Research Article

Reweighted Covariance Fitting Based on Nonconvex Schatten-p Minimization for Gridless Direction of Arrival Estimation

Lama Zien Alabideen , Oumayma Al-Dakkak, and Khaldoun Khorzom

Department of Telecommunication, Higher Institute for Applied Sciences and Technology, Damascus 31983, Syria

Correspondence should be addressed to Lama Zien Alabideen; lama.zienalabideen@hiast.edu.sy

Received 12 December 2019; Accepted 25 March 2020; Published 27 April 2020

Academic Editor: Luigi Rodino

Copyright © 2020 Lama Zien Alabideen et al. 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 reformulate the gridless direction of arrival (DoA) estimation problem in a novel reweighted covariance fitting (CF) method. The proposed method promotes joint sparsity among different snapshots by means of nonconvex Schatten-p quasi-norm penalty. Furthermore, for more tractable and scalable optimization problem, we apply the unified surrogate for Schatten-p quasi-norm with two-factor matrix norms. Then, a locally convergent iterative reweighted minimization method is derived and solved efficiently via a semidefinite program using the optimization toolbox. Finally, numerical simulations are carried out in the background of unknown nonuniform noise and under the consideration of coprime array (CPA) structure. The results illustrate the superiority of the proposed method in terms of resolution, robustness against nonuniform noise, and correlations of sources, in addition to its applicability in a limited number of snapshots.

1. Introduction

DoA estimation is a major task in array signal processing. It has wide applications in radar, sonar, wireless communications, underwater acoustics, and seismology [1–3]. The main goal of DoA estimation is to retrieve the direction information of several sources from the outputs of the antenna array. The study of DoA estimation has a long history. One of an outstanding class of methods is designated as subspace-based methods such as root multiple signal classification (RootMUSIC) [4] and the estimation of parameters by rotational invariant techniques (ES-PRIT) [5], and their variants. In general, these methods have high estimation accuracy with low computational complexity, but they suffer from considerable limitations. For example, they require a priori knowledge of the number of sources that may be difficult to acquire in practice. Moreover, they need a sufficient number of data snapshots to estimate the data covariance matrix precisely. In addition, they are sensitive to source correlations that may lead to a rank deficiency in the sample data covariance matrix.

With the development of sparse representation and compressed sensing (CS) [6, 7], the last decade has received remarkable advances resulting in a number of sparse methods for DoA estimation. Sparse methods can be applied in several challenging scenarios, including cases with no knowledge of the number of sources, a limited number of snapshots, and highly correlated sources. The early grid-based sparse methods have been focused on discrete linear systems, meaning that the DoAs have to lie on a set of predefined grid points so that the collected snapshots can be sparsely represented under a finite discrete dictionary. However, in reality, the DoAs are continuously valued and nonlinear with the measured data; hence, there is no guarantee that the true DoAs will be on the grid. This results in a grid or basis mismatch problem [8, 9]. To address this issue, several off-grid methods have been proposed [10–12]. However, these methods bring complexity issue while the grid selection remains a major problem since gridding still exists. Recently, gridless sparse methods have been developed [13, 14]. These methods eliminate the requirement of gridding by working directly in a continuous infinite dictionary and recover the DoAs by solving a semidefinite programming (SDP) problem.

2. Related Works

The gridless sparse methods in DoA estimation are categorized into two methodologies. The first one is based on the concept of atomic norm minimization (ANM) [15]. While, the second methodology makes use of sparse representation of second-order statistics of the array output, which is designated as gridless CF.

ANM has been successfully applied in estimating a spectrally-sparse signal without any statistical assumptions in [12, 16]. It has been shown that the frequencies can be exactly estimated if they are sufficiently separated by at least $(4/n)$, where n denotes the size of full data. Later, this condition is reduced to $(2.52/n)$ in [17]. Initially, the atomic norm soft-thresholding (AST) method was introduced in [18] for full data case in the presence of stochastic noise. Then, the study was extended to the case of compressive data for a single measurement vector (SMV), where the problem was referred to as continuous CS (CCS) in [12]. Next, CCS was expanded to the case of multiple measurement vectors (MMVs) in [19]. In order to alleviate the limitation of estimation resolution due to the effect of separation condition, the reweighted atomic-norm minimization (RAM) was proposed in [20] with applicability to DoA estimation.

The gridless CF technique in DoA estimation is not new. The covariance matching estimation technique (COMET) in [21] and the asymptotic maximum likelihood (AML) technique in [22] are both consistent gridless DoA estimators as the signal-to-noise ratio (SNR) increases. In addition, the continuous sparse recovery (CSR) method [23] and the nuclear norm minimization method followed by MUSIC (NNM-MUSIC) [13], do not require the number of sources to estimate the DoAs precisely, while the knowledge of error bounds is essential. The gridless sparse iterative covariance-based estimation (GL-SPICE or GLS) method [24] (also referred to as the sparse and parametric approach (SPA) [25]) has many merits. Mainly, GLS is a hyper-parameter-free method that can estimate the noise variance(s) consistently, also it is applicable to an arbitrary number of snapshots and robust against moderate sources correlation. However, in the case of finite snapshots, GLS suffers from the separation condition due to the theoretical relation with ANM. Recently, the covariance matrix reconstruction approach (CMRA) was developed in [26] as a super-resolution method. Nevertheless, CMRA faces difficulties in highly correlated sources scenarios. Furthermore, the study in [27] has proposed an improved version of CMRA (ICMRA) by applying a family of nonconvex penalties on the singular values of the covariance matrix. ICMRA has shown an enhancement in sparsity and resolution compared to CMRA. On the contrary, ICMRA suffers from unsatisfactory performance in the presence of low SNR and low number of snapshots. A clearer comparison of the abovementioned methods is provided in Table 1.

Concerning noise, many sparse methods assume that the antenna array noises are spatially uniform white with equal variances. This assumption helps in choosing a suitable regularization parameter in a sparse representation optimization formula. In several practical applications, this may be violated

since the noise levels among all antennas exhibit some perturbations due to the imperfection of the array calibration, the nonidealities of the communication channels, or the mutual coupling between antennas. In many sparse methods, the antennas are sparsely placed so their noises are spatially uncorrelated and modeled as a diagonal covariance matrix [28, 29], but the antennas noise variances are no longer identical (i.e., nonuniform noise). Such a noise model is relevant in situations with CF-based methods [14], while ANM-based methods are favored when the noise is uniform with bounded power or when the sources are highly correlated. In fact, the nonuniform noise is a challenging problem in sparse estimation methods, where the sample covariance matrix is no longer low-rank. To eliminate this problem the recent research studies get rid of noise in two major methods, the first one is by using linear transform primarily and then solving noise-free optimization problem as in [30, 31]. However, this method needs to design a linear transform able to separate the noise from the signal, which is pointless in the general cases, specifically when the transformation requires prior knowledge of the noise parameters. The second method treats the noise variance as a variable to be optimized in the optimization problem [14, 25]. In this paper, the latter method is adopted.

Despite the convexity of the preliminary gridless optimization models, it has been shown that the estimation performance of such convex relaxation will degrade in the presence of measurement noise or when some assumptions are violated (e.g., spatial sources separation limit and signals correlation). Hence, the solution may significantly deviate from the original (i.e., rank minimization). In order to achieve a better approximation of the original solution, many nonconvex surrogate functions have been proposed, including L_p norm, Logarithm, Laplace, and Schatten- p norm which are well-studied in the literature [32, 33]. In [27], the DoA problem has been formulated in a nonconvex minimization problem using a family of nonconvex penalties on the singular values of the covariance matrix resulting in an enhancement in both sparsity and resolution. Recently, we have proposed a high-resolution gridless method for noiseless joint sparse frequency estimation named modified reweighted atomic norm minimization (MRAM) [34]. MRAM was based on nonconvex relaxation of ANM by Schatten- p norm. This approximation has offered improved performance compared to trace norm relaxation. Hence, due to the one-to-one relation between frequency and direction information, we look forward to achieve a better gridless DoA estimation using Schatten- p quasi-norm penalty.

3. Main Contributions

Since, gridless CF is preferable in the presence of a non-uniform noise background, in this paper we develop a new gridless sparse method for DoA estimation based on recently developed gridless CF methods. By exploiting the positive semidefinite (PSD) and Toeplitz structure of the parameterized data covariance matrix, a gridless CF criterion is formulated in an efficient joint sparsity optimization problem. More specifically, we propose a novel objective function based on introducing the nonconvex Schatten- p quasi-norm as a sparsity metric. Then, for a more tractable

TABLE 1: Gridless DoA estimation methods and their comparisons.

Gridless sparse methodology Specifications	Methods										
	Atomic norm				Covariance fitting						
	AST	[12]	RAM	MRAM	SPA/GLS	CMRA	ICMRA	AML	COMET	CSR	NNM-MUSIC
Need for error bound	✓		✓			✓	✓			✓	✓
Need for model order	✓	✓							✓		
Angle separation condition	✓	✓	✓	✓	✓		✓			✓	✓
Sparsity promoting	✓	✓	✓	✓		✓	✓			✓	✓
Noise estimation	✓				✓			✓	✓		
Feasibility to low number of snapshots	✓	✓	✓	✓	✓						
Immunity to sources correlation	✓	✓	✓	✓	✓						
High detection probability in low SNR						✓		✓	✓		✓

and scalable optimization problem, the unified surrogate for Schatten-p quasi-norm with two-factor matrix norms is applied. Consequently, an iterative reweighted minimization method is derived with a SDP formulation. Regarding the major role of the array geometry and in order to acquire a high degree of freedom with less mutual coupling and a reasonable number of physical antennas, the coprime array (CPA) structure [23, 35] is adopted. Then, the new method has been tested, where the numerical results have shown an improved performance as compared to the state-of-art methods.

Notations 1. Throughout this paper, the bold letters are reserved for matrices and vectors. The transpose is denoted by $(\cdot)^T$, and the complex conjugate or Hermitian is denoted by $(\cdot)^H$. $\text{tr}(\cdot)$ and $\text{rank}(\cdot)$ represent matrix trace and rank, respectively. $E[\cdot]$ stands for the expectation. $\|\cdot\|_2$, $\|\cdot\|_*$, and $\|\cdot\|_F$ refer to l_2 , nuclear, and Forbenius norms, respectively. $\mathbf{A} \geq 0$ means that $\mathbf{A}$ is a positive semidefinite matrix.

4. Problem Formulation

Assume there are K uncorrelated narrow-band far-field source signals s_k , impinge from directions $\boldsymbol{\theta} = [\theta_1, \theta_2, \dots, \theta_K]^T$ on a sparse linear CPA array. In particular, for a coprime number pair (N, M) with $N > M$. CPA consists of two uniform linear subarrays, the first consists of M elements having interelement spacing of Nd , whereas the other consists of N elements having interelement spacing of Md , where d is chosen to be half of the wavelength $\lambda/2$. According to the properties of coprime integers, the elements of the subarrays do not overlap with each other when they are aligned except the first element. Therefore, CPA consists of $M + N - 1$ physical antennas in total. The CPA output data matrix can be considered as the output of virtual $N \times M$ elements uniform linear array (ULA) by retaining the antennas indexed by $\boldsymbol{\Omega} = [r_1, r_2, \dots, r_{M+N-1}] \subset \{1, \dots, NM\}$. The observation model of CPA at t snapshot can be modeled as follows:

$$\mathbf{y}_\Omega(t) = \sum_{k=1}^K \mathbf{a}_\Omega(\theta_k) s_k(t) + \mathbf{n}_\Omega(t) = \mathbf{A}_\Omega(\boldsymbol{\theta}) \mathbf{s}(t) + \mathbf{n}_\Omega(t),$$

$$t = 1, \dots, L, \quad (1)$$

where L is the number of snapshots that are jointly sparse, in the sense that they share the same support where the DoAs are encoded, and $\mathbf{y}_\Omega(t) \in \mathbb{C}^{(M+N-1) \times 1}$, $\mathbf{s}(t) \in \mathbb{C}^{K \times 1}$, and $\mathbf{n}_\Omega(t) \in \mathbb{C}^{(M+N-1) \times 1}$ are the array output, the source signal vector, and the noise powers vector, respectively. $\mathbf{A}_\Omega(\boldsymbol{\theta}) = [\mathbf{a}_\Omega(\theta_1), \dots, \mathbf{a}_\Omega(\theta_K)] \in \mathbb{C}^{(M+N-1) \times K}$ denotes the steering matrix, where $\mathbf{a}_\Omega(\theta_k)$ is the steering vector of the k^{th} source which is determined by the geometry of the antenna array and is given by

$$\mathbf{a}_\Omega(\theta_k) = \left[e^{j2\pi r_1(d/\lambda)\sin(\theta_k)}, e^{j2\pi r_2(d/\lambda)\sin(\theta_k)}, \dots, e^{j2\pi r_{M+N-1}(d/\lambda)\sin(\theta_k)} \right]^T, \quad (2)$$

where r_i is the physical positions of the i^{th} CPA element. Furthermore, it is assumed that $\{\mathbf{s}\}_1^K$ are spatially and temporarily uncorrelated with independently and uniformly distributed phases. It follows that $E[\mathbf{s}\mathbf{s}^H] = \text{diag}(\mathbf{P})$, where $\mathbf{P} = [P_1, P_2, \dots, P_K] \in \mathbb{R}_+^{1 \times K}$ is the vector of the powers of sources.

Moreover, suppose the noise $\{\mathbf{n}_\Omega\}_1^{M+N-1}$ is independent of $\{\mathbf{s}\}_1^K$ and satisfies $E[\mathbf{n}_\Omega \mathbf{n}_\Omega^H] = \text{diag}(\boldsymbol{\sigma}_\Omega)$, where $\boldsymbol{\sigma}_\Omega \in \mathbb{R}_+^{(M+N-1) \times 1}$ denotes the noise variances vector whose elements are assumed to be unequal in this study. Briefly, the array output for L snapshots can be modeled as

$$\mathbf{Y}_\Omega = [\mathbf{y}_\Omega(1), \dots, \mathbf{y}_\Omega(L)] = \mathbf{A}_\Omega(\boldsymbol{\theta}) \mathbf{S} + \mathbf{N}_\Omega \in \mathbb{C}^{(M+N-1) \times L}, \quad (3)$$

where $\mathbf{S} = [\mathbf{s}(1), \dots, \mathbf{s}(L)] \in \mathbb{C}^{K \times L}$ is the sources signals matrix and $\mathbf{N}_\Omega = [\mathbf{n}_\Omega(1), \dots, \mathbf{n}_\Omega(L)] \in \mathbb{C}^{(M+N-1) \times L}$ is the measurement noise matrix.

