IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

SPECIAL SECTION ON THEORY, ALGORITHMS, AND APPLICATIONS OF SPARSE RECOVERY

Received September 15, 2018, accepted October 1, 2018, date of publication October 23, 2018, date of current version November 30, 2018.

*Digital Object Identifier 10.1109/ACCESS.2018.2877727*

# Robust Sparse Bayesian Learning for Off-Grid DOA Estimation With Non-Uniform Noise

**HUAFEI WANG[1], XIANPENG WANG[1], (Member, IEEE), LIANGTIAN WAN[2], (Member, IEEE), AND MENGXING HUANG[1]**

[1]State Key Laboratory of Marine Resource Utilization in South China Sea, College of Information Science and Technology, Hainan University, Haikou 570228, China

[2]Key Laboratory for Ubiquitous Network and Service Software of Liaoning Province, School of Software, Dalian University of Technology, Dalian 116620, China

Corresponding authors: Xianpeng Wang (wxpeng1986@126.com) and Liangtian Wan (wanliangtian@dlut.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61701144 and Grant 61801076, in part by the Program of Hainan Association for Science and Technology Plans to Youth Research and Development Innovation under Grant QCXM201706, in part by the Scientific Research Projects of University in Hainan Province under Grant Hnky2018ZD-4, in part by the Major Science and Technology Project of Hainan Province under Grant ZDKJ2016015, in part by the Natural Science Foundation of Hainan Province under Grant 617024, and in part by the Scientific Research Setup Fund of Hainan University under Grant KYQD(ZR)1731.

**ABSTRACT** The performance of traditional sparse representation-based direction-of-arrival (DOA) estimation algorithm is substantially degraded in the presence of non-uniform noise and off-grid gap caused by the discretization processes. In this paper, a robust sparse Bayesian learning method is proposed for off-grid DOA estimation with non-uniform noise. In the proposed method, the covariance matrix of non-uniform noise is reconstructed by a modified inverse iteration method. Then, the discrete sampling grid points in the spatial domain are treated as dynamic parameters, and the expectation–maximization algorithm is used to iteratively refine the position of the discretization grid points. This refinement procedure is implemented by solving a polynomial. The simulation results indicate that the proposed method can maintain excellent DOA estimation performance with uniform or non-uniform noise. Furthermore, it can also achieve satisfactory performance under a coarse grid condition.

**INDEX TERMS** Array signal processing, direction-of-arrival estimation, non-uniform noise, off-grid, sparse Bayesian learning.

## I. INTRODUCTION

Direction-of-arrival (DOA) estimation of impinging signals has always played an important role in the field of array signal processing, and it has attracted a considerable attention for its widely application in mobile communication, radar and sonar, etc. [1]. Moreover, the application of DOA estimation for target location and tracking [2]–[4] has also attracted particular research attention. A large number of methods, for example, the subspace-based method [5], [6] and maximum likelihood (ML) method [7]–[9], have been proposed in the past few decades to provide solutions for DOA estimation. The subspace-based algorithms, such as multiple signal classification (MUSIC) [5] and estimation of signal parameters via rotational invariance techniques (ESPRIT) [6], are the most mature classic algorithms. But the performance of these subspace-based algorithms is usually limited by snapshot number and signal-to-noise ratio (SNR) due to its dependence on eigenvalue decomposition of covariance matrix. To deal with these issues, a large number of sparse signal reconstruction (SSR) based DOA estimation methods [10]–[14] have been proposed, including sparse Bayesian learning (SBL) [10], $l_1$-norm based singular value decomposition ($l_1$-SVD) [11] and the variants of $l_1$-SVD [12]–[14]. Because these SSR-based methods are less sensitive to SNR and snapshot number, their performance is superior to that of subspace-based methods with lower SNR or/and limited number of snapshots. This merit results in the SSR-based method widely used in DOA estimation for MIMO radar [15]–[17]. In these SSR-based methods, the performance of $l_1$-SVD method is limited due to its use of $l_1$-norm optimization. However, the SBL method, which has less error and higher estimation precision [18], has become one of the most concerned SSR-based DOA estimation method.

But it cannot be ignored that the satisfactory performance of all these methods mentioned above is based on the assumption that the noise is uniform white Gaussian noise.

However, in practice, this assumption is hard to accomplish due to the existing of the non-uniform sensor response and the non-ideality receiving channel [19], [20]. