solving

$$\min_{\zeta_q \in \Omega, \rho \geq 0, \tau > 0} \sum_{g \neq q} [\log(\rho_g) + \hat{P}_g^{(r,q)} / \rho_g] + \log(\text{Det}(\mathbf{N}_q)) + \text{Tr}(\mathbf{N}_q^{-1} \hat{\mathbf{P}}_q^{(r)}). \quad (63)$$

We, thus, have  $\rho_g^{(r,q)} = \hat{P}_g^{(r,q)}$ ,  $\forall g \neq q$ , and  $\zeta_q^{(r)}$ ,  $\rho_q^{(r,q)}$ , and  $\tau^{(r,q)}$  are obtained by solving

$$\min_{\zeta_q \in \Omega, \sigma_q \geq 0, \tau > 0} K \log(\tau) + \log(\text{Det}(\mathbf{Q}_q)) + \frac{1}{\tau} \text{Tr}(\mathbf{Q}_q^{-1} \hat{\mathbf{P}}_q^{(r)}), \quad (64)$$

where  $\mathbf{N}_q = \tau \mathbf{Q}_q$  with  $\mathbf{Q}_q = \sigma_q \mathbf{b}(\zeta_q) \mathbf{b}^H(\zeta_q) + \mathbf{I}_K$  and  $\sigma_q = \rho_q / \tau$ . Following (45)–(47),  $\zeta_q^{(r)}$ ,  $\rho_q^{(r,q)}$ , and  $\tau^{(r,q)}$  are obtained by

$$\zeta_q^{(r)} = \zeta_q^{(r,q)} = \arg \max_{\zeta_q \in \Omega} \text{Tr}(\Gamma_q \hat{\mathbf{P}}_q^{(r)}), \quad (65)$$

$$\tau^{(r,q)} = \begin{cases} d_q^{(r)} / (K-1) \geq 0, & e_q^{(r)} > \text{Tr}(\hat{\mathbf{P}}_q^{(r)}) / K, \\ \text{Tr}(\hat{\mathbf{P}}_q^{(r)}) / K \geq 0, & e_q^{(r)} \leq \text{Tr}(\hat{\mathbf{P}}_q^{(r)}) / K, \end{cases} \quad (66)$$

$$\rho_q^{(r,q)} = \max\{(e_q^{(r)} - \text{Tr}(\hat{\mathbf{P}}_q^{(r)}) / K) / (K-1), 0\}, \quad (67)$$

where  $\tau^{(r,q)} = 0$  is possible although its probability is very low. For example, if  $G = 2$ ,  $L = 1$ , and  $\rho_1^{(1,1)} = 0$  at 1st cycle of the 1st iteration, at the 2nd cycle of the 1st iteration we will have

$$\mathbf{N}_w^{(1,1)} = \rho_2^{(1,1)} \mathbf{b}(\zeta_2^{(1,1)}) \mathbf{b}^H(\zeta_2^{(1,1)}) + \tau^{(1,1)} \mathbf{I}_K = \mathbf{N}_2^{(1,1)} \quad (68)$$

and then  $\hat{\mathbf{P}}_2^{(1)} = \mathbf{w}(1) \mathbf{w}^H(1)$  by (62). Furthermore, if  $\mathbf{w}(1) = h \mathbf{b}(\bar{\zeta})$  ( $h \neq 0, \bar{\zeta} \in \Omega$ ), we will obtain  $\zeta_2^{(1)} = \bar{\zeta}$  by (65) and  $\tau^{(1,2)} = d_2^{(1)} / (K - 1) = 0$  by (66).

To avoid  $\tau^{(r,q)} = 0$ , we use the following two CM-steps to re-estimate  $\zeta_q, \rho_q$ , and  $\tau$  based on problem (64) if  $\tau^{(r,q)} = 0$  in (66).

- *First CM-step:* Estimate  $\zeta_q$  and  $\rho_q$  but hold  $\tau = \tau^{(r,q-1)}$  fixed. Then, problem (64) is simplified to

$$\min_{\zeta_q \in \Omega, \sigma_q \geq 0} \log(\text{Det}(\mathbf{Q}_q)) + \frac{1}{\tau^{(r,q-1)}} \text{Tr}(\mathbf{Q}_q^{-1} \hat{\mathbf{P}}_q^{(r)}). \quad (69)$$

Following (21) and (22),  $\zeta_q^{(r)}$  and  $\rho_q^{(r,q)}$  are obtained by

$$\zeta_q^{(r)} = \zeta_q^{(r,q)} = \arg \max_{\zeta_q \in \Omega} \text{Tr}(\mathbf{\Gamma}_q \hat{\mathbf{P}}_q^{(r)}), \quad (70)$$

$$\rho_q^{(r,q)} = \sigma_q^{(r,q)} \tau^{(r,q-1)} = \max \{ (e_q^{(r)} - \tau^{(r,q-1)}) / K, 0 \}. \quad (71)$$

- *Second CM-step:* Estimate  $\tau$  but hold  $\zeta_q = \zeta_q^{(r)}$  and  $\sigma_q = \sigma_q^{(r,q)}$  fixed. Then, problem (64) is simplified to

$$\min_{\tau > 0} K \log(\tau) + \frac{1}{\tau} \text{Tr}([\mathbf{Q}_q^{(r)}]^{-1} \hat{\mathbf{P}}_q^{(r)}), \quad (72)$$

where  $\mathbf{Q}_q^{(r)} = \sigma_q^{(r,q)} \mathbf{b}(\zeta_q^{(r)}) \mathbf{b}^H(\zeta_q^{(r)}) + \mathbf{I}_K$ . We obtain  $\tau^{(r,q)}$  by

$$\tau^{(r,q)} = \text{Tr} \{ [\mathbf{Q}_q^{(r)}]^{-1} \hat{\mathbf{P}}_q^{(r)} \} / K = (\tau^{(r,q-1)} + d_q^{(r)}) / K, \quad (73)$$

where  $d_q^{(r)} = \text{Tr}((\mathbf{I}_K - \mathbf{\Gamma}_q^{(r)}) \hat{\mathbf{P}}_q^{(r)}) \geq 0$  and  $\tau^{(r,q)} > 0$  if  $\tau^{(r,q-1)} > 0$ .

Moreover, the other parameter estimate(s) is (are) not updated at this cycle and the iterative value(s) is (are)

$$\zeta_g^{(r,q)} = \zeta_g^{(r,q-1)}, \forall g \neq q. \quad (74)$$

## 6. Properties of the Proposed EM, MEM, and SAGE Algorithms

### 6.1. Convergence Point

It is easy to verify that the above EM-type algorithms satisfy standard regularity conditions [5,21,26] and always converge to stationary points of  $\mathcal{J}(\Psi, \tau)$ . Of course, the convergence points of these EM-type algorithms depend on their initial points. To generate appropriate initial points, we can employ the method presented in [10] using the deterministic signal model.

### 6.2. Complexity and Stability

At the  $r$ th iteration, the computational burdens of the above EM-type algorithms lie in solving the  $G$  maximization problems

$$\zeta_g^{(r)} = \arg \max_{\zeta_g \in \Omega} \text{Tr}(\mathbf{\Gamma}_g \hat{\mathbf{P}}_g^{(r)}), \forall g. \quad (75)$$

Thus, these EM-type algorithms have almost the same computational complexity at every iteration and if an algorithm has faster convergence, its number of iterations required will be smaller and its computations will be fewer.

However, when the powers of sources are unequal, we have found via simulation that the DOA estimates of multiple sources, updated by (75), tend to be consistent with the true DOA of the source with the largest power. Accordingly, these EM-type algorithms may be unstable. To address this issue, we can reduce the difference between  $\zeta_g^{(r)}$  and  $\zeta_g^{(r-1)}$  by increasing  $\text{Tr}(\mathbf{\Gamma}_g \hat{\mathbf{P}}_g^{(r)})$  rather than maximizing it, i.e.,

$$\text{Tr}(\mathbf{\Gamma}_g^{(r)} \hat{\mathbf{P}}_g^{(r)}) \geq \text{Tr}(\mathbf{\Gamma}_g^{(r-1)} \hat{\mathbf{P}}_g^{(r)}), \forall g, \quad (76)$$

which guarantees the monotonicity of these EM-type algorithms [3]. As a good example, Algorithm 1 in the next section has given excellent simulation results.

---

#### Algorithm 1 Gradient ascent with backtracking line search

---

- 1:  $y(\eta_g) = K \times \text{Tr}(\mathbf{\Gamma}_g \hat{\mathbf{P}}_g^{(r)})$ ,  $\eta_g = \eta_g^{(r-1)} \in (0, \pi)$  (radian).
  - 2: **while**  $|y'(\eta_g)| > 0.001$  **do**
  - 3:  $s = 0.1 \times \begin{cases} (\pi - \eta_g)/y'(\eta_g), & y'(\eta_g) > 0, \\ -\eta_g/y'(\eta_g), & y'(\eta_g) < 0. \end{cases}$
  - 4: **while**  $y(\eta_g + sy'(\eta_g)) < y(\eta_g) + 0.3s(y'(\eta_g))^2$  **do**
  - 5:  $s = 0.5s$ .
  - 6: **end while**
  - 7:  $\eta_g = \eta_g + sy'(\eta_g) \in (0, \pi)$  (radian).
  - 8: **end while**
  - 9:  $\eta_g^{(r)} = \eta_g$ .
- 

## 7. Simulation Results

We give simulation results to compare the proposed EM-type algorithms. For convenience, the array is assumed to be a uniformly spaced linear array, in which  $\zeta_k = [\frac{\lambda}{2}(k-1) \ 0 \ 0]^T$ . We set  $G = 2$ ,  $L = 20$ , and  $\beta = \mathbf{1}/G$  in the EM algorithm. Here,  $\mu_1 = \mu_2 = 90^\circ$  is known while  $\eta_1$  and  $\eta_2$  need to be estimated.  $\mathbf{M}$  in the deterministic signal model is also generated by the independent random numbers  $m_g(t) \sim \mathcal{CN}(0, \rho_g)$ . All the algorithms adopt the stopping criterion  $\|\zeta^{(r+1)} - \zeta^{(r)}\| \leq 0.001^\circ$ . Algorithm 1 is designed to search the  $\eta_g^{(r)}$ 's in (76) [27]. Moreover,  $\mathbf{M}^{(0)} = [\mathbf{1} \ \mathbf{1}]^T$ ,  $\rho^{(0)} = \mathbf{1}$ ,  $\tau^{(0)} = \mathbf{1}/G$ , and  $\tau^{(0)} = 1$ . In this section, the EM, MEM, and SAGE algorithms are simply written as EM, MEM, and SAGE, respectively.

### 7.1. Deterministic Signal Model

To compare the convergence of EM, MEM, and SAGE, Figure 1 plots their  $\mathcal{J}(\Psi^{(r)}, \tau^{(r)})$ 's,  $\eta_1^{(r)}$ 's, and  $\eta_2^{(r)}$ 's under one trial. In Figure 1,  $K = 10$ ,  $\eta_1 = 20^\circ$ ,  $\eta_2 = 80^\circ$ ,  $\rho_1 = -2$  dB,  $\rho_2 = 4$  dB,  $\tau = 4$  dB,  $\eta_1^{(0)} = 24^\circ$ , and  $\eta_2^{(0)} = 84^\circ$ . It is easy to see that EM, MEM, and SAGE converge to a consistent  $(\eta_1, \eta_2)$  estimate given an accurate initial point. Moreover, EM has a similar convergence with MEM while SAGE converges faster than EM and MEM.

Figures 2 and 3 show two scatter plots of  $(\eta_1, \eta_2)$  estimates under 200 trials. In Figure 2,  $K = 10$ ,  $\eta_1 = 25^\circ$ ,  $\eta_2 = 75^\circ$ ,  $\rho_1 = -4$  dB,  $\rho_2 = 2$  dB,  $\tau = 4$  dB,  $\eta_1^{(0)} = 40^\circ$ , and  $\eta_2^{(0)} = 60^\circ$ . Moreover, the numbers of desirable points obtained by EM, MEM, and SAGE are 68, 72, and 179, respectively. In Figure 3,  $K = 10$ ,  $\eta_1 = 70^\circ$ ,  $\eta_2 = 78^\circ$ ,  $\rho_1 = -2$  dB,  $\rho_2 = 4$  dB,  $\tau = 4$  dB,  $\eta_1^{(0)} = 50^\circ$ , and  $\eta_2^{(0)} = 58^\circ$ . Moreover, the numbers of desirable points obtained by EM, MEM, and SAGE are 159, 157, and 190, respectively. Figures 2 and 3 imply that given a poor initial point, SAGE can avoid the convergence to an undesirable stationary point of  $\mathcal{J}(\Psi, \tau)$  more efficiently than EM and MEM.

