

Research Article

Covariance Matrix Reconstruction for Direction Finding with Nested Arrays Using Iterative Reweighted Nuclear Norm Minimization

Weijie Tan  and Xi'an Feng

School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an 710072, China

Correspondence should be addressed to Weijie Tan; tanweijie@hotmail.com

Received 30 July 2018; Accepted 10 December 2018; Published 18 March 2019

Guest Editor: Raed A. Abd-Alhameed

Copyright © 2019 Weijie Tan and Xi'an Feng. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, we address the direction finding problem in the background of unknown nonuniform noise with nested array. A novel gridless direction finding method is proposed via the low-rank covariance matrix approximation, which is based on a reweighted nuclear norm optimization. In the proposed method, we first eliminate the noise variance variable by linear transform and utilize the covariance fitting criteria to determine the regularization parameter for insuring robustness. And then we reconstruct the low-rank covariance matrix by iteratively reweighted nuclear norm optimization that imposes the nonconvex penalty. Finally, we exploit the search-free DoA estimation method to perform the parameter estimation. Numerical simulations are carried out to verify the effectiveness of the proposed method. Moreover, results indicate that the proposed method has more accurate DoA estimation in the nonuniform noise and off-grid cases compared with the state-of-the-art DoA estimation algorithm.

1. Introduction

Source localization is always a significant research direction in the past decades and today which is widely used in various fields including radar, sonar, wireless communication, and acoustics [1–4]. However, achieving superresolution is a great challenge for more sources with the least elements. Compared with the uniform linear arrays (ULAs), the sparse linear arrays (SLAs) can provide much more degree of freedoms (DOFs) and resolve more sources with the same number of physical elements. Meanwhile, the SLAs can efficiently reduce the cost and power consumption. Therefore, the study of SLAs has attracted more and more researchers' great attention [5, 6].

So far, there are some works about SLA configurations. The minimum redundancy arrays (MRAs) are class optimum lag SLAs [5], which provide a complete set of spatial lags between pairs of elements with minimum redundancies. However, it is not easy to attain the optimum design of

MRAs, especially for a large array; it is time-consuming to search the optimum structure. In the past decades, two potential array configurations, i.e., coprime arrays (CPA) and nested arrays (NA), have drawn the researchers' attentions since they have exactly closed-form expressions for sensor locations and it is easy to predict the attainable DOFs [7, 8]. Coprime arrays can resolve up to $o(MN)$ sources with only $M + N - 1$ elements; however, the CPAs do not generate a filled coarray since CPA has holes in the difference coarray [7]. So it cannot directly apply the augmentation techniques. By contrast, nested arrays can generate a filled difference coarray except the mentioned advantages [8]. Although the NAs and CPAs are not optimum lag arrays like MRAs, they are the most attractive array configuration because they are easy to build in the last decade.

The processing of SLA mainly has array interpolation (AI) [9–11] and the Toeplitz completion method [12–14]. The classical AI technique [10] can effectively obtain the element data of virtual ULAs, which imposes a linear

interpolation process on the element data of a real SLA and selects interpolator coefficients by minimizing the interpolation error for a source coming from a certain angular sector. The drawback is that this technique needs to know the angular sector. In AI techniques, Wiener array interpolation (WAI) [15] is a practically attractive method, which exploits a maximum likelihood method to estimate the power of signal and noise and use the calibration angles to recover the array steering matrix; hence, it can approximate the MSE optimum solution. However, WAI requires the initial DoA estimation. In the classical Toeplitz completion method, the direct augmentation approach (DAA) is a widely used method for improving the covariance matrix of the SLA [12]. DAA constructs a Toeplitz matrix using the sample covariance matrix; since there is one-to-one correspondence between covariance lags and spatial lags, the diagonal elements could be obtained by redundant averaging. Unfortunately, Toeplitz completion method does not guarantee positive semidefinite augmented covariance matrix. In order to construct a positive definite augmented matrix, the authors in [13, 14] proposed an iterative DAA algorithm; however, the complicated iteration procedure cannot guarantee the global convergence. Besides, the coarray MUSIC algorithm is used to estimated angle parameters [16] which utilizes the Toeplitz properties of matrices.

Sparse signal representation (SSR) framework has attracted a great interest in direction finding [17–23], which exploits the spatial sparsity of the signal arriving angles. Despite these aforementioned SSR-based methods have some attractive features, there are still two problems need to be considered. The first problem is nonuniform noise [23, 24]; many of SSR-based methods assume that the element noises are spatially uniform, and the reason for this is the uniform noise assumption benefits to choose an appropriate regularization parameter in a sparse representation model. However, in some practical applications, due to the nonidealistic of the practical arrays, such as the nonideality of the receiving channel, the nonuniformity of the element response, and the mutual coupling between elements, the uniform white noise assumption among all elements may not be satisfied. Therefore, the diagonal elements in noise covariance matrix should be considered as arbitrary values which represent the noise levels. Once the assumption misfits the true noise levels, the performance of conventional SSR-based DoA estimation approaches may thereby degrade severely. The second problem is the off-grid problem (also called basis mismatch) [17, 25]; in fact, a great deal of SSR-based DoA estimation methods assume that the true target directions exactly lie on the prespecified angular grid. In practical application, the angle parameters are continuous variables, the on-grid assumption rarely holds, and the basis mismatch always exists, which leads to the accuracy degradation of DoA estimation.

In this paper, we propose a gridless direction finding method under the unknown nonuniform noise over with nested array, which is based on low-rank covariance matrix approximation by an iteratively reweighted nuclear norm minimization. The advantages of the proposed method are in three aspects: Firstly, we mitigate the noise variance variable by linear transform to reduce the effect of

unknown nonuniform noise. Secondly, we utilize the covariance fitting criteria to determine the regularization parameter for the accuracy error fitting. Last but not least, we develop a novel reweighted nonconvex penalty objective function for exactly approximating the low-rank covariance matrix. After attaining the covariance matrix, we exploit the search-free DoA method to estimate parameters. Numerical simulation is carried out to verify the effectiveness of the proposed method.

The rest of this paper is organized as follows: In Section 2, we present the signal model with nested array. In Section 3, the DoA estimation with iteratively covariance matrix reconstruction (ICMR) is introduced to the nested array, which is based on an iteratively reweighted nuclear norm minimization algorithm. Section 4 validates the proposed method by numerical simulations and comparing the algorithm with state-of-the-art DoA estimation methods based on nested array. Section 5 concludes this paper.

Notation. We define a vector \mathbf{x} with a boldface lowercase and a matrix \mathbf{X} with a boldface uppercase. The symbol $(\cdot)^*$ denotes the complex conjugate of vector \mathbf{x} or matrix \mathbf{X} , and the transpose and the Hermitian transpose of vector or matrix are expressed by $(\cdot)^T$ and $(\cdot)^H$, respectively. $\mathbb{E}\{\cdot\}$ denote the statistical expectation, $\text{vec}(\cdot)$ denotes stacking element-wise in a column, $\text{rank}(\cdot)$ and $\|\cdot\|_*$ denote the rank and nuclear norm of the matrix, respectively, \otimes denotes the Kronecker product, \odot denotes the element-wise multiplication between two matrices with the same dimensional, $\text{diag}(\cdot)$ denotes the diagonal operator that takes a vector to a matrix with vector on the diagonal, sign is signum function, $\sigma(\mathbf{X})$ denote the eigenvalue of matrix \mathbf{X} , $\sigma_i(\mathbf{X})$ denote the i -th largest eigenvalue value, \mathbf{I}_k denotes a $K \times K$ dimensional identity matrix, $\|\cdot\|_2$ denotes the l_2 norm, and $\text{Pr}(\cdot)$ denotes probability density function.

2. Array Model for Nested Array

We consider a two-level nested array consisting of two ULA subarrays, where the first ULA subarray has N_1 elements with the interelement spacing $d_1 = d$, while the second ULA subarray has N_2 elements with the interelement spacing $(N_1 + 1)d_1$; the total number of element is $M = N_1 + N_2$, as shown in Figure 1.

Let the first element of the array as the phase reference point, the position set of element in ascending is denoted as

$$\mathcal{P} = \{p_1, \dots, p_M\}. \quad (1)$$

Assume that K uncorrelated narrow-band far-field sources from $\boldsymbol{\theta} = [\theta_1, \dots, \theta_K]^T$ impinge on the nested array. The array output at time t can be represented as

$$\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t) + \mathbf{n}(t), \quad (2)$$

where $\mathbf{x}(t) \in \mathbb{C}^{M \times 1}$ is the output of nested array. $\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_K(t)]^T$ is the $K \times 1$ receive signal source vector, and $\mathbf{n}(t) \in \mathbb{C}^{M \times 1}$ denotes the additive white Gaussian

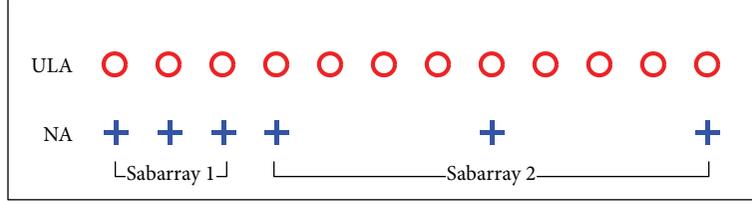


FIGURE 1: The nested array structure.

noise vector. The $M \times K$ array manifold matrix in array model (2) is defined by

$$\mathbf{A} = [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \dots, \mathbf{a}(\theta_K)], \quad (3)$$

where $\mathbf{a}(\theta_k) = [e^{j\pi p_1 \sin(\theta_k)}, e^{j\pi p_2 \sin(\theta_k)}, \dots, e^{j\pi p_M \sin(\theta_k)}]$ is the array steering vector and the element position is defined in (1). We note that each position p_i is measured with half wavelength of indent signal source. Let the \tilde{M} denote the maximum interelement distance (i.e., the array aperture). Then \tilde{M} is an integral multiple of the half signal wavelength. Nested array possesses the same set of interelement distances as the ULA which composed of \tilde{M} elements separated by half signal wavelength.

From the above discussion, we can derive that nested array can be regarded as spatial sparse sampling on the ULA, i.e., we can take the position set \mathcal{P} of nested array from the position subset $\tilde{\mathcal{P}} = [0, 1, \dots, \tilde{M} - 1]$ of the elements of the ULA, as shown in Figure 1. Thus, nested array can be represented by its element index set $\mathcal{P} \subset \tilde{\mathcal{P}}$. For the k -th signal source, the steering vector $\mathbf{a}_{\tilde{\mathcal{P}}}(\theta_k)$ of the ULA can be given as $\mathbf{a}_{\tilde{\mathcal{P}}}(\theta_k) = [e^{j2\pi\tilde{\mathcal{P}}_1(d/\lambda) \sin(\theta_k)}, \dots, e^{j2\pi\tilde{\mathcal{P}}_{\tilde{M}}(d/\lambda) \sin(\theta_k)}]$.

According to the index set \mathcal{P} of the NA, we design a selection matrix $\Gamma \in \{0, 1\}^{M \times \tilde{M}}$ such that the i -th row element contains $\mathbf{0}_s$ except a single 1 at the $\tilde{\mathcal{P}}_i$ -th location. From the spatial sparse sampling viewpoint, we have $\mathbf{a}(\theta_k) = \Gamma \mathbf{a}_{\tilde{\mathcal{P}}}(\theta_k)$. Then, the nested array model can be expressed as

$$\mathbf{x}(t) = \mathbf{A}(\theta_k)\mathbf{s}(t) + \mathbf{n}(t) = \Gamma \mathbf{A}_{\tilde{\mathcal{P}}}\mathbf{s}(t) + \Gamma \mathbf{n}_{\tilde{\mathcal{P}}}(t). \quad (4)$$

Under ideal condition, the array covariance matrix of the output $\mathbf{x}(t)$ of nested array can be calculated as

$$\begin{aligned} \mathbf{R} &= \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mathbf{x}(t_n) \mathbf{x}^H(t_n) = \mathbf{A} \mathbf{R}_s \mathbf{A}^H + \text{diag}(\boldsymbol{\sigma}^2) \\ &= \Gamma \mathbf{A}_{\tilde{\mathcal{P}}} \mathbf{R}_s \mathbf{A}_{\tilde{\mathcal{P}}}^H \Gamma^T + \Gamma \text{diag}(\boldsymbol{\sigma}_{n, \tilde{\mathcal{P}}}^2) \Gamma^T \\ &= \Gamma \text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}) \Gamma^T + \Gamma \text{diag}(\boldsymbol{\sigma}_{n, \tilde{\mathcal{P}}}^2) \Gamma^T, \end{aligned} \quad (5)$$

where N is the number of collected snapshot, $\mathbf{R}_s = \text{diag}([\sigma_{s,1}^2, \dots, \sigma_{s,K}^2])$, and $\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}) = \mathbf{A}_{\tilde{\mathcal{P}}} \mathbf{R}_s \mathbf{A}_{\tilde{\mathcal{P}}}^H$ for some $\mathbf{u}_{\tilde{\mathcal{P}}} \in \mathbb{C}^{\tilde{M} \times 1}$. Furthermore, the noise-free covariance matrix $\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u})$ is a Hermitian Toeplitz matrix, which is

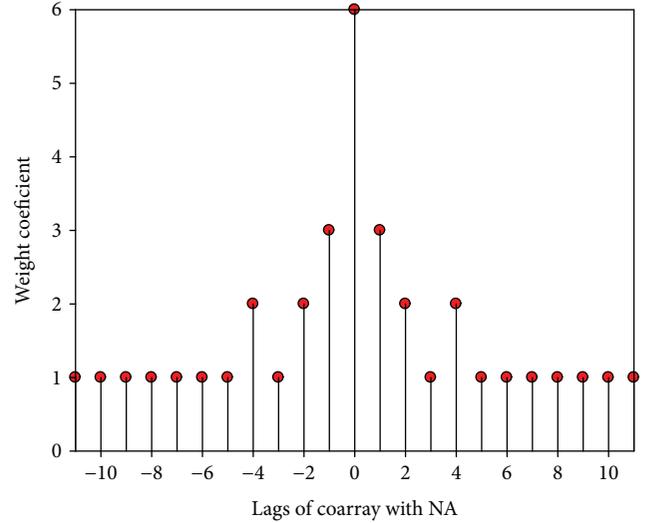


FIGURE 2: Weight coefficient of NA coarray with six elements.

determined by M complex numbers.

$$\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}) = \begin{bmatrix} u_1 & u_2 & \cdots & u_{\tilde{M}} \\ u_2^* & u_1 & \cdots & u_{\tilde{M}-1} \\ \vdots & \vdots & \ddots & \vdots \\ u_{\tilde{M}}^* & u_{\tilde{M}-1}^* & \cdots & u_1 \end{bmatrix}. \quad (6)$$

It is evident that $\text{rank}(\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u})) = K \leq \tilde{M} - 1$, and $\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u})$ is positive definite, i.e., $\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}) \geq 0$. For nested array, $\text{Toep}(\mathbf{u})$ holds that

$$\begin{aligned} \text{Toep}(\mathbf{u})_{m,n} &= \sum_{k=1}^K \sigma_{s,k}^2 a_m(\theta_k) a_n^*(\theta_k) \\ &= \sum_{k=1}^K \sigma_{s,k}^2 e^{j2\pi d/\lambda(m-n) \sin(\theta_k)}. \end{aligned} \quad (7)$$

Let $\mathcal{D} = \{m - n : m, n \in \mathcal{P}, m \geq n\} \subset \tilde{\mathcal{P}}$. For the NA, \mathcal{D} is a redundancy array which can be regarded as a coarray defines on the NA. The matrix $\text{Toep}(\mathbf{u})$ contains all elements $\mathbf{u}_{\tilde{\mathcal{P}}}$ of explicitly by (6). It implies that the relation $\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}) \rightarrow \text{Toep}(\mathbf{u})$ is one-to-one.

A coarray of nested array with six elements is shown in Figure 2. When the positions of nested array are $\mathcal{P} = \{0, 1, 2, 3, 6, 10\}$, the coarray of nested array composes of $\tilde{M} + 1 = 11$

elements with the position $\tilde{\mathcal{P}} = \{0, 1, \dots, 10\}$. From Figure 2, we can observe that the coarray is a redundancy array, which guaranteed the reconstruction of the covariance matrix [17].

One of the aims of this paper is to recover the noise-free covariance matrix $\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u})$ of the nested array output under nonuniform noise environment. In the next section, we introduce the reweighted nuclear norm minimization to approximate the low-rank noise-free covariance matrix $\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u})$.

3. Iterative Covariance Matrix Reconstruction for Direction Finding

3.1. Nuclear Norm Minimization Optimization. As mentioned above, the key problem is recovery of the noise-free covariance matrix $\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u})$. On one hand, we note that the covariance matrix $\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u})$ has three key features: Toeplitz-structured, low rank, and positive definite (PSD); on the other hand, in application, the sample covariance matrix of NA is calculated by $\hat{\mathbf{R}} = \sum_{n=1}^N \mathbf{x}(t_n) \mathbf{x}^H(t_n) / N$. Due to the finite samples, there exists the perturbation error term between the $\hat{\mathbf{R}}$ and \mathbf{R} , whose distribution benefits to choose the regularized parameter [17]. Making full use of the two constraints, we formulate the recovery problem of the covariance matrix as the following optimization problem:

$$\begin{aligned} \min_{\mathbf{u}, \sigma} \quad & \text{rank}(\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u})) \\ \text{s.t.} \quad & \left\| \hat{\mathbf{C}}^{-1/2} \text{vec}(\Delta \mathbf{R}) \right\|_2^2 \leq \eta \\ & \text{Toep}(\mathbf{u}) \geq 0, \end{aligned} \quad (8)$$

where $\hat{\mathbf{C}} = (1/N) \mathbf{R}^T \otimes \mathbf{R}$ represents the prewhitening weighted matrix, $\Delta \mathbf{R} = \hat{\mathbf{R}} - \mathbf{R}$ is the perturbation error term, and η is an upper bound of $\left\| \hat{\mathbf{C}}^{-1/2} \text{vec}(\Delta \mathbf{R}) \right\|_2^2$, which determines the quality of fitting term. According to the covariance fitting criteria [17, 18, 26], the vectorization form of the perturbation error term $\Delta \mathbf{R}$ satisfies the asymptotic normal (AsN) distribution as follows:

$$\text{vec}(\Delta \mathbf{R}) \sim \text{AsN} \left(\mathbf{0}_{M^2, 1}, \frac{1}{N} \mathbf{R}^T \otimes \mathbf{R} \right). \quad (9)$$

After prewhitening processing, we obtain

$$\hat{\mathbf{C}}^{-1/2} \text{vec}(\Delta \mathbf{R}) \sim \text{AsN}(0, \mathbf{I}_{M^2}). \quad (10)$$

We can further derive the following formulation:

$$\left\| \hat{\mathbf{C}}^{-1/2} \text{vec}(\Delta \mathbf{R}) \right\|_2^2 \sim \text{As}\chi^2(M^2), \quad (11)$$

where $\text{As}\chi^2(M^2)$ represents the asymptotic chi-square distribution with M^2 degrees of freedom. According to (11), it is quite obvious how to choose the parameter η . We choose the parameter η large enough such that the inequality $\left\| \hat{\mathbf{C}}^{-1/2} \text{vec}(\Delta \mathbf{R}) \right\|_2^2 \leq \eta$ is satisfied with a high probability $1 - p$,

e.g., 0.999, $(\Pr(\left\| \hat{\mathbf{C}}^{-1/2} \text{vec}(\Delta \mathbf{R}) \right\|_2^2 \leq \eta) = 1 - p)$, which is calculated by the matlab function $\text{chi2inv}(1 - p, M^2)$, so we have $\Pr\{\chi_{M^2}^2 \leq \eta\} = p$ and $\eta = \chi_p^2\{M^2\}$.

Unfortunately, this problem in (8) is nonconvex optimization problem and it is also a NP-hard problem. To sidestep the nonconvexity, we exploit convex relaxation technique, i.e., we relax the rank norm of $\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u})$ to the nuclear norm of $\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u})$ for approximating the low-rank matrix. Therefore, the convex relaxation form for (8) can be represented as

$$\begin{aligned} \min_{\mathbf{u}, \sigma} \quad & \left\| \text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}) \right\|_* \\ \text{s.t.} \quad & \left\| \hat{\mathbf{C}}^{-1/2} \text{vec}(\Delta \mathbf{R}) \right\|_2^2 \leq \eta \\ & \text{Toep}(\mathbf{u}) \geq 0. \end{aligned} \quad (12)$$

Combining (5) and (12), the optimization problem is rewritten as

$$\begin{aligned} \min_{\mathbf{u}, \sigma} \quad & \left\| \text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}) \right\|_* \\ \text{s.t.} \quad & \left\| \hat{\mathbf{C}}^{-1/2} \text{vec}(\hat{\mathbf{R}} - \Gamma(\text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}) + \text{diag}(\sigma_n^2))\Gamma^T) \right\|_2^2 \leq \eta \\ & \text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}) \geq 0. \end{aligned} \quad (13)$$

Remark 1. It is worth noting that there is noise variance variable that needs to be optimized in model (13). In fact, we can eliminate this variable by estimating the noise variance in advance, but the estimation of the nonuniform noise variance is not easy to attain directly, especially in the underdetermined DoA estimation circumstances while the sample covariance matrix no longer contains a low-rank component spanning the noise subspace. We can use two methods to handle this case, the first method is that we treat noise variance as a variable to be optimized in the optimal problem such as (13). The second method, we get rid of noise variables by linear transform firstly and then solve the optimization problem; in this paper, we focused on the latter.

It is interesting to observe that the noise term has M nonzero elements which contain in the diagonal element of the covariance matrix and correspond to the unknown noise variance $\{\sigma_n^2\}$. Based on this observation, we carry out the linear transform to remove the element of the noise variance terms. Mathematically, this operation can be formulated as

$$\begin{aligned} \text{Jvec}(\Delta \mathbf{R}) &= \text{Jvec}(\hat{\mathbf{R}} - (\Gamma \text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}) \Gamma^H + \Gamma \text{diag}(\sigma_n^2) \Gamma^H)) \\ &= \text{Jvec}(\hat{\mathbf{R}}) - \text{Jvec}(\Gamma \text{Toep}_{\tilde{\mathcal{P}}}(\mathbf{u}) \Gamma^H), \end{aligned} \quad (14)$$

where \mathbf{J} is an $M(M-1) \times M^2$ linear transform selecting matrix and can be represented as

$$\mathbf{J} = [\mathbf{J}_1, \mathbf{J}_2, \dots, \mathbf{J}_{M-1}]^T, \quad (15)$$

where

$$\mathbf{J}_m = \left[\mathbf{e}_{(m-1)(M-1)+2}, \mathbf{e}_{(m-1)(M-1)+3}, \dots, \mathbf{e}_{m(M-1)} \right], \quad (16)$$

$$m = 1, 2, \dots, M-1,$$

with $\mathbf{e}_i (i = (m-1)(M-1) + 2, \dots, m(M-1))$ being an $M^2 \times 1$ column vector with one at the i th position and zeros elsewhere. $\mathbf{J}_m \in \mathbb{R}^{M^2 \times M}$. By applying the above operation, the unknown noise variance is eliminated effectively and easily. This elimination operation avoids the estimation of noise variance, therefore facilitating a noise-free low-rank covariance matrix reconstruction.

Through the above derivation, (13) can further reformatted as

$$\begin{aligned} \min_{\mathbf{u}} \quad & \|\text{Toep}_{\tilde{\varphi}}(\mathbf{u})\|_* \\ \text{s.t.} \quad & \left\| \tilde{\mathbf{C}}^{-1/2} (\text{Jvec}(\hat{\mathbf{R}}) - \text{Jvec}(\mathbf{\Gamma} \text{Toep}_{\tilde{\varphi}}(\mathbf{u}) \mathbf{\Gamma}^H)) \right\|_2^2 \leq \hat{\eta} \\ & \text{Toep}_{\tilde{\varphi}}(\mathbf{u}) \geq 0. \end{aligned} \quad (17)$$

From (17), we observe that there is only one vector \mathbf{u} that needs to be optimized by the above processing. At the same time, the prewhitening matrix and regularization parameter should be changed. Combining the distribution of the perturbation error term with (14) yields

$$\text{Jvec}(\Delta \mathbf{R}) \sim \text{AsN} \left(\mathbf{0}_{M(M-1),1}, \frac{1}{N} \mathbf{J}(\mathbf{R}^T \otimes \mathbf{R}) \mathbf{J}^T \right). \quad (18)$$

Applying the prewhitening processing [19], we have

$$\tilde{\mathbf{C}}^{-1/2} \text{Jvec}(\Delta \mathbf{R}) \sim \text{AsN} \left(\mathbf{0}_{M(M-1),1}, \mathbf{I}_{M(M-1)} \right), \quad (19)$$

where the prewhitening matrix $\tilde{\mathbf{C}}^{-1/2}$ is Hermitian square root of $\tilde{\mathbf{C}}^{-1}$, $\tilde{\mathbf{C}} = \mathbf{J}(\mathbf{R}^T \otimes \mathbf{R}) \mathbf{J}^T / N$, and $\mathbf{I}_{M(M-1)}$ is an $M(M-1) \times M(M-1)$ identity matrix. Like the choosing parameter $\hat{\eta}$, the parameter $\tilde{\eta}$ can be determined by $\tilde{\eta} = \chi_p^2 \{M(M-1)\}$.

3.2. Iteratively Reweighted Nuclear Norm Minimization. We utilize the nuclear norm $\|\text{Toep}_{\tilde{\varphi}}(\mathbf{u})\|_* = \sum_i \sigma_i(\text{Toep}_{\tilde{\varphi}}(\mathbf{u})) = \sum_i g_1(\sigma_i(\text{Toep}_{\tilde{\varphi}}(\mathbf{u})))$ in (12) instead of the rank norm $\text{rank}(\text{Toep}_{\tilde{\varphi}}(\mathbf{u})) = \|\sigma(\text{Toep}_{\tilde{\varphi}}(\mathbf{u}))\|_0 = \sum_i g_0(\sigma_i(\text{Toep}_{\tilde{\varphi}}(\mathbf{u})))$ in (8), where $g_0(x) = \text{sign}(x)$, $x \geq 0$, $g_1(x) = x$, and $x \geq 0$. The advantage is effective computable and the best convex approximation, but the drawback is the worst fitting for rank approximation. We can observe that there is a large gap between the rank and nuclear norm for robust and exact reconstruction of low-rank covariance matrix. Therefore, it is

essential to find a surrogate function of rank function which not only have the better approximation but also needs a low computational complexity. Some methods utilize a class of nonconvex surrogate functions to approximate the rank function, such as the works in [27, 28].