Concerning the main goal of DoA estimation, in this paper, we seek to estimate the unknown parameters $(\boldsymbol{\theta}, \mathbf{P}, \boldsymbol{\sigma}_\Omega)$ given the measurement matrix $\mathbf{Y}_\Omega$.

5. Gridless CF with Nonconvex Schatten-p Minimization

In this section, we propose a new gridless method to estimate the parameters $(\boldsymbol{\theta}, \mathbf{P}, \boldsymbol{\sigma}_\Omega)$ by introducing the nonconvex Schatten-p norm regularization into the same CF criterion as in Sparse Iterative Covariance-based Estimation (SPICE) method [36] and its gridless version method (GL-SPICE or GLS) [24, 25]. Mainly, the proposed method inherits the

merits of GLS, with enhanced performance as it is verified by the numerical simulations in Section 6.

5.1. Gridless CF. The key idea of gridless CF is exploiting the structure of the data covariance matrix and performing the estimation under ungrided parameterized domain with specific statistical considerations (i.e., uncorrelated sources and reasonable number of snapshots). Based on the assumptions in Section 4, the uncorrelated data $\{\mathbf{y}_\Omega\}$ have the data covariance matrix $\mathbf{R}_\Omega \in \mathbb{C}^{(M+N-1) \times (M+N-1)}$ given by

$$\mathbf{R}_\Omega = E\{\mathbf{y}_\Omega(t)\mathbf{y}_\Omega^H(t)\} = \mathbf{A}_\Omega(\boldsymbol{\theta})\text{diag}(\mathbf{P})\mathbf{A}_\Omega^H(\boldsymbol{\theta}) + \text{diag}(\boldsymbol{\sigma}_\Omega). \quad (4)$$

Note that $\mathbf{R}_\Omega^{-1}$ exists in the presence of noise for $\boldsymbol{\sigma}_\Omega > 0$. $\mathbf{R}_\Omega$ is a highly nonlinear function of the DoAs parameters. To linearize the problem, a reparameterization of $\mathbf{R}_\Omega$ using a PSD Toeplitz matrix $\mathbf{T}(\mathbf{u})$ is performed. Hence, according to Vandermonde decomposition lemma (see, e.g., [3]), $\mathbf{R}_\Omega$ can be linearly reparametrized as

$$\mathbf{R}_\Omega(\mathbf{u}, \boldsymbol{\sigma}_\Omega) = \mathbf{T}_{\mathbf{u}, \Omega} + \text{diag}(\boldsymbol{\sigma}_\Omega), \quad \mathbf{T}(\mathbf{u}) \geq 0, \boldsymbol{\sigma}_\Omega > 0, \quad (5)$$

where $\mathbf{T}_{\mathbf{u}, \Omega} = \Gamma_\Omega \mathbf{T}(\mathbf{u}) \Gamma_\Omega^T = \mathbf{A}_\Omega(\boldsymbol{\theta})\text{diag}(\mathbf{P})\mathbf{A}_\Omega^H(\boldsymbol{\theta})$ for $\mathbf{T}(\mathbf{u}) \in \mathbb{C}^{NM \times NM}$ is a PSD Toeplitz matrix formed by the vector $\mathbf{u} \in \mathbb{C}^{1 \times NM}$. $\Gamma_\Omega \in \{0, 1\}^{(M+N-1) \times NM}$ denotes a selection matrix, such that the i^{th} row of Γ_Ω contains all zeros but a single 1 at the r_i position. The importance of the parameterization in equation (5) is that $\mathbf{T}(\mathbf{u})$ is linear in the vector $\mathbf{u}$, which comprises the interesting parameters.

Recall, the sample covariance matrix is set as $\tilde{\mathbf{R}}_\Omega = (1/L)\mathbf{Y}_\Omega \mathbf{Y}_\Omega^H$. In this paper, we are mainly interested in the case, where $L \geq M + N - 1$ (i.e., $\tilde{\mathbf{R}}_\Omega^{-1}$ exists with probability one) and $K \leq NM - 1$ (i.e., $\mathbf{T}(\mathbf{u})$ is rank-deficient, then $(\boldsymbol{\theta}, \mathbf{P})$ can be uniquely determined [25]). As well, we consider the same CF criterion as in the SPICE method, which is given by

$$\begin{aligned} f(\mathbf{u}, \boldsymbol{\sigma}_\Omega) &= \left\| \mathbf{R}_\Omega^{-(1/2)} (\tilde{\mathbf{R}}_\Omega - \mathbf{R}_\Omega) \tilde{\mathbf{R}}_\Omega^{-(1/2)} \right\|_F^2 \\ &= \text{tr}(\mathbf{R}_\Omega^{-1} \tilde{\mathbf{R}}_\Omega) + \text{tr}(\tilde{\mathbf{R}}_\Omega^{-1} \mathbf{R}_\Omega) - 2NM. \end{aligned} \quad (6)$$

By using the linear parameterization in equation (5), the CF can be accomplished by minimizing f leading to the following equivalences:

$$\begin{aligned} \min_{\mathbf{u}, \{\boldsymbol{\sigma}_\Omega \geq 0\}} \quad & \text{tr}(\tilde{\mathbf{R}}_\Omega^{(1/2)} (\mathbf{T}_{\mathbf{u}, \Omega} + \text{diag}(\boldsymbol{\sigma}_\Omega))^{-1} \tilde{\mathbf{R}}_\Omega^{(1/2)}) \\ & + \text{tr}(\tilde{\mathbf{R}}_\Omega^{-1} \mathbf{T}_{\mathbf{u}, \Omega}) + \text{tr}(\tilde{\mathbf{R}}_\Omega^{-1} \text{diag}(\boldsymbol{\sigma}_\Omega)) \end{aligned} \quad (7)$$

subject to $\mathbf{T}(\mathbf{u}) \geq 0$,

$$\begin{aligned} \min_{\mathbf{X}, \mathbf{u}, \{\boldsymbol{\sigma}_\Omega \geq 0\}} \quad & \text{tr}(\mathbf{X}) + \text{tr}(\tilde{\mathbf{R}}_\Omega^{-1} \mathbf{T}_{\mathbf{u}, \Omega}) + \text{tr}(\tilde{\mathbf{R}}_\Omega^{-1} \text{diag}(\boldsymbol{\sigma}_\Omega)) \\ \text{subject to} \quad & \mathbf{T}(\mathbf{u}) \geq 0, \mathbf{X} \geq \tilde{\mathbf{R}}_\Omega^{(1/2)} (\mathbf{T}_{\mathbf{u}, \Omega} + \text{diag}(\boldsymbol{\sigma}_\Omega))^{-1} \tilde{\mathbf{R}}_\Omega^{(1/2)}, \end{aligned} \quad (8)$$

$$\begin{aligned} \min_{\mathbf{X}, \mathbf{u}, \{\boldsymbol{\sigma}_\Omega \geq 0\}} \quad & \text{tr}(\mathbf{X}) + \text{tr}(\tilde{\mathbf{R}}_\Omega^{-1} \mathbf{T}_{\mathbf{u}, \Omega}) + \text{tr}(\tilde{\mathbf{R}}_\Omega^{-1} \text{diag}(\boldsymbol{\sigma}_\Omega)) \\ \text{subject to} \quad & \begin{bmatrix} \mathbf{X} & \tilde{\mathbf{R}}_\Omega^{(-1/2)} \\ \tilde{\mathbf{R}}_\Omega^{(-1/2)} & \mathbf{T}_{\mathbf{u}, \Omega} + \text{diag}(\boldsymbol{\sigma}_\Omega) \end{bmatrix} \geq 0. \end{aligned} \quad (9)$$