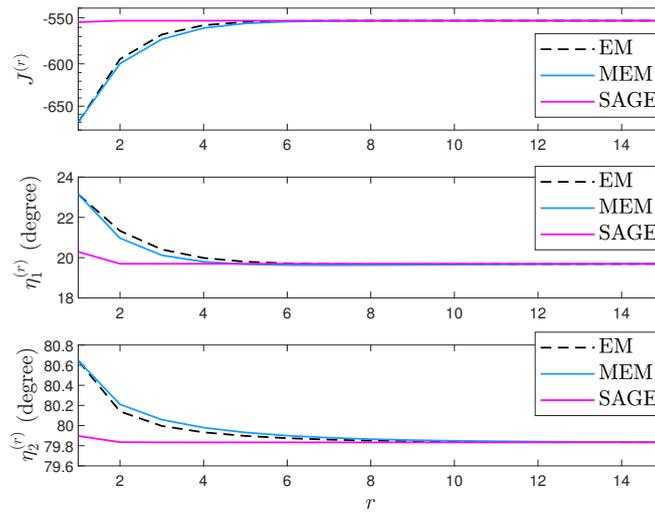


Figure 1.  $\mathcal{J}(\Psi^{(r)}, \tau^{(r)})$ ,  $\eta_1^{(r)}$ , and  $\eta_2^{(r)}$  comparison.

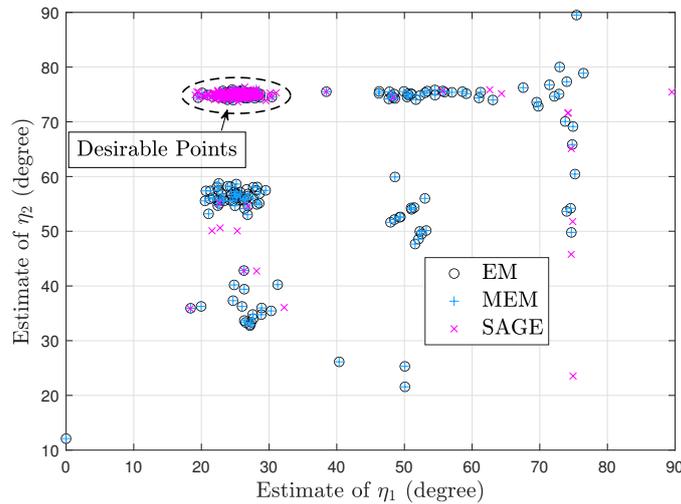


Figure 2. Scatter plot of  $(\eta_1, \eta_2)$  estimates.

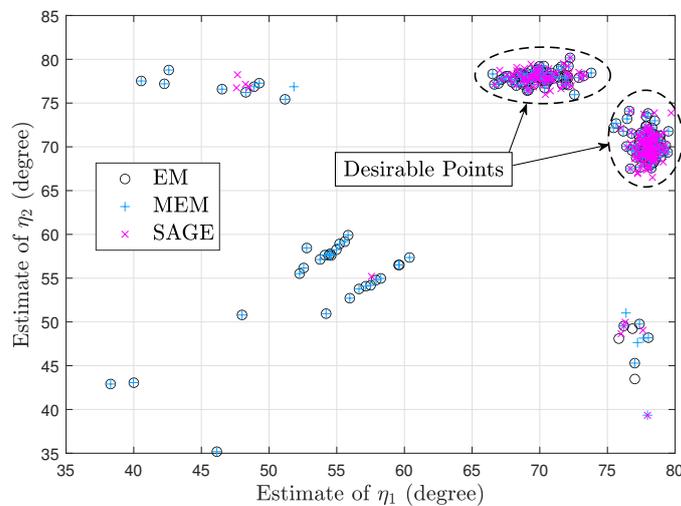


Figure 3. Scatter plot of  $(\eta_1, \eta_2)$  estimates.

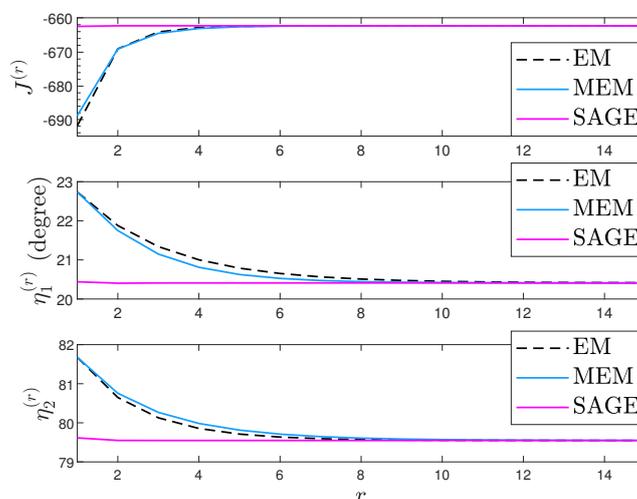
Note that in each of Figures 2 and 3, the SAGE algorithm requires the least processing time due to the fastest convergence and thus performs the fewest computations required.

Moreover, both sources in Figure 2 are not closely spaced, so it is very difficult to mix up both sources and the desirable points in Figure 2 are centered around the true position ( $25^\circ, 75^\circ$ ). However, both sources in Figure 3 are closely spaced and the desirable points are centered around ( $78^\circ, 70^\circ$ ) or ( $70^\circ, 78^\circ$ ), i.e., these EM-type algorithms are likely to mix up closely spaced sources.

According to these simulations, we can conclude that (1) EM has similar convergence with MEM, and (2) SAGE outperforms EM and MEM.

## 7.2. Random Signal Model

To compare the convergence of EM, MEM, and SAGE, Figure 4 plots their  $\mathcal{J}(\Psi^{(r)}, \tau^{(r)})$ 's,  $\eta_1^{(r)}$ 's, and  $\eta_2^{(r)}$ 's under one trial. In Figure 4,  $K = 10$ ,  $\eta_1 = 20^\circ$ ,  $\eta_2 = 80^\circ$ ,  $\rho_1 = -4$  dB,  $\rho_2 = 4$  dB,  $\tau = 4$  dB,  $\eta_1^{(0)} = 24^\circ$ , and  $\eta_2^{(0)} = 84^\circ$ . We can also observe that given an accurate initial point, EM, MEM, and SAGE converge to a consistent  $(\eta_1, \eta_2)$  estimate. Moreover, EM has similar convergence with MEM while SAGE converges faster than EM and MEM.



**Figure 4.**  $\mathcal{J}(\Psi^{(r)}, \tau^{(r)})$ ,  $\eta_1^{(r)}$ , and  $\eta_2^{(r)}$  comparison.

Figures 5 and 6 show two scatter plots of  $(\eta_1, \eta_2)$  estimates under 200 trials. In Figure 5,  $K = 10$ ,  $\eta_1 = 25^\circ$ ,  $\eta_2 = 75^\circ$ ,  $\rho_1 = -4$  dB,  $\rho_2 = 2$  dB,  $\tau = 4$  dB,  $\eta_1^{(0)} = 40^\circ$ , and  $\eta_2^{(0)} = 60^\circ$ . Moreover, the numbers of desirable points obtained by EM, MEM, and SAGE are 185, 186, and 175, respectively. In Figure 6,  $K = 10$ ,  $\eta_1 = 70^\circ$ ,  $\eta_2 = 78^\circ$ ,  $\rho_1 = -2$  dB,  $\rho_2 = -1$  dB,  $\tau = 4$  dB,  $\eta_1^{(0)} = 55^\circ$ , and  $\eta_2^{(0)} = 63^\circ$ . Moreover, the numbers of desirable points obtained by EM, MEM, and SAGE are 161, 161, and 172, respectively. Figures 5 and 6 imply that EM has similar convergence with MEM but compared to EM and MEM, SAGE is less and more efficient for avoiding the convergence to an undesirable stationary point of  $\mathcal{J}(\Psi, \tau)$  in Figures 5 and 6, respectively. Note that in each of Figures 5 and 6, the SAGE algorithm requires the least processing time due to the fastest convergence and thus performs the fewest computations required. In addition, these EM-type algorithms are likely to mix up closely spaced sources, so the desirable points in Figure 5 are centered around ( $25^\circ, 75^\circ$ ) and the desirable points in Figure 6 are centered around ( $78^\circ, 70^\circ$ ) or ( $70^\circ, 78^\circ$ ).

Based on the above figures, we can conclude that (1) EM has similar convergence with MEM, and (2) SAGE cannot always outperform EM and MEM.