Among the nonconvex approximation functions, we find $f(x) = 1 - e^{-x/\epsilon}$ closely matches $g_0(x)$, so we use this function $h^\epsilon(x) = f(x) = 1 - e^{-x/\epsilon}$ as the approximation function of $g_0(x)$, i.e., rank norm. Combine the model (17), the following general nonconvex optimization model is proposed:

$$\begin{aligned} \min_{\mathbf{u}} \quad & h^\epsilon(\text{Toep}_{\tilde{\varphi}}(\mathbf{u})) \\ \text{s.t.} \quad & \left\| \tilde{\mathbf{C}}^{-1/2} (\text{Jvec}(\hat{\mathbf{R}}) - \text{Jvec}(\mathbf{\Gamma} \text{Toep}_{\tilde{\varphi}}(\mathbf{u}) \mathbf{\Gamma}^H)) \right\|_2^2 \leq \hat{\eta} \\ & \text{Toep}_{\tilde{\varphi}}(\mathbf{u}) \geq 0, \end{aligned} \quad (20)$$

where $h^\epsilon(\sigma(\text{Toep}_{\tilde{\varphi}}(\mathbf{u}))) = \sum_i f(\sigma_i(\text{Toep}_{\tilde{\varphi}}(\mathbf{u})))$ denotes the rank approximating function. We expect that the proposed nonconvex optimization model can bridge the gap between the rank norm model (8) and nuclear norm (12).

Since the model in (20) is nonconvex optimization problem, how to guarantee for convergence to a global minimum is a key problem. Now, we propose an iteratively reweighted nuclear norm minimization to obtain a suboptimal solution, which is based on the framework of the majorization-maximization (MM) method. The MM method is an iterative approach, it iteratively solves a sequence of weighted convex surrogate optimization problems whose weights are attained by the previous optimal value. Thus, by calculating the sequence of convex problems given in (22), we can minimize the optimization function in (20). For the k -th iteration, the weights \mathbf{u} are calculated by the previous optimal variable \mathbf{u}_{k-1} . Since $h^\epsilon(\text{Toep}_{\tilde{\varphi}}(\mathbf{u})) = \sum_i f(\sigma_i(\text{Toep}_{\tilde{\varphi}}(\mathbf{u})))$ is concave when $f(x) = g_1(x)$, we have

$$\begin{aligned} h^\epsilon(\text{Toep}_{\tilde{\varphi}}(\mathbf{u})) & \leq h^\epsilon(\text{Toep}_{\tilde{\varphi}}(\mathbf{u}_k)) \\ & \quad + \|\nabla h^\epsilon[\text{Toep}_{\tilde{\varphi}}(\mathbf{u})]\| \text{Toep}_{\tilde{\varphi}}(\mathbf{u} - \mathbf{u}_k)\|_*. \end{aligned} \quad (21)$$

Therefore, the optimization problem for the k -th iteration becomes

$$\begin{aligned} \min_{\mathbf{u}} \quad & \|\mathbf{W}_k^\epsilon \text{Toep}_{\tilde{\varphi}}(\mathbf{u})\|_* \\ \text{s.t.} \quad & \left\| \tilde{\mathbf{C}}^{-1/2} (\text{Jvec}(\hat{\mathbf{R}}) - \text{Jvec}(\mathbf{\Gamma} \text{Toep}_{\tilde{\varphi}}(\mathbf{u}) \mathbf{\Gamma}^H)) \right\|_2^2 \leq \hat{\eta} \\ & \text{Toep}_{\tilde{\varphi}}(\mathbf{u}) \geq 0, \end{aligned} \quad (22)$$

where $\mathbf{W}_k^\epsilon \triangleq \nabla h^\epsilon(\text{Toep}_{\tilde{\varphi}}(\mathbf{u}_k))$. In order to calculate \mathbf{W}_k^ϵ , we start by introducing a key useful proposition as follows.

Proposition 1. Suppose that $h^\epsilon(\sigma(\mathbf{Z}))$ is expressed as $h^\epsilon(\sigma(\mathbf{Z})) = \sum_i f(\sigma_i(\mathbf{Z}))$, where \mathbf{Z} is positive semidefinite

```

1: Input:  $\mathbf{X} \in \mathbb{C}^{M \times K}$ .
2: Initialize:  $k = 0$ ,  $\mathbf{u}_0 = \mathbf{0}_{M \times 1}$ ,  $\epsilon$ ,  $\gamma$ ,  $\epsilon = 10^{-4}$ ,  $\text{Iter}_{\max}$ 
3: Calculate the sample covariance matrix  $\hat{\mathbf{R}}_S$  of NA using
   the observe snapshots  $\mathbf{x}_\Omega(t)$ ;
4: while  $k \leq \text{Iter}_{\max}$  do
5:   Update the weight matrix  $\mathbf{W}_k^\epsilon$  by the equation
6:   Update  $\mathbf{u}_{k+1}$  by the optimization problem (22), (24);
7:   if the convergence condition  $|\mathbf{u}_{k+1} - \mathbf{u}_k| \leq \epsilon$  is satisfied
8:     then
9:       break
10:    else
11:       $\epsilon = \epsilon/\gamma$ ,  $k = k + 1$ ;
12:    end if
13: end while
14: Calculate the covariance matrix  $\hat{\mathbf{R}} = \text{Toep}(\hat{\mathbf{u}})$ ;
15: Using search-free DoA estimation method to estimate the
   angle parameters.

```

ALGORITHM 1: Gridless DoA estimation using iteratively reweighted covariance matrix reconstruction.

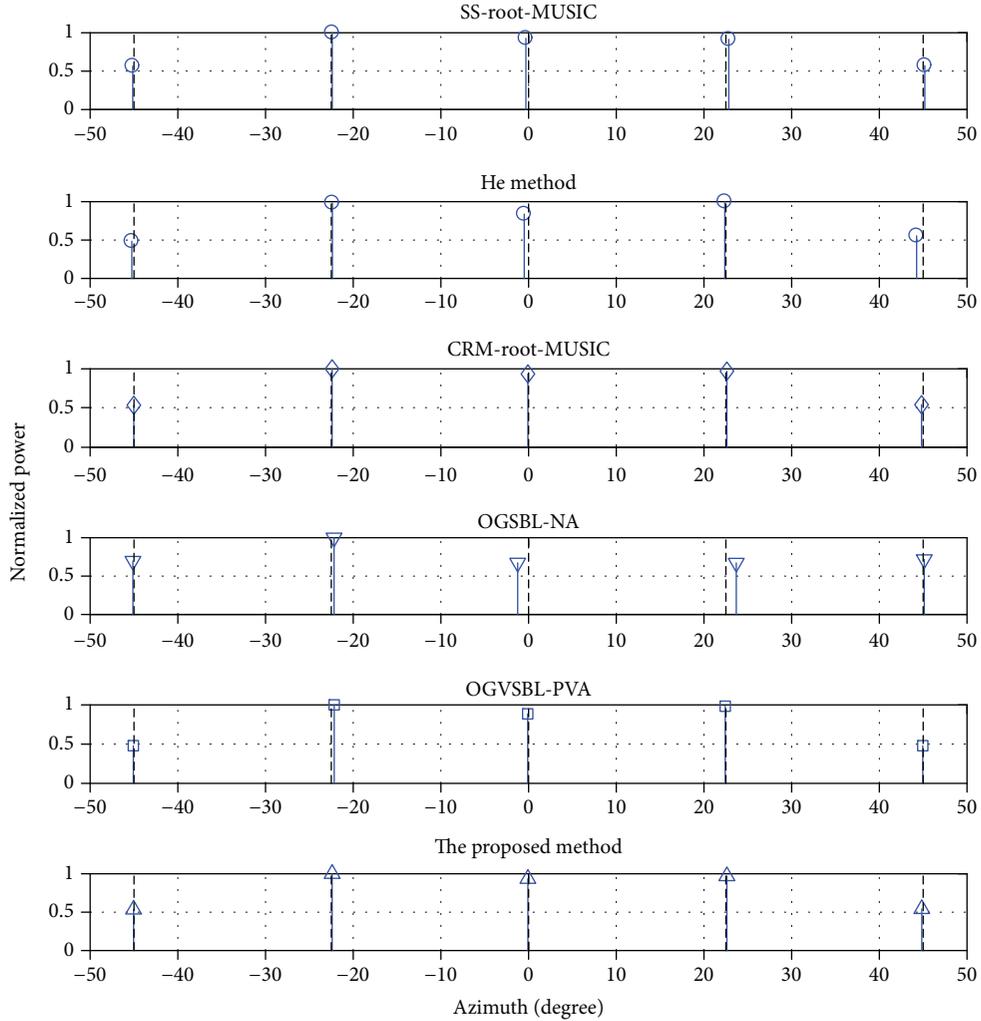


FIGURE 3: Normalized spatial spectra in nonuniform noise case for $M = 6$, $K = 5$, $N = 500$, and $\text{SNR} = 0$ dB.