Hence, if f minimization is casted as SDP, then it is convex. Problem (9) was derived under the development of the GLS method. In particular, the GLS method transforms the direction estimation problem into the estimation of a PSD Toeplitz matrix in which the parameters of the directions are encoded. Therefore, once the SDP in problem (9) is solved, the parameters of interest can be obtained from the estimated covariance matrix by either using the conventional subspace methods or applying the Vandermonde decomposition lemma.

In spite of the strong theoretical guarantees of the GLS method, it suffers from considerable limitations according to Table 1. In fact, we believe that its performance can be further improved by employing the nonconvex penalty, which to the best of our knowledge, has not been studied in the literature. Note that the nonconvex Schatten- p quasi-norm regularization is able to achieve a more accurate and sparser solution than the trace norm. In this paper, we reformulate the minimization problem of f in a new objective function by introducing Schatten- p quasi-norm. Then, the resultant optimization problem is modeled as a SDP formula.

5.2. Schatten- p Norm. Let $\|\cdot\|_p$ be the Schatten- p norm of the matrix $\mathbf{M} \in \mathbb{C}^{n \times m}$, which is defined as the l_p norm of its singular values $\sigma_k(\mathbf{M})$, i.e.,

$$\begin{aligned} \|\mathbf{M}_p\| &= \left(\sum_{k=1}^{\min(n,m)} \sigma_k^p(\mathbf{M}) \right)^{(1/p)} = \left(\text{tr}((\mathbf{M}^T \mathbf{M})^{(p/2)}) \right)^{(1/p)}, \\ &0 < p < \infty. \end{aligned} \quad (10)$$

Recently, Schatten- p norm has received significant attention from researchers in various domains [27, 33], due to its outstanding properties that make it a good choice in many optimization problems (e.g., rank function relaxation). Specifically, $\|\mathbf{M}\|_p$ equals to the nuclear norm when $p = 1$. As well, if $0 < p < 1$, then the quasi-norm (for $0 < p < 1$, $\mathbf{M}_p$ is a quasi-norm because it does not satisfy the triangle inequality) $\|\mathbf{M}\|_p$ tends to rank($\mathbf{M}$). Moreover, Zhang et al. [37] verified theoretically that $\|\mathbf{M}\|_p$ with small p requires significantly fewer measurements, with superior performance over the conventional nuclear norm in the majority of the problems.

5.3. The Proposed Method. Motivated by the properties of Schatten- p norm and GLS method, we develop a new method by replacing the $\text{tr}(\tilde{\mathbf{R}}_\Omega^{-1} \mathbf{T}_{\mathbf{u}, \Omega})$ term in problem (8) by its Schatten- p quasi-norm, leading to the following optimization problem:

$$\begin{aligned} \min_{\mathbf{X}, \mathbf{u}, \{\sigma \geq 0\}} \quad & tr(\mathbf{X}) + \lambda \|\mathbf{Z}(\mathbf{u})\|_p^p + tr(\tilde{\mathbf{R}}_\Omega^{-1} \text{diag}(\boldsymbol{\sigma}_\Omega)), \quad 0 < p < 1 \\ \text{subject to} \quad & \mathbf{T}(\mathbf{u}) \geq 0, \mathbf{X} \geq \tilde{\mathbf{R}}_\Omega^{(1/2)} (\mathbf{T}_{\mathbf{u}, \Omega} + \text{diag}(\boldsymbol{\sigma}_\Omega))^{-1} \tilde{\mathbf{R}}_\Omega^{(1/2)}, \end{aligned} \quad (11)$$

where $\lambda > 0$ is a regularization parameter and $\mathbf{Z}(\mathbf{u}) = \tilde{\mathbf{R}}_\Omega^{-1} \mathbf{T}_{\mathbf{u}, \Omega}$ is a PSD matrix clearly:

$$\|\mathbf{Z}(\mathbf{u})\|_p^p = \sum_{k=1}^{M+N-1} \sigma_k^p(\mathbf{Z}(\mathbf{u})), \quad 0 < p < 1. \quad (12)$$

According to equation (10), $\|\mathbf{Z}(\mathbf{u})\|_p^p$ is given as where $\{\sigma_k(\mathbf{Z}(\mathbf{u})) \geq 0\}_1^{M+N-1}$ are descendingly sorted singular values of $\mathbf{Z}(\mathbf{u})$.

Since $0 < p < 1$, then for any $p, p_1, p_2 > 0$ satisfying

$$\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2}, \quad (13)$$

the bi-Schatten- p norm surrogate can be applied successfully [38]. In other words, the Schatten- p quasi-norm of $\mathbf{Z}(\mathbf{u})$ matrix is equivalent to minimizing the product of the Schatten- p_1 norm (or quasi-norm) and Schatten- p_2 norm (or quasi-norm) of its two factor matrices. So, as an alternative to the nonconvex quasi-norm in (12), the following efficient factorization is used:

$$\frac{1}{p} \|\mathbf{Z}(\mathbf{u})\|_p^p = \min_{\substack{\tilde{\mathbf{R}}_\Omega^{-1} \\ \mathbf{T}_{\mathbf{u}, \Omega} \\ \mathbf{Z} = \tilde{\mathbf{R}}_\Omega^{-1} \mathbf{T}_{\mathbf{u}, \Omega}}} \frac{1}{p_1} \|\tilde{\mathbf{R}}_\Omega^{-1}\|_{p_1}^{p_1} + \frac{1}{p_2} \|\mathbf{T}_{\mathbf{u}, \Omega}\|_{p_2}^{p_2}. \quad (14)$$

Notice that $\tilde{\mathbf{R}}_\Omega^{-1}$ is fixed, so whatever the value of p_1 is on minimizing Problem (14), one only needs to minimize $(1/p_2) \|\mathbf{T}_{\mathbf{u}, \Omega}\|_{p_2}^{p_2}$. Substituting the spectral regularization in Problem (11) with Problem (14) and setting $\lambda = (1/p)$, we infer the following problem:

$$\begin{aligned} \min_{\mathbf{X}, \mathbf{u}, \{\sigma \geq 0\}} \quad & tr(\mathbf{X}) + \frac{1}{p_2} \|\mathbf{T}_{\mathbf{u}, \Omega}\|_{p_2}^{p_2} + tr(\tilde{\mathbf{R}}_\Omega^{-1} \text{diag}(\boldsymbol{\sigma}_\Omega)) \\ \text{subject to} \quad & \mathbf{T}(\mathbf{u}) \geq 0, \mathbf{X} \geq \tilde{\mathbf{R}}_\Omega^{(1/2)} (\mathbf{T}_{\mathbf{u}, \Omega} + \text{diag}(\boldsymbol{\sigma}_\Omega))^{-1} \tilde{\mathbf{R}}_\Omega^{(1/2)}. \end{aligned} \quad (15)$$

With respect to equation (13), if $p_2 \geq 1$, Problem (15) is strongly convex; thus, it can be easily solved by off-the-shelf algorithms. When $p_2 < 1$, then Problem (15) becomes nonconvex and must be relaxed. However, we empirically found that the later problem has superior properties in terms of estimation performance and solution sparsity. Therefore, in the following, we consider only the nonconvex model and work on the relaxation issue of Problem (15).