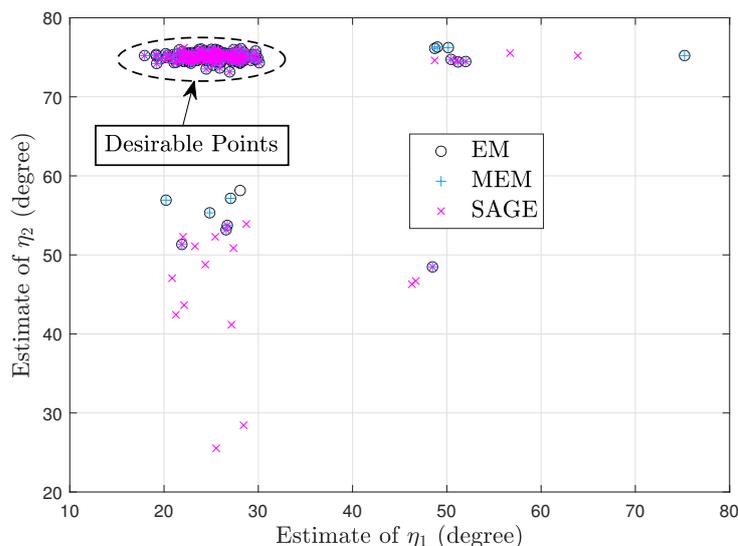


Figure 5. Scatter plot of  $(\eta_1, \eta_2)$  estimates.

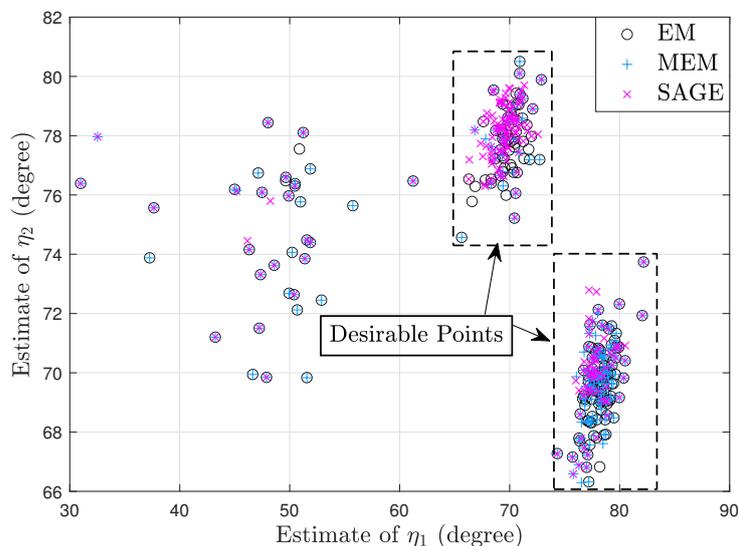


Figure 6. Scatter plot of  $(\eta_1, \eta_2)$  estimates.

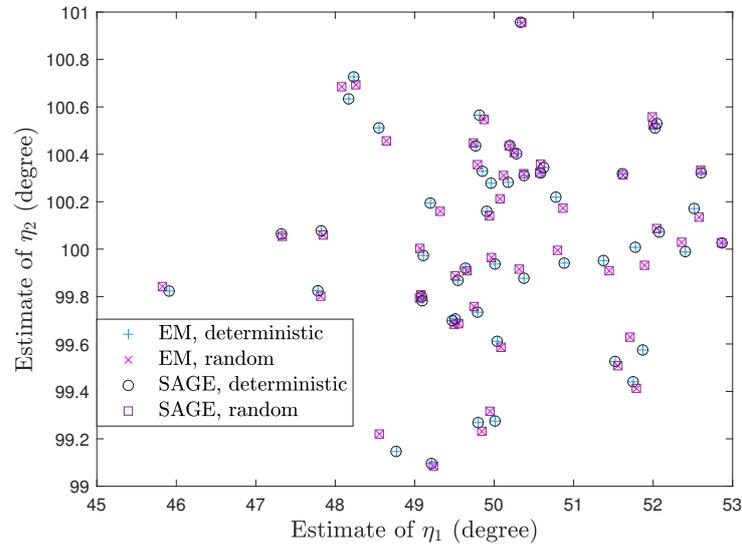
### 7.3. Deterministic and Random Signal Models

Snapshots from the random signal model can be processed by these EM-type algorithms for the deterministic signal model, which means that we can compare these algorithms for both signal models. The above simulation results have shown that EM has similar convergence with MEM, so we only compare EM and SAGE for both signal models in this subsection for simplicity.