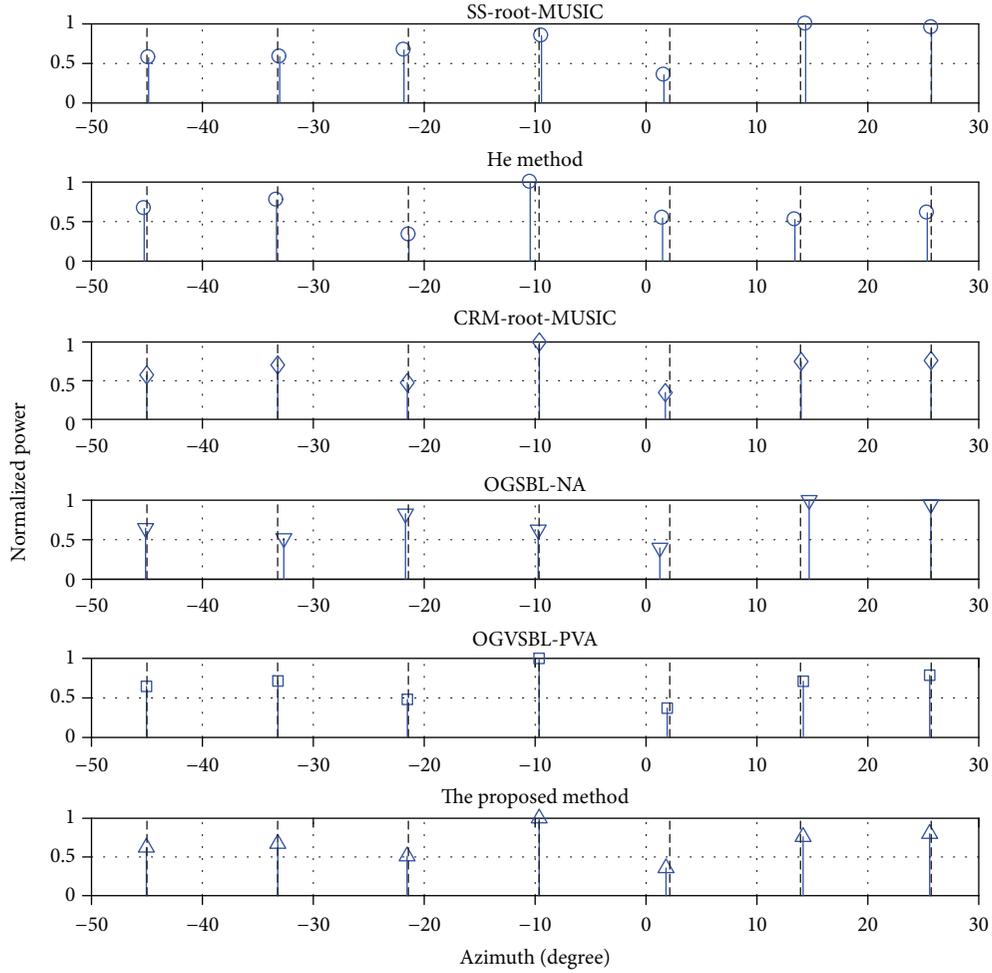


FIGURE 4: Normalized spatial power in nonuniform noise case for $M = 6$, $K = 7$, $N = 500$, and $\text{SNR} = 0$ dB.

matrices with the eigenvalue decomposition (EVD) $\mathbf{Z} = \mathbf{U} \text{diag}(\boldsymbol{\sigma}(\mathbf{Z})) \mathbf{U}^H$, $\mathbf{Z} \geq 0$, and h^ϵ and f are concave and differentiable, then the gradient of $f(\mathbf{Z})$ at \mathbf{Z} is calculated as

$$\nabla h^\epsilon(\mathbf{Z}) = \mathbf{U} \text{diag}(\boldsymbol{\eta}(\mathbf{Z})) \mathbf{U}^H, \quad (23)$$

where $\boldsymbol{\eta}(\mathbf{Z}) = \nabla h^\epsilon(\boldsymbol{\sigma}(\mathbf{Z}))$ denotes the gradient of h^ϵ at $\boldsymbol{\sigma}(\mathbf{Z})$.

In (22), $\mathbf{Z} = \text{Toep}_{\hat{\varphi}}(\mathbf{u}_k)$, for the sake of simplicity, we denote $\boldsymbol{\sigma}(\text{Toep}_{\hat{\varphi}}(\mathbf{u}_k))$ as $\boldsymbol{\sigma}_k$ and apply the result in [27]; we obtain

$$\mathbf{W}_k^\epsilon = \mathbf{U}_k \text{diag}(\nabla h^\epsilon(\boldsymbol{\sigma}_k)) \mathbf{U}_k^H, \quad (24)$$

where $\mathbf{U}_k \text{diag}(\boldsymbol{\sigma}_k) \mathbf{U}_k^H$ is the EVD of $\text{Toep}_{\hat{\varphi}}(\mathbf{u}_k)$.

We observe that ϵ is a key parameter, which controls the closeness between $f(\boldsymbol{\sigma})$ and $g_0(\boldsymbol{\sigma})$ or $g_1(\boldsymbol{\sigma})$. In particular, $f(\boldsymbol{\sigma})$ is close to $g_0(\boldsymbol{\sigma})$ when ϵ is small, but it is subjected to many local minima. On the other hand, a large value of ϵ causes $f(\boldsymbol{\sigma})$ toward the convex $g_1(\boldsymbol{\sigma})$, whereas with poor approximation quality to $g_0(\boldsymbol{\sigma})$. In order to avoid getting

local minima, the algorithm starts with a large value of ϵ , and then the value ϵ decreases gradually during the iteration.

It is worth mentioning that this strategy is adopted in model (22). We define the weight matrix \mathbf{W}_k^ϵ in each iteration, which uses the solution of the previous iteration for avoiding local minima. The convergence analysis of the proposed method for (20) is given as the following theorem, which is proven in the appendix.

Theorem 1. Denote \mathbf{u}_k as the k -th iteration solution in (22), after that, the optimal variable \mathbf{u}_k satisfies the following properties:

- (i) $h^\epsilon[\text{Toep}_{\hat{\varphi}}(\mathbf{u}_k)]$ is monotonically decreasing and bounded as k tends to ∞
- (ii) The optimal variable \mathbf{u}_k converges to a local minimum of (20)

3.3. Steps of the Proposed Method. To sum up, the steps of the proposed method are presented as follows. Once the solution \mathbf{u} of model (20) is obtained, we can determine the number of

sources \hat{K} by the rank of $\text{Toep}_{\hat{\mathcal{P}}}(\mathbf{u})$. Given $\text{Toep}_{\hat{\mathcal{P}}}(\hat{\mathbf{u}})$, the remaining task is to estimate the directions of the signal $\hat{\boldsymbol{\theta}}$. According to classical Vandermonde decomposition lemma [28], we can estimate the DoAs using positive semidefinite Toeplitz matrices $\text{Toep}_{\hat{\mathcal{P}}}(\mathbf{u})$. So we can apply the grid-free DoA estimation method, such as Prony's method to efficiently estimate $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\sigma}}_s^2$. Firstly, by solving a linear system involving only $\hat{\mathbf{u}}$, we can obtain $\hat{\boldsymbol{\theta}}$ from zeros of a polynomial. Then, $\hat{\boldsymbol{\sigma}}_s^2$ is solved by a polynomial rooting-based estimation method.

In the proposed method, firstly, it recovers the complete noiseless matrix $\text{Toep}_{\hat{\mathcal{P}}}(\mathbf{u})$ in advance under the nonuniform noise background, which is attained by solving an iteratively weighted nuclear norm minimization optimization. Compare with the ULAs which have the same number of physical elements, the proposed method forms a larger aperture ULA and gets more degree of freedoms, so it has the potential ability to estimate the DoA in the underdetermined case. Moreover, exploiting the grid-free DoA estimation method to estimate the DoAs, which do not need to discretize the angle grid, avoids the grid mismatch in the grid-based sparse method.