Let's set $p_2 = p$ for simplicity. Since the proposed objective function in Problem (15) is nonconvex, then it must be linearized. A popular locally convergent approach to the minimization of such a problem is the majorization maximization (MM) method [39]. Accordingly, at the l^{th} iteration, the Schatten- p quasi-norm term will be replaced by its tangent plane, which is given as

$$(\sigma_k(\mathbf{T}_{\mathbf{u}, \Omega}^l) + \varepsilon)^p + \frac{p}{(\sigma_k(\mathbf{T}_{\mathbf{u}, \Omega}^l) + \varepsilon)^{1-p}} (\sigma_k(\mathbf{T}_{\mathbf{u}, \Omega}) - \sigma_k(\mathbf{T}_{\mathbf{u}, \Omega}^l)), \quad (16)$$

where $\varepsilon > 0$ is a smoothing parameter. Notice that $\mathbf{T}_{\mathbf{u}, \Omega}^l$ is constant. Consequently, by applying equation (16) into Problem (15), the resultant optimization formula will be rewritten as

$$\begin{aligned} \min_{\mathbf{X}, \mathbf{u}, \{\sigma \geq 0\}} \quad & tr(\mathbf{X}) + \frac{1}{p} \|\mathbf{W}^l \mathbf{T}_{\mathbf{u}, \Omega}\|_* + tr(\tilde{\mathbf{R}}_\Omega^{-1} \text{diag}(\boldsymbol{\sigma}_\Omega)) \\ \text{subject to} \quad & \mathbf{T}(\mathbf{u}) \geq 0, \mathbf{X} \geq \tilde{\mathbf{R}}_\Omega^{(1/2)} (\mathbf{T}_{\mathbf{u}, \Omega} + \text{diag}(\boldsymbol{\sigma}_\Omega))^{-1} \tilde{\mathbf{R}}_\Omega^{(1/2)}. \end{aligned} \quad (17)$$

In Problem (17), the Schatten- p quasi-norm term is reduced to the weighted nuclear norm (WNN), where $\|\mathbf{W}^l \mathbf{T}_{\mathbf{u}, \Omega}\|_* = tr(\mathbf{W}^l \mathbf{T}_{\mathbf{u}, \Omega}) = \sum_{k=1}^{M+N-1} (w_k^l \sigma_k(\mathbf{T}_{\mathbf{u}, \Omega}))$ and $\mathbf{W}^l = \mathbf{D} \cdot \text{diag}(w_1^l, w_2^l, \dots, w_{N+M-1}^l) \cdot \mathbf{D}^H$ is the reweighting matrix in the l^{th} iteration, whose elements are $w_k^l = (p/(\sigma_k(\mathbf{T}_{\mathbf{u}, \Omega}^l) + \varepsilon))^{1-p}$. Therefore, we name our method as reweighted p -GLS (Rwp-GLS). The matrix $\mathbf{T}_{\mathbf{u}, \Omega}^l$ is updated based on the latest solution of the problem, and $\mathbf{D}$ is obtained by using the eigen decomposition of $\mathbf{T}_{\mathbf{u}, \Omega}^l$ so that $\mathbf{D} \mathbf{D}^H = \mathbf{I}$.

Since $\|\mathbf{T}_{\mathbf{u}, \Omega}\|_p^p$ is a concave function when $0 < p < 1$, then at each iteration its value decreases significantly, faster than the decrease of its tangent plane. It follows that, by iteratively solving Problem (17), the objective function monotonically decreases and converges to a local minimum. Following the same steps in deriving Problem (9), at each iteration of Rwp-GLS, we need to solve the following SDP:

$$\begin{aligned} \min_{\mathbf{X}, \mathbf{u}, \{\sigma \geq 0\}} \quad & tr(\mathbf{X}) + \frac{1}{p} \|\mathbf{W}^l \mathbf{T}_{\mathbf{u}, \Omega}\|_* + tr(\tilde{\mathbf{R}}_\Omega^{-1} \text{diag}(\boldsymbol{\sigma}_\Omega)) \\ \text{subject to} \quad & (\text{constraint of (9)}). \end{aligned} \quad (18)$$

By solving the SDP in Problem (18) by off-the-shelf SDP solvers, e.g., SDPT3 [40], $\hat{\mathbf{R}}_\Omega = \mathbf{T}(\hat{\mathbf{u}}) + \text{diag}(\hat{\boldsymbol{\sigma}}_\Omega)$ is obtained given the solution $(\hat{\mathbf{u}}, \hat{\boldsymbol{\sigma}}_\Omega)$. Once $\hat{\mathbf{R}}_\Omega$ is obtained, the following task is to estimate the parameters $(\hat{\theta}, \hat{P})$. Note that Rwp-GLS does not assume any knowledge of K , which must be determined in order to acquire the final estimation step. Recall that $\text{rank}(\mathbf{T}(\hat{\mathbf{u}})) = K$; hence, K determination can be formulated as $\text{rank}(\mathbf{T}(\hat{\mathbf{u}}))$ estimation. In Rwp-GLS, based on the eigenvalue decomposition of $\mathbf{T}(\hat{\mathbf{u}})$, K is set as the number of eigenvalues that are greater than a predefined threshold. Considering the assumptions in Section 5.1, we acquire the parameters $(\hat{\theta}, \hat{P})$ from $\mathbf{T}(\hat{\mathbf{u}})$ uniquely by applying the Vandermonde decomposition lemma as $\mathbf{T}(\hat{\mathbf{u}}) = \mathbf{A}_\Omega(\hat{\theta}) \text{diag}(\hat{P}) \mathbf{A}_\Omega^H(\hat{\theta})$ (readers are referred to [25], for more information).

5.4. Convergence Analysis. As shown, Rwp-GLS method converges to a local minimum, here we analyze its convergence from other side. It is well known that WNN

minimization problem is nonconvex in general, and no efficient methods in theory can guarantee the global minimum. According to [41], WNN converges weakly to the solution if the weights are nondescendingly ordered, through a weighted soft-thresholding operator. By observing that $\{\sigma_k(\mathbf{T}_{\mathbf{u},\Omega})\}_{k=1}^{M+N-1}$ are nonnegative monotonically decreasing, then the ascending order of the resultant weights $\{w_k\}_{k=1}^{M+N-1}$ will be kept throughout the reweighting process. Hence, we can conclude that Rwp-GLS method converges weakly to the solution.