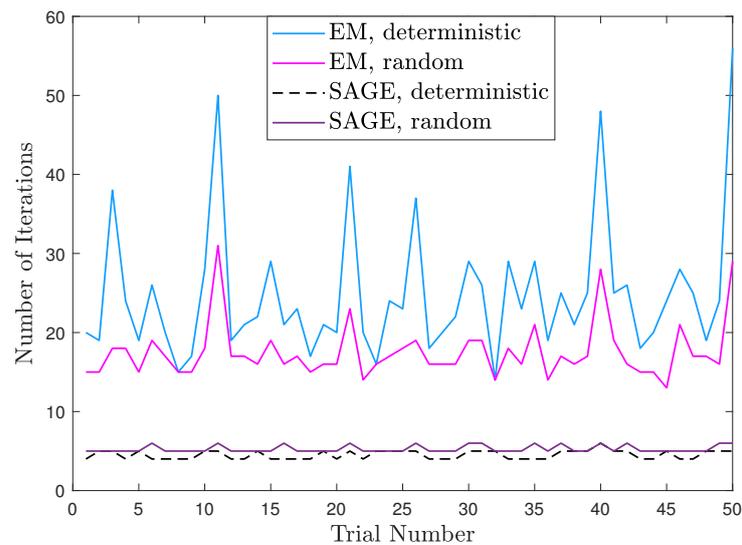
Since both signal models have the same DOA parameter  $\zeta$ , the stopping criterion  $\|\zeta^{(r+1)} - \zeta^{(r)}\| \leq 0.001^\circ$  is suitable. Figure 7 shows a scatter plot of  $(\eta_1, \eta_2)$  estimates under 50 trials. In Figure 7,  $K = 10$ ,  $\eta_1 = 50^\circ$ ,  $\eta_2 = 100^\circ$ ,  $\rho_1 = -4$  dB,  $\rho_2 = 4$  dB,  $\tau = 4$  dB,  $\eta_1^{(0)} = 55^\circ$ , and  $\eta_2^{(0)} = 95^\circ$ . We can see that EM and SAGE for both signal models yield close  $(\eta_1, \eta_2)$  estimates for each trial.

Based on Figure 7, Figure 8 compares the numbers of iterations required by these algorithms. We can observe that EM for the deterministic signal model generally needs more iterations than EM for the random signal model. The reason is that EM for deterministic signal model needs to update more parameter estimates at each iteration and, thus, has slower convergence than EM for the random signal model. Moreover, SAGE for the deterministic signal model generally needs fewer iterations than SAGE for the random

signal model. More importantly, SAGE for the deterministic signal model always requires the fewest iterations for each trial. Thus, we can conclude that SAGE for the deterministic signal model is superior to the other algorithms in the computational cost.



**Figure 7.** Scatter plot of  $(\eta_1, \eta_2)$  estimates.



**Figure 8.** Numbers of iterations of different algorithms.

Figure 9 shows the root mean square error (RMSE) performances of DOA estimation obtained by the SAGE algorithm for the deterministic and random signal models. In Figure 9,  $\eta_1 = 60^\circ$ ,  $\eta_2 = 120^\circ$ ,  $\rho_1 = 0$  dB,  $\rho_2 = 1$  dB,  $\tau = 3$  dB,  $\eta_1^{(0)} = 55^\circ$ , and  $\eta_2^{(0)} = 115^\circ$ . Each RMSE is computed from 1000 independent realizations. We can observe that as the number of sensors  $K$  increases, the SAGE algorithm for each signal model yields small RMSEs, which indicates that increasing the number of sensors  $K$  can improve the accuracy of DOA estimation.

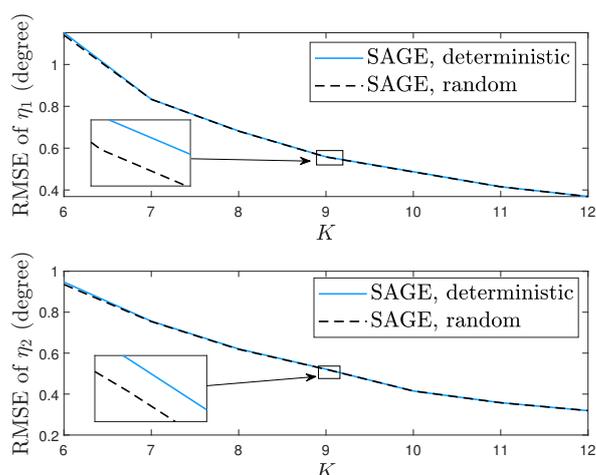


Figure 9. RMSEs of the SAGE algorithm.

## 8. Conclusions

In this paper, we applied and designed the EM and SAGE algorithms for DOA estimation in unknown uniform noise and proposed a new MEM algorithm applicable to the noise assumption. Next, we improved these EM-type algorithms to ensure the stability when the powers of sources are unequal. After being improved, simulation results illustrated that the EM algorithm has similar convergence with the MEM algorithm, the SAGE algorithm outperforms the EM and MEM algorithms for the deterministic signal model, and the SAGE algorithm cannot always outperform the EM and MEM algorithms for the random signal model. In addition, simulation results indicated that when these EM-type algorithms process the same snapshots from the random signal model, the SAGE algorithm for the deterministic signal model can require the fewest iterations and computations.