4. Simulation Results

In this section, we evaluate the DoA estimation performance of the proposed method and compare the proposed method with the others: the state-of-the-art method, spatial smoothing root MUSIC (SS-root-MUSIC) [7], covariance matrix reconstruction approach with root MUSIC (CMR-root-MUSIC) [22], He et al.'s method [23], OGSBL-NA [25], and OGVSBL-PVA [24]. To provide a benchmark for evaluating the effectiveness of these methods, the Cramer-Rao lower bound (CRLB) for nonuniform noise given in [23] is also included though it is a lower bound for unbiased estimators.

In the following simulations, we consider a two-level nested array with optimal configuration [8], whose total element number is six. The first ULA subarray contains three elements with interelement spacing $d_1 = d$; the second ULA subarray contains three elements with spacing $d_2 = 4d_1$, where d is equal to half wavelength. We assume that all the signals are far-field equal power uncorrelated signals; the additive noise is spatially nonuniform which is modeled as an independent complex Gaussian random vector with zero mean, and its covariance matrix is $\mathbf{R}_n = \text{diag}\{\boldsymbol{\sigma}_n^2\}$, since the input signal-to-noise ratio (SNR) can be defined as

$$\text{SNR} = 10 \log_{10} \left[\frac{1}{M} \sum_{m=1}^M \frac{\sigma_s^2}{\sigma_{n,m}^2} \right], \quad (25)$$

where $M = N_1 + N_2$ is the number of NA elements, σ_s^2 denotes the signal source power, and $\sigma_{n,m}^2$ denotes the noise power in difference elements. For the performance metric of DoA estimation, the root-mean-square error (RMSE) is

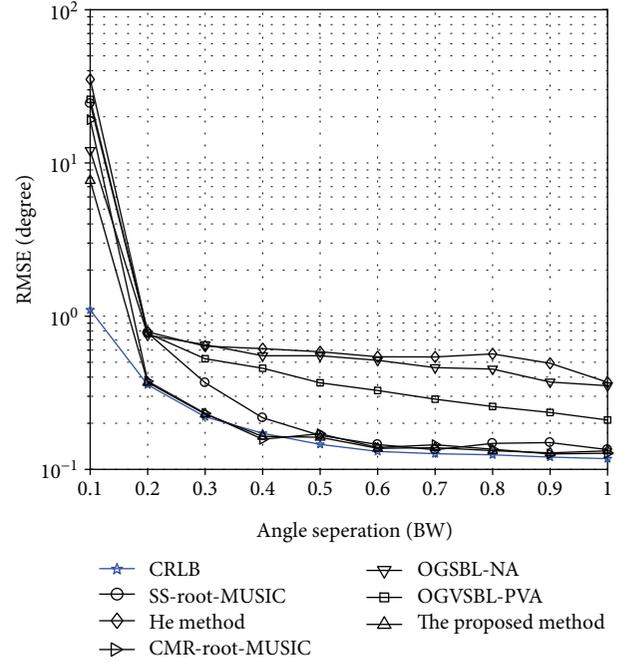


FIGURE 5: RMSEs of DoA estimation versus angle separation for $M = 6$, $N = 500$, and $\text{SNR} = 0$ dB.

selected for evaluating the estimation accuracy of the DoAs, which is defined as

$$\text{RMSE} = \left[\frac{1}{LK} \sum_{k=1}^K \sum_{l=1}^L \left(\theta_k^r - \hat{\theta}_{k,l} \right)^2 \right]^{1/2}, \quad (26)$$

where $\hat{\theta}_{k,l}$ denote the estimated θ_k^r of the k -th signal source from the l -th Monte Carlo trial.

In the first experiment, we investigate the resolution ability under the overdetermined case with the nested array geometries, respectively. We consider $K = 5$ spatial uncorrelated sources which are uniformly distributed within -45° to 45° . The SNR is set to 0 dB, and 500 snapshots are collected. The nonuniform noise is $\boldsymbol{\sigma}_n^2 = [10, 5.6, 8.5, 4.2, 9.5, 0.5]$. In He et al.'s method, we set the grid size is 0.5° . Figure 3 shows the normalized power of the five targets for the nested array, respectively. We use the dashed lines to represent the true DoAs. It is observed that the proposed method can exactly resolve all targets among all methods, and it has the most superior estimation performance since the DoA estimations are closest to the true target directions. The reason is that the proposed method makes the best of the features of covariance matrix, i.e., low rank, Toeplitz structure, and PSD, and approach the true rank by the iterative weighted idea. The OGVSBL-PVA and the CMR-root-MUSIC are a little worse than the proposed method, so we confirmed that the proposed method achieved higher DoA estimation accuracy.

In the second experiment, we investigate the detectable capability of all methods for the underdetermined case. We consider $K = 7$ narrowband uncorrelated sources whose number is greater than the number of physical elements,

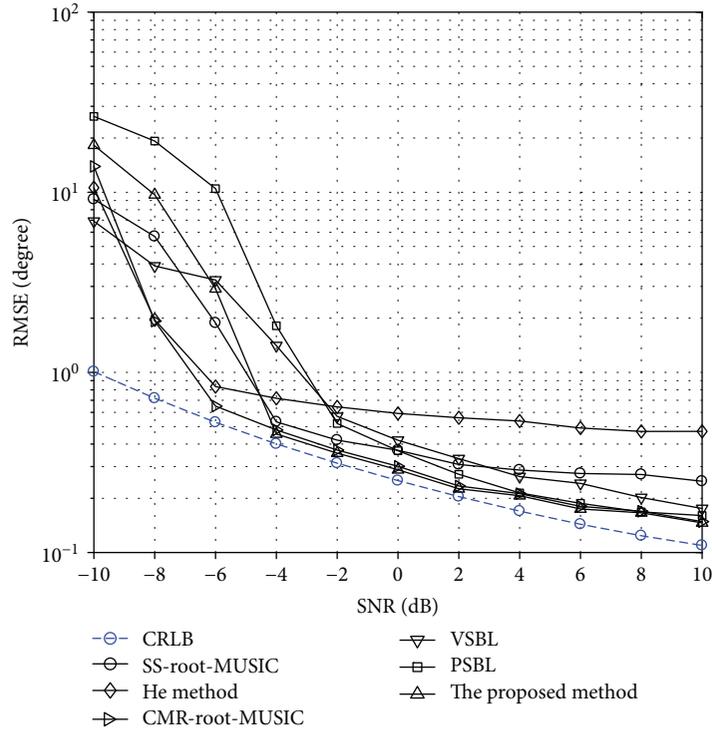


FIGURE 6: RMSEs of DoA estimation versus SNR for $M = 6$, $K = 7$, and $N = 500$.

the signal source directions are uniformly distributed within the range from $-\pi/4$ to $-\pi/7$, and the nonuniform noise power is assumed that $\sigma_n^2 = [10, 5.6, 8.5, 4.2, 9.5, 0.5]$; the SNR is set to 0 dB, and the number of snapshots is 500. Figure 4 displays the normalized power of all methods, respectively. It is observed that the OGVSBL-PVA method and the proposed method can detect all the signal sources exactly. He et al.'s method has a certain deviation; this method uses the denoising operation and covariance matrix fitting to estimate the DoAs, which is based on the on-grid model. From Figure 4, it is easy to observe that the DoA estimation of the proposed method is closest to the true DoAs among all the methods; the reason is that the NA array provides more DOFs than the number of physical elements, and the low-rank covariance matrix approximation is more exact than the covariance matrix reconstruction method by exploiting the iteratively reweighted nuclear norm minimization; moreover, the proposed method is not limited by the grid size like the traditional grid-based sparsity method. Thus, the proposed methods have the superior estimation performance in all methods.

In the third experiment, in order to further illustrate the superior performance of the proposed method, the RMSE of DoA estimation performance versus the angular separation between the source is examined for all methods. In the experiment, the DoA of the first source is fixed and equal to -5.3° , whereas the DoA of the second source is varied and the varied step size is equal to 0.1 times beam width. When the angle separation is larger than 0.2 times beam width, the performance of the proposed method and

CMR-root-MUSIC is achieving the CRLB; since the covariance matrix reconstruction method is based on the low-rank property, it circumvents the traditional grid search, which is the gridless method. The OGSBL-NA method and OGVSBL-PVA method also have the better DoA performance due to the linear approximation manifold vector. As shown in Figure 5, compared with other methods, the performance of the proposed method demonstrates the better performance even when the target is very close.

In the following experiments, unless otherwise specified, the element number in nested array is set to six; seven uncorrelated equal power signals are considered which are uniformly distributed between $-\pi/4$ and $-\pi/7$. In this case, when grid size is set to 0.5° , for He et al.'s method, the DoAs do not fall into the predefined angular grid, i.e., grid mismatch. The nonuniform noise power is assumed that $\sigma_n^2 = [10, 5.6, 8.5, 4.2, 9.5, 0.5]$. All simulation results were obtained by averaging over 500 independent Monte Carlo trials.