5.5. Computational Complexity. The overall computational complexity of Rwp-GLS has two dominated parts, the first one is related to the computation of the reweighted CF, which is an iterative process, and the second is the cost of $(\hat{\theta}, \hat{P})$ estimation. To make the following discussion more sense, let's set $\bar{M} = M + N - 1$, and $\bar{N} = MN$. According to the iterative part, the overall computational complexity equals the cost per iteration times the number of iterations. In each iteration, Rwp-GLS requires computing $\tilde{R}_\Omega^{(1/2)}$ and $\mathbf{W}^l$, where the numbers of required floating-point operations (flops) are $O(\bar{M}^2 L + \bar{M}^3)$ and $O(\bar{M}^3)$, respectively. Then, solving the SDP must be performed, where the interior-point based SDPT3 solver requires $O(\bar{M}^4 \bar{N}^{2.5} + \bar{N}^{4.5})$ flops [40], in the case of $(\bar{N} + \bar{M}^2)$ variable size, and $(\bar{N} + \bar{M})$ as the dimension of the PSD matrix. Finally, applying the Vandermonde decomposition to estimate $(\hat{\theta}, \hat{P})$ consumes up to $O(\bar{N}^3)$.

6. Simulations and Results

In this section, the performance of Rwp-GLS method is tested and compared with state-of-art methods via numerical simulations. The implementation of Rwp-GLS is done using MATLAB v.14. a by SDPT3 solver of CVX toolbox [42], on a PC with 2 GHz CPU.

6.1. Simulation Setups. All the simulations are carried out for two equal-power uncorrelated signals (i.e. $K = 2$), impinge onto 7-elements CPA with $M = 3$, and $N = 5$ with half-wavelength element spacing. Consider both signals and noise are i.i.d. Gaussian. The root mean squared error (RMSE) of the DoA estimation is computed as $\sqrt{(1/TK) \sum_{i=1}^T \sum_{k=1}^K (\theta_k - \hat{\theta}_{k,i})^2}$, where $\{\theta_k\}_{k=1}^K$ is the set of true directions, $\{\hat{\theta}_k\}_{k=1}^K$ indicates the set of estimated directions and T is the number of Monte Carlo runs. The SNR is defined for equal-power sources ($P_1 = P_2 = \dots = P_K = P$) as [28],

$$\text{SNR} = 10 \log_{10} \frac{P}{M + N - 1} \sum_{i=1}^{M+N-1} \sigma_{\Omega_i}^{-1}, \quad (19)$$

where $\{\sigma_{\Omega_i}\}_{i=1}^{M+N-1}$ are spatially uncorrelated white Gaussian random. Furthermore, to measure the variation of the noise on the antenna array, we use the worst noise power ratio (WNPR) as $\text{WNPR} = (\sigma_{\max}/\sigma_{\min})$ [28].

At first, we discuss the choice of ε and p parameters and study the influence of applying different values of both on

the performance of the proposed method. Since Problem (15) puts a penalty on $(1/p) \sum_{k=1}^K (\sigma_k(\mathbf{T}(\mathbf{u})) + \varepsilon)^p$, we plot the function $h_p(\sigma) = (1/p)(\sigma + \varepsilon)^p$ together with l_0 , and l_1 norms in Figure 1. As shown in Figure 1(a), $h_p(\sigma)$ converges faster toward l_0 whenever $p \rightarrow 0$ and yields to a sparser optimizer. While, by fixing p as small as 0.1 and using different values of ε , then $h_p(\sigma)$ will get close to l_0 norm as $\varepsilon \rightarrow \infty$, and approach l_1 norm when $\varepsilon \rightarrow 0$ as depicted in Figure 1(b).

However, after linearization, we expect that the resultant reweighting function in Problem (18) will play the same role as $h_p(\sigma)$ when $\varepsilon \rightarrow 0$ and $p \rightarrow 0$. Hence, throughout the iterative process, ε is initiated by 10^{-2} and reduced with a factor η in each iteration. Regarding the iterative process, the iteration is stopped if the relative change of $\hat{\mathbf{u}}$ at two consecutive iterations is less than 10^{-4} , or the maximum number of iterations which sets to 10 is reached. Finally, the results in each experiment are evaluated after 100 Monte Carlo runs.

6.2. DoA Estimation Performance Metrics. Consider two uncorrelated sources with spatial separation limit $\Delta\theta$ is defined as [23],

$$\Delta\theta = \min_{\theta_k, \theta_l, \theta_k \neq \theta_l} |\sin(\theta_k) - \sin(\theta_l)|. \quad (20)$$

Taking for example $\boldsymbol{\theta} = [17.46^\circ, 27.9^\circ]$ (i.e. $\Delta\theta = (2.52/NM)$), SNR = 0 dB with $\boldsymbol{\sigma}_\Omega = [1, 4, 7.3, 10, 8, 9, 5]$ (i.e. WNPR = 10) and $L = 200$ snapshots.

Initially, the success probability (SP) performance as a function of the number of iterations is examined. In Figure 2(a), we plot SP for two different values of $p \in \{0.1, 0.5\}$, while each is repeated for three values of factor $\eta \in \{2, 5, 10\}$. As shown, SP of Rwp-GLS is the best all over the number of iterations when $p = 0.1$ regardless of η value. In contrast to the case of $p = 0.5$, where SP slightly varies with η , and further it gets worse when the number of iteration becomes larger. However, using larger η is somewhat better. This indicates again the outperformance of Rwp-GLS with remarkable convergence characteristics for small values of p where the impact of η becomes negligible.

Figure 2(b) points to the influence of the reweighting function choice that is applied in the first iteration. In the case of $p = 0.1$, it is observed that SP is better when a nontrivial weighting based on the sample covariance matrix of the measurements is used rather than the weightless first iteration.

Further, we test the performance of Rwp-GLS method for different values of $0 < p < 1$ with respect to different settings of WNPR and SNR. Figure 3(a) shows that the RMSE of Rwp-GLS perform the best, over the range $1 \leq \text{WNPR} \leq 200$ for $0.1 \leq p \leq 0.2$, where SNR = 0 dB. In Figure 3(b), we fix WNPR = 10 and repeat the test for $-20 \leq \text{SNR (dB)} \leq 20$. The results show that the best performance is realized for $0.1 \leq p \leq 0.4$.

Regarding the results of the previous experiments, in the following simulations, we report the results considering the following settings: $p = 0.1$, $\varepsilon = 10^{-2}$, and $\eta = 10$, besides to

FIGURE 1: Property of $h_p(\sigma)$ when (a) $\varepsilon = 1$ and different values of p , (b) $p = 0.1$ and different values of ε .

FIGURE 2: Success probability performance versus number of iterations and various: (a) p and η , (b) reweighting function for $p = 0.1$ and $\eta = 10$.

FIGURE 3: RMSE of Rwp-GLS with respect to p and (a) $WNPR$ for $SNR = 0$ dB and (b) SNR for $WNPR = 10$.

adopt the weighted first iteration as the one tested in Figure 2(b).

6.3. Comparisons with Prior Arts. In this section, the DoA estimation performance of Rwp-GLS is evaluated and

compared with the conventional RootMUSIC method, which is applied directly to the aforementioned CPA. In addition, we consider two remarkable gridless CF methods: SPA (we compare with SPA because it uses the same gridless CF procedure as GLS and agrees with the comparative methods in the model order estimation technique, while

GLS adopts different technique) and ICMRA with Schatten- p penalty (ICMRA_{lp}) to perform this comparison. These methods are derived under similar assumptions on sources and noise and applied with their best settings. We fix WNPR = 10 and $L = 200$ all over the following simulations unless otherwise is mentioned. Moreover, to deliver a benchmark for the evaluation, the stochastic Cramer–Rao lower bound (CRLB) for nonuniform noise given in [30] is included.

Suppose two signals impinge onto the CPA from directions $\theta = [5.74^\circ, 15.55^\circ]$. By varying the SNR from -15 dB to 25 dB, Figure 4 shows that the performance of Rwp-GLS is the best, and it improves persistently with increasing SNR and follows the CRLB when $\text{SNR} \geq -5$ dB. Thus, Rwp-GLS can separate the two incident signals precisely when SNR is as low as -5 dB, in contrast to RootMUSIC, which suffers fixed error at high SNR due to the limited number of array elements.

Figure 5 presents the CPU times of the experiment associated with Figure 4. Rwp-GLS is an iterative method with significant convergence characteristics due to well initialization and its computational complexity in one iteration has a similar complexity as SPA; then, Rwp-GLS costs a similar CPU time to SPA and differs in weighting computing. Note that the time usages of both Rwp-GLS and SPA become slightly greater when increasing the SNR because of the σ_Ω estimation. It is worth mentioning, ICMRA_{lp} is an iterative method with convergence characteristics as Rwp-GLS, but without the ability to estimate the noise variances, which explains the slight decrease in CPU time-consuming.

In the following simulation, we evaluate the estimation performance of the Rwp-GLS versus the number of snapshots varying from 10 to 250. Consider the settings of $\theta = [5.74^\circ, 15.55^\circ]$, $\text{SNR} = 0$ dB, and $\sigma_\Omega = [1, 8, 2, 10, 3, 6, 5]$. Figure 6 shows that the estimation performance of Rwp-GLS outperforms the other methods and it follows well the CRLB as $L \geq 50$. In contrast to ICMRA_{lp}, whose performance degrades significantly when $L \leq 150$.

Next, we examine the performance of DoA estimation versus angle separation. Assume two signals impinge onto the CPA from directions $[1^\circ, 1^\circ + \Delta\theta]$. The RMSEs of the comparative methods are shown in Figure 7, with $\Delta\theta$ varying from 0.5° to 5° , $\text{SNR} = 15$ dB, and $\sigma_\Omega = [1, 8, 2, 10, 3, 6, 5]$. It can be observed that Rwp-GLS surpasses the other methods in localizing spatially adjacent signals. In particular, Rwp-GLS is able to coincide with the CRLB when the sources are only 1° apart.

In Figure 8, the RMSEs of DoA estimation using the aforementioned methods are plotted against the variation of correlation factor $0.1 \leq \rho \leq 0.9$. Furthermore, we carry out this simulation for two closely spaced sources, for example, let $\theta = [5.74^\circ, 11.54^\circ]$ (i.e., $\Delta\theta = (1.5/NM)$), where, $\text{SNR} = 0$ dB and $\sigma_\Omega = [3.5, 5, 1, 8, 10, 4.5, 2.2]$. In Figure 8, we observe that, as the correlation factor increases the performances of the compared methods degrade, in particular, ICMRA_{lp} is damaged significantly. However, the Rwp-GLS method has the best performance and exhibits immunity to the incorporated correlations among sources.

FIGURE 4: RMSEs of RootMUSIC, SPA, ICMRA_{lp}, Rwp-GLS, and CRLB with SNR.

FIGURE 5: CPU times comparisons of RootMUSIC, SPA, ICMRA_{lp}, and Rwp-GLS.

FIGURE 6: RMSEs of RootMUSIC, SPA, ICMRA_{lp}, Rwp-GLS, and CRLB with number of snapshots L .

FIGURE 7: RMSEs of RootMUSIC, SPA, ICMRA_{Ip}, Rwp-GLS, and CRLB with sources separation.

FIGURE 8: RMSEs of RootMUSIC, SPA, ICMRA_{Ip}, and Rwp-GLS with different correlation factors.

To further validate the robustness of the Rwp-GLS against the nonuniform noise, we present the RMSEs versus the WNPR in Figure 9, where WNPR is varied from 20 to 200, under the settings of SNR = 0 dB and $\theta = [5.74^\circ, 11.54^\circ]$ for two uncorrelated signals. It can be seen that, as WNPR increments, Rwp-GLS has the lowest RMSE over the studied WNPR range. This indicates again that the proposed method is robust against the nonuniform noise.

6.4. Power Estimation. Finally, we attempt to estimate the DoAs and powers of 7 equal-power signals impinging onto the CPA from directions $\theta = [2^\circ, 8^\circ, 19^\circ, 29^\circ, 41^\circ, 55^\circ, 77^\circ]$ simultaneously. SNR is set as 10 dB and $\sigma_n = [1, 4, 3, 8, 5, 7, 10]$ (i.e., WNPR = 10). We carry out 500 independent trials and show the results in Figure 10. Black circles denote the positions of the true powers and corresponding DoAs, while the colored dots denote the estimated

FIGURE 9: RMSEs of RootMUSIC, SPA, ICMRA_{Ip}, and Rwp-GLS with WNPR.

FIGURE 10: DoAs and power estimation of Rwp-GLS for $\theta = [2^\circ, 8^\circ, 19^\circ, 29^\circ, 41^\circ, 55^\circ, 77^\circ]$.

ones. As shown, the Rwp-GLS method can correctly locate both DoAs and powers for all sources.

7. Conclusions

In this paper, we have proposed a gridless sparse method for DoA estimation named Rwp-GLS. The new method reformulates the DoA estimation, by introducing the unified surrogate of the nonconvex Schatten- p quasi-norm to penalize the parametrized covariance matrix in the gridless CF criterion. Then, an iterative reweighted minimization method was derived with a SDP formulation. The effectiveness of Rwp-GLS is verified through numerical simulations under the consideration of CPA structure and unknown nonuniform noise background. The results show the superiority of Rwp-GLS in terms of resolution, applicability in a limited number of snapshots, and immunity against nonuniform noise and correlation of sources. In the future, we look forward to reasonable computational