**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 conflicts of interest.

## References

1. Krim, H.; Viberg, M. Two decades of array signal processing research: The parametric approach. *IEEE Signal Process. Mag.* **1996**, *13*, 67–94. [\[CrossRef\]](#)
2. Godara, L.C. Application of antenna arrays to mobile communications. II. Beam-forming and direction-of-arrival considerations. *Proc. IEEE* **1997**, *85*, 1195–1245. [\[CrossRef\]](#)
3. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Ser. B (Methodol.)* **1977**, *39*, 1–38.
4. Meng, X.; Dyk, D.V. The EM algorithm—an old folk-song sung to a fast new tune. *J. R. Stat. Soc. Ser. B (Methodol.)* **1997**, *59*, 511–567. [\[CrossRef\]](#)

5. Feder, M.; Weinstein, E. Parameter estimation of superimposed signals using the EM algorithm. *IEEE Trans. Acoust. Speech Signal Process.* **1988**, *36*, 477–489. [[CrossRef](#)]
6. 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](#)]
7. Fessler, J.A.; Hero, A.O. Space-alternating generalized expectation-maximization algorithm. *IEEE Trans. Signal Process.* **1994**, *42*, 2664–2677. [[CrossRef](#)]
8. 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](#)]
9. Gong, M.-Y.; Lyu, B. Alternating maximization and the EM algorithm in maximum-likelihood direction finding. *IEEE Trans. Veh. Technol.* **2021**, *70*, 9634–9645. [[CrossRef](#)]
10. Ziskind, I.; Wax, M. Maximum likelihood localization of multiple sources by alternating projection. *IEEE Trans. Acoust. Speech Signal Process.* **1988**, *36*, 1553–1560. [[CrossRef](#)]
11. Stoica, P.; Nehorai, A. MUSIC, maximum likelihood, and Cramer-Rao bound. *IEEE Trans. Acoust. Speech Signal Process.* **1989**, *37*, 720–741. [[CrossRef](#)]
12. 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](#)]
13. 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](#)]
14. 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](#)]
15. 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](#)]
16. Esfandiari, M.; Vorobyov, S.A. A novel angular estimation method in the presence of nonuniform noise. In Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Singapore, 22–27 May 2022.
17. Yang, J.; Yang, Y.; Lu, J.; Yang, L. Iterative methods for DOA estimation of correlated sources in spatially colored noise fields. *Signal Process.* **2021**, *185*, 108100. [[CrossRef](#)]
18. Selva, J. Efficient computation of ML DOA estimates under unknown nonuniform sensor noise powers. *Signal Process.* **2023**, *205*, 108879. [[CrossRef](#)]
19. Pesavento, M.; Gershman, A.B. Maximum-likelihood direction-of-arrival estimation in the presence of unknown nonuniform noise. *IEEE Trans. Signal Process.* **2001**, *49*, 1310–1324. [[CrossRef](#)]
20. 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](#)]
21. Meng, M.; Rubin, D.B. Maximum likelihood estimation via the ECM algorithm: A general framework. *Biometrika* **1993**, *80*, 267–278. [[CrossRef](#)]
22. Chung, P.; Bohme, J.F. DOA estimation using fast EM and SAGE algorithms. *Signal Process.* **2002**, *82*, 1753–1762. [[CrossRef](#)]
23. Rhodes, I.B. A tutorial introduction to estimation and filtering. *IEEE Trans. Autom. Control* **1971**, *16*, 688–706. [[CrossRef](#)]
24. Jaffer, A.G. Maximum likelihood direction finding of stochastic sources: A separable solution. In Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP), New York, NY, USA, 11–14 April 1988.
25. Stoica, P.; Nehorai, A. On the concentrated stochastic likelihood function in array signal processing. *Circuits Syst. Signal Process.* **1995**, *14*, 669–674. [[CrossRef](#)]
26. Jeff Wu, C.F. On the convergence properties of the EM algorithm. *Ann. Stat.* **1983**, *11*, 95–103.
27. Boyd, S.; Vandenberghe, L. *Convex Optimization*, 1st ed.; Cambridge University Press: New York, NY, USA, 2004; pp. 463–466.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.