In the fourth experiment, we examine the RMSE of DoA estimation performance with regard to the SNR for different methods. Figure 6 shows the RMSE of DoA estimation performance as a function of SNR when the number of snapshots is 500. Figure 7 shows that the RMSE versus the number of snapshots with the SNR is fixed to 0 dB. As we can see in Figures 6 and 7, the DoA estimation performance of the proposed method is improved with the SNR and increasing snapshot number, since the covariance matrix approximation is more accurate with larger SNR and more snapshots. The proposed method and the CMR-root-MUSIC outperform OGSBL-NA,

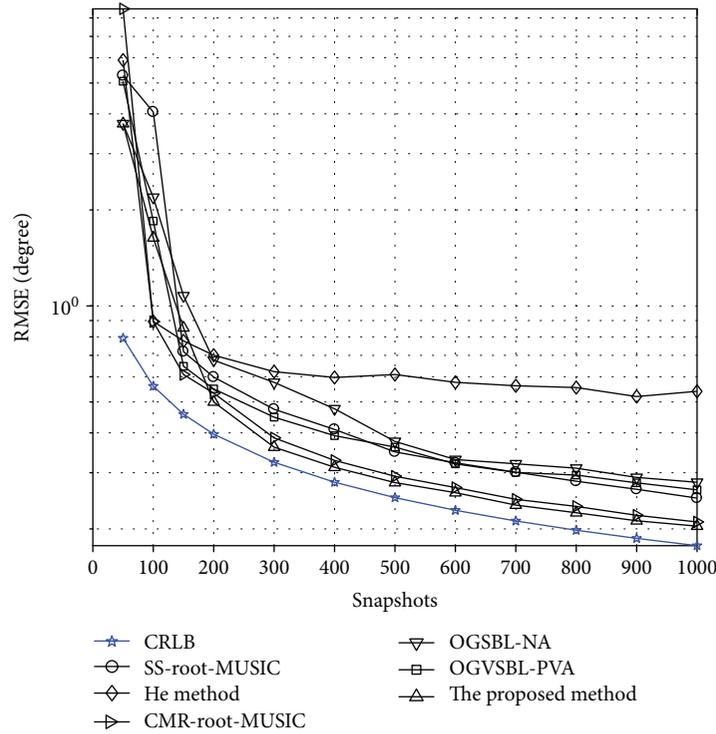


FIGURE 7: RMSEs of DoA estimation versus snapshot number for $M = 6$, $K = 7$, and $SNR = 0$ dB.

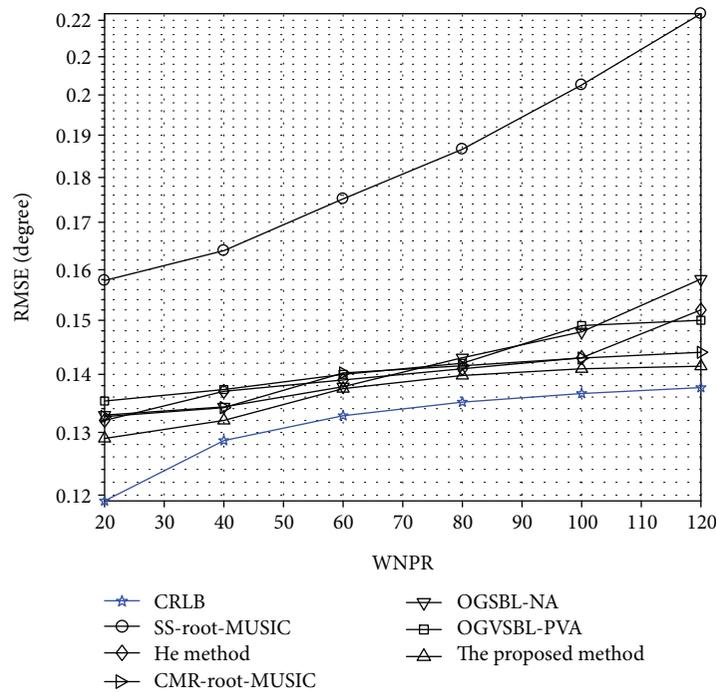


FIGURE 8: RMSEs of DoA estimation versus WNPR for $M = 6$, $N = 500$, and $SNR = 0$ dB.

OGVSBL-PVA, and the other methods. CMR-root-MUSIC method and the proposed method provide almost similar estimation performance, while the proposed method is slightly better than CMR-root-MUSIC. Since the proposed method not only exploits the full noiseless covariance matrix but also

utilizes the iteratively reweighted strategy to enhance the sparsity of matrix rank resulting in better covariance matrix approximation, we can see that in most of the cases He et al.'s method showcases the worst performance with the largest error; the reason is that He et al.'s method only utilizes

the part of the noiseless covariance matrix due to the denoising operation and when the DoAs do not fall into the predefined angular grid, the performance will degrade. OGSBL-NA method and OGVsBL-PVA method both improve the DoA performance by linear approximation manifold vector, but the precise depends on the order of the Taylor expansion. It needs to point that the proposed method considerably enhances the sparsity of the covariance matrix due to the iteratively weighted nuclear norm strategy; moreover, the search-free method is free of grid restriction. Thus, we can see that the proposed method has the best performance of DoA estimation accuracy in all tested methods.

In the last experiment, we measure the DoA estimation performance in the nonuniform noise environment. We define the worst noise power ratio (WNPR) $WNPR = \sigma_{n_{\max}}^2 / \sigma_{n_{\min}}^2$ [29], where $\sigma_{n_{\max}}^2$ and $\sigma_{n_{\min}}^2$ denote the maximum and the minimum noise power, respectively. The RMSE of DoA estimation performance for all methods is shown in Figure 8 as a function of the WNPR. The SNR is fixed to 0 dB. We vary the WNPR from 20 to 120 by changing the noise power of the first element from 10 to 60. It is observed that the proposed method performs better than the other methods at all WNPR levels; this can be explained by the fact that using denoising operation mitigates the influence of the nonuniform noise and enhances the robustness of the proposed method; furthermore, by iteratively reweighted idea, the proposed method can exactly approximate the rank of covariance matrix, and search-free DoA method is not subject to the grid constraint.

5. Conclusion

In this paper, we addressed the direction finding problem in the presence of unknown nonuniform noise over nested array. We proposed a gridless DoA estimation method based on the low-rank covariance matrix approximation. In order to mitigate the effect of nonuniform noise, we eliminated the nonuniform noise variables on the diagonal of the covariance matrix by linear transform. In order to recover the noise-free covariance matrix, we proposed an iterative reweighted non-convex optimization method, whose regularization parameter is determined by the covariance fitting criteria. Finally, we exploited the search-free DoA estimation method to estimate the exact DoA parameters. Numerical simulations were carried out to verify the effectiveness of the proposed method, compared with the state-of-the-art methods, which achieves more accurate DoA estimation performance for the nonuniform noise environment in the presence of grid bias.

Appendix

Proof. Assume $h^\epsilon(\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u})) = \sum_i f(\sigma_i(\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u})))$. If function f is twice differentiable, strictly concave, for any bounded k and $0 \leq x \leq k$, $f'' \leq -m'_k < 0$, and then $h^\epsilon(\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u}))$ is strictly concave. For any bounded $\mathbf{u}, \mathbf{v} (\mathbf{u} \neq \mathbf{v})$, there exists some $\lambda > 0$ such that

$$\begin{aligned} & h^\epsilon(\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u})) - h^\epsilon(\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{v})) \\ & \leq \langle \text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u} - \mathbf{v}), \nabla f^\epsilon(\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{v})) \rangle - \frac{\lambda}{2} \|\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u} - \mathbf{v})\|_F^2. \end{aligned} \quad (\text{A.1})$$

We first show that $h^\epsilon(\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u}_k))$ is convergent. In terms of the concavity of $h^\epsilon(\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u}_k))$ and classical Vandermonde decomposition lemma [28], we have

$$\begin{aligned} & h^\epsilon(\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u}_{k+1})) - h^\epsilon(\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u}_k)) \\ & \leq \|\nabla h^\epsilon(\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u}_k))\|_* \|\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u}_{k+1} - \mathbf{u}_k)\|_* \\ & \quad - \frac{\lambda}{2} \|\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u}_{k+1} - \mathbf{u}_k)\|_F^2, \end{aligned} \quad (\text{A.2})$$

which directly results in

$$\begin{aligned} & h^\epsilon(\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u}_k)) - h^\epsilon(\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u}_{k+1})) \\ & \geq \|\nabla h^\epsilon(\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u}_k))\|_* \|\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u}_k - \mathbf{u}_{k+1})\|_*. \end{aligned} \quad (\text{A.3})$$

According to the model (22), we update \mathbf{u}_{k+1} and deduce that

$$\begin{aligned} & \|\nabla h^\epsilon(\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u}_k))\|_* \|\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u}_{k+1})\|_* \\ & \leq \|\nabla h^\epsilon(\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u}_k))\|_* \|\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u}_k)\|_*, \end{aligned} \quad (\text{A.4})$$

which together with (A.3) confirms that

$$h^\epsilon(\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u}_k)) - h^\epsilon(\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u}_{k+1})) \geq 0. \quad (\text{A.5})$$