efficiency implementation and arbitrary array geometries incorporation.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] A. Liu, D. Yang, S. Shi, Z. Zhu, and Y. Li, "Augmented subspace MUSIC method for DOA estimation using acoustic vector sensor array," *IET Radar, Sonar and Navigation*, vol. 13, no. 6, pp. 969–975, 2019.
- [2] Q. Xie, X. Pan, M. Huang, J. Chen, and S. Xiao, "Sparsity-based direction-of-departure and direction-of-arrival estimation for bistatic multiple-input multiple-output radar," *IEEE Access*, vol. 7, pp. 118826–118838, 2019.
- [3] Z. Yang, J. Li, P. Stoica, and L. Xie, "Sparse methods for direction-of-arrival estimation," in *Academic Press Library in Signal Processing*, vol. 7, pp. 509–581, Academic Press, Cambridge, MA, USA, 2017.
- [4] A. J. Barabell, "Improving the resolution performance of eigenstructure-based direction-finding algorithms," in *Proceedings of the ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing*, Boston, MA, USA, April 1983.
- [5] R. Roy and T. Kailath, "ESPRIT-Estimation of signal parameters via rotational invariance techniques," *Adaptive Antennas for Wireless Communications*, vol. 37, no. 7, pp. 984–995, 2009.
- [6] R. R. G. Baraniuk, "Compressive sensing," *IEEE Signal Processing Magazine*, vol. 24, no. 4, 2007.
- [7] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [8] D. H. Chae, P. Sadeghi, and R. A. Kennedy, "Effects of basis-mismatch in compressive sampling of continuous sinusoidal signals," in *Proceedings of the 2010 2nd International Conference on Future Computer and Communication*, Wuha, China, May 2010.
- [9] Y. Chi, L. L. Scharf, A. Pezeshki, and A. R. Calderbank, "Sensitivity to basis mismatch in compressed sensing," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 2182–2195, 2011.
- [10] Q. Liu, H. C. So, and Y. Gu, "Off-grid DOA estimation with nonconvex regularization via joint sparse representation," *Signal Processing*, vol. 140, pp. 171–176, 2017.
- [11] Z. Tan and A. Nehorai, "Sparse direction of arrival estimation using co-prime arrays with off-grid targets," *IEEE Signal Processing Letters*, vol. 21, no. 1, pp. 26–29, 2014.
- [12] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, "Compressed sensing off the grid," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7465–7490, 2013.
- [13] P. Pal and P. P. Vaidyanathan, "A grid-less approach to underdetermined direction of arrival estimation via low rank matrix denoising," *IEEE Signal Processing Letters*, vol. 21, no. 6, pp. 737–741, 2014.
- [14] Z. Yang and L. Xie, "On gridless sparse methods for multi-snapshot direction of arrival estimation," *Circuits, Systems, and Signal Processing*, vol. 36, no. 8, pp. 3370–3384, 2017.
- [15] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky, "The convex geometry of linear inverse problems," *Foundations of Computational Mathematics*, vol. 12, no. 6, pp. 805–849, 2012.
- [16] E. J. Candès and C. Fernandez-Granda, "Towards a mathematical theory of super-resolution," *Communications on Pure and Applied Mathematics*, vol. 67, no. 6, pp. 906–956, 2014.
- [17] C. Fernandez-Granda, "Super-resolution of point sources via convex programming," in *Proceedings of the 2015 IEEE 6th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, Cancun, Mexico, December 2015.
- [18] B. N. Bhaskar, G. Tang, and B. Recht, "Atomic norm denoising with applications to line spectral estimation," *IEEE Transactions on Signal Processing*, vol. 61, no. 23, pp. 5987–5999, 2013.
- [19] Z. Yang and L. Xie, "Continuous compressed sensing with a single or multiple measurement vectors," in *Proceedings of the IEEE Workshop on Statistical Signal Processing Proceedings*, Gold Coast, Australia, June–July 2014.
- [20] Z. Yang and L. Xie, "Enhancing sparsity and resolution via reweighted atomic norm minimization," *IEEE Transactions on Signal Processing*, vol. 64, no. 4, pp. 995–1006, 2016.
- [21] 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.
- [22] H. Li, P. Stoica, and J. Li, "Computationally efficient maximum likelihood estimation of structured covariance matrices," *IEEE Transactions on Signal Processing*, vol. 47, no. 5, pp. 1314–1323, 1999.
- [23] Z. Tan, Y. C. Eldar, and A. Nehorai, "Direction of arrival estimation using co-prime arrays: a super resolution viewpoint," *IEEE Transactions on Signal Processing*, vol. 62, no. 21, pp. 5565–5576, 2014.
- [24] 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.
- [25] Z. Yang, L. Xie, and C. Zhang, "A discretization-free sparse and parametric approach for linear array signal processing," *IEEE Transactions on Signal Processing*, vol. 62, no. 19, pp. 4959–4973, 2014.
- [26] X. Wu, W. P. Zhu, and J. Yan, "A Toeplitz covariance matrix reconstruction approach for direction-of-arrival estimation," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 9, pp. 8223–8237, 2017.
- [27] 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.
- [28] 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, no. 7, pp. 1310–1324, 2001.
- [29] B. Liao, S. C. Chan, L. Huang, and C. Guo, "Iterative methods for subspace and DOA estimation in nonuniform noise," *IEEE Transactions on Signal Processing*, vol. 64, no. 12, pp. 3008–3020, 2016.
- [30] Z.-Q. He, Z.-P. Shi, and L. Huang, "Covariance sparsity-aware DOA estimation for nonuniform noise," *Digital Signal Processing*, vol. 28, no. 1, pp. 75–81, 2014.
- [31] W. Tan and X. Feng, "Covariance matrix reconstruction for direction finding with nested arrays using iterative reweighted nuclear norm minimization," *International Journal of*

- Antennas and Propagation*, vol. 2019, Article ID 7657898, 13 pages, 2019.
- [32] C. Lu, J. Tang, S. Yan, and Z. Lin, "Generalized nonconvex nonsmooth low-rank minimization," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA, June 2014.
  - [33] M. Malek-Mohammadi, M. Babaie-Zadeh, and M. Skoglund, "Performance guarantees for Schatten-p quasi-norm minimization in recovery of low-rank matrices," *Signal Processing*, vol. 114pp. 225–230, C, 2015.
  - [34] L. Zien Alabideen, O. Al-Dakkak, and K. Khorzom, "Frequency recovery resolution enhancement using generalized reweighted nonconvex relaxation," *UPB Scientific Bulletin, Series C: Electrical Engineering*, vol. 81, no. 1, pp. 81–98, 2019.
  - [35] C. Zhou, Y. Gu, Y. D. Zhang, Z. Shi, T. Jin, and X. Wu, "Compressive sensing-based coprime array direction-of-arrival estimation," *IET Communications*, vol. 11, no. 11, pp. 7465–7490, 2017.
  - [36] P. Stoica, P. Babu, and J. Li, "SPICE: a sparse covariance-based estimation method for array processing," *IEEE Transactions on Signal Processing*, vol. 59, no. 2, pp. 629–638, 2011.
  - [37] M. Zhang, Z.-H. Huang, and Y. Zhang, "Restricted p-Isometry properties of nonconvex matrix recovery," *IEEE Transactions on Information Theory*, vol. 59, no. 7, pp. 4316–4323, 2013.
  - [38] C. Xu, Z. Lin, and H. Zha, "A unified convex surrogate for the Schatten-p norm," in *Proceedings of the 31st AAAI Conference on Artificial Intelligence*, San Francisco, CA, USA, February 2017.
  - [39] M. Fazel, H. Hindi, and S. P. Boyd, "Log-det heuristic for matrix rank minimization with applications to Hankel and Euclidean distance matrices," in *Proceedings of the American Control Conference*, Denver, CO, USA, June 2003.
  - [40] K. C. Toh, M. J. Todd, and R. H. Tütüncü, "SDPT3—a MATLAB software package for semidefinite programming, version 1.3," *Optimization Methods and Software*, vol. 11, no. 1–4, pp. 545–581, 1999.
  - [41] S. Gu, Q. Xie, D. Meng, W. Zuo, X. Feng, and L. Zhang, "Weighted nuclear norm minimization and its applications to low level vision," *International Journal of Computer Vision*, vol. 121, no. 2, pp. 183–208, 2017.
  - [42] M. Grant and S. Boyd, *CVX: Matlab Software for Disciplined Convex Programming*, CVX Research, Inc., Austin, TX, USA, 2017.