With the aid of the result on the $h^\epsilon(\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u})) = \sum_i h^\epsilon(\sigma_i \text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u})) \geq 0$, we can draw the first property. In order to derive the second property, we prove that the optimal variable $\{\mathbf{u}_k\}$ is convergent. By applying classical Vandermonde decomposition lemma [28] on $h^\epsilon(\text{Toep}(\mathbf{u}))$, we have

$$\begin{aligned} & h^\epsilon(\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u}_k)) - h^\epsilon(\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u}_{k+1})) \\ & \geq \|\nabla h^\epsilon(\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u}_k))\|_* \|\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u}_k - \mathbf{u}_{k+1})\|_* \\ & \quad + \frac{\lambda}{2} \|\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u}_{k+1} - \mathbf{u}_k)\|_F^2 \geq \frac{\lambda}{2} \|\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u}_{k+1} - \mathbf{u}_k)\|_F^2. \end{aligned} \quad (\text{A.6})$$

For all $k \geq 0$ summing the inequality above, we can attain

$$\frac{\lambda}{2} \sum_{k=0}^{+\infty} \|\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u}_{k+1} - \mathbf{u}_k)\|_F^2 \leq h^\epsilon(\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u}_0)), \quad (\text{A.7})$$

which implies that $\lim_{k \rightarrow \infty} \|\text{Toep}_{\hat{\mathcal{D}}}(\mathbf{u}_{k+1} - \mathbf{u}_k)\|_F = 0$. Thus, $\{\mathbf{u}_k\}$ is convergent. In order to prove that $\{\mathbf{u}_k\}$ converges to a local minimum, it is noted that the models (20) and (22) have the same Karush-Kuhn-Tucker (KKT) conditions when the optimal variable $\{\mathbf{u}_k\}$ converges, i.e., $\mathbf{u}_k \rightarrow \mathbf{u}^*$ as $k \rightarrow \infty$. The reason is that the gradients of the objective functions and the constraints in the rank norm are the same as in nuclear norm models when the MM method

converges. We conclude that \mathbf{u}^* tend to be a local minimum of (20) since the objective function in (20) monotonically decreases. This completes the proof.

Data Availability

The authors claim that all data generated or analysed in this article are provided by their simulations and are included in this published article. The matlab data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant no. 61671378). The authors would like to thank Dr. Weiqiang Tan and Professor Guo Wang for their comments that greatly improved the manuscript.

References

- [1] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE Signal Processing Magazine*, vol. 13, no. 4, pp. 67–94, 1996.
- [2] W. Tan, M. Matthaiou, S. Jin, L. Fan, and G. K. Karagiannidis, "On the spectral efficiency of massive MIMO systems with hybrid DFT processing," *IEEE Transactions on Wireless Communications*, vol. 1, no. 1, pp. 1–13, 2018.
- [3] X. Wu, W. P. Zhu, and J. Yan, "A high-resolution DOA estimation method with a family of nonconvex penalties," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 6, pp. 4925–4938, 2018.
- [4] W. Tan, M. Matthaiou, S. Jin, and X. Li, "Spectral efficiency of DFT-based processing hybrid architectures in massive MIMO," *IEEE Wireless Communications Letters*, vol. 6, no. 5, pp. 586–589, 2017.
- [5] A. Moffet, "Minimum-redundancy linear arrays," *IEEE Transactions on Antennas and Propagation*, vol. 16, no. 2, pp. 172–175, 1968.
- [6] P. P. Vaidyanathan and P. Pal, "Sparse sensing with co-prime samplers and arrays," *IEEE Transactions on Signal Processing*, vol. 59, no. 2, pp. 573–586, 2011.
- [7] P. Pal and P. P. Vaidyanathan, "Coprime sampling and the music algorithm," in *2011 Digital Signal Processing and Signal Processing Education Meeting (DSP/SPE)*, pp. 289–294, Sedona, AZ, USA, 2011.
- [8] P. Pal and P. P. Vaidyanathan, "Nested arrays: a novel approach to array processing with enhanced degrees of freedom," *IEEE Transactions on Signal Processing*, vol. 58, no. 8, pp. 4167–4181, 2010.
- [9] A. J. Weiss and M. Gavish, "Direction finding using esprit with interpolated arrays," *IEEE Transactions on Signal Processing*, vol. 39, no. 6, pp. 1473–1478, 1991.
- [10] B. Friedlander and A. J. Weiss, "Direction finding using spatial smoothing with interpolated arrays," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 28, no. 2, pp. 574–587, 1992.
- [11] B. Friedlander, "The root-music algorithm for direction finding with interpolated arrays," *Signal Processing*, vol. 30, no. 1, pp. 15–29, 1993.
- [12] S. Pillai and F. Haber, "Statistical analysis of a high resolution spatial spectrum estimator utilizing an augmented covariance matrix," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, no. 11, pp. 1517–1523, 1987.
- [13] Y. I. Abramovich, D. A. Gray, A. Y. Gorokhov, and N. K. Spencer, "Positive-definite Toeplitz completion in DOA estimation for nonuniform linear antenna arrays. I. Fully augmentable arrays," *IEEE Transactions on Signal Processing*, vol. 46, no. 9, pp. 2458–2471, 1998.
- [14] Y. I. Abramovich, N. K. Spencer, and A. Y. Gorokhov, "Positive-definite Toeplitz completion in DOA estimation for nonuniform linear antenna arrays. II. Partially augmentable arrays," *IEEE Transactions on Signal Processing*, vol. 47, no. 6, pp. 1502–1521, 1999.
- [15] T. E. Tuncer, T. K. Yasar, and B. Friedlander, "Direction of arrival estimation for nonuniform linear arrays by using array interpolation," *Radio Science*, vol. 42, no. 4, 2007.
- [16] C. L. Liu and P. P. Vaidyanathan, "Remarks on the spatial smoothing step in coarray MUSIC," *IEEE Signal Processing Letters*, vol. 22, no. 9, pp. 1438–1442, 2015.
- [17] Z. Yang and L. Xie, "On gridless sparse methods for line spectral estimation from complete and incomplete data," *IEEE Transactions on Signal Processing*, vol. 63, no. 12, pp. 3139–3153, 2015.
- [18] Z. M. Liu, Z. T. Huang, and Y. Y. Zhou, "Sparsity-inducing direction finding for narrowband and wideband signals based on array covariance vectors," *IEEE Transactions on Wireless Communications*, vol. 12, no. 8, pp. 1–12, 2013.
- [19] J. Yin and T. Chen, "Direction-of-arrival estimation using a sparse representation of array covariance vectors," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4489–4493, 2011.
- [20] Y. Tian and X. Sun, "Passive localization of mixed sources jointly using MUSIC and sparse signal reconstruction," *AEU - International Journal of Electronics and Communications*, vol. 68, no. 6, pp. 534–539, 2014.
- [21] W. Tan, X. Feng, X. Ye, and L. Jing, "Direction-of-arrival of strictly non-circular sources based on weighted mixed-norm minimization," *EURASIP Journal on Wireless Communications and Networking*, vol. 2018, no. 1, 2018.
- [22] W. Tan and X. Feng, "Covariance matrix reconstruction for source localization over impaired uniform linear array," in *2017 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*, pp. 1–5, Xiamen, China, 2017.
- [23] Z. Q. He, Z. P. Shi, and L. Huang, "Covariance sparsity-aware DOA estimation for nonuniform noise," *Digital Signal Processing*, vol. 28, pp. 75–81, 2014.
- [24] C. Wen, X. Xie, and G. Shi, "Off-grid DOA estimation under nonuniform noise via variational sparse Bayesian learning," *Signal Processing*, vol. 137, pp. 69–79, 2017.
- [25] J. Yang, G. Liao, and J. Li, "An efficient off-grid DOA estimation approach for nested array signal processing by using sparse Bayesian learning strategies," *Signal Processing*, vol. 128, pp. 110–122, 2016.

- [26] B. Ottersten, P. Stoica, and R. Roy, "Covariance matching estimation techniques for array signal processing applications," *Digital Signal Processing*, vol. 8, no. 3, pp. 185–210, 1998.
- [27] M. Malek-Mohammadi, M. Babaie-Zadeh, and M. Skoglund, "Iterative concave rank approximation for recovering low-rank matrices," *IEEE Transactions on Signal Processing*, vol. 62, no. 20, pp. 5213–5226, 2014.
- [28] U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications*, University of California Press, 2001.
- [29] M. Pesavento and A. B. Gershman, "Maximum-likelihood direction-of-arrival estimation in the presence of unknown nonuniform noise," *IEEE Transactions on Signal Processing*, vol. 49, pp. 1310–1324, 2002.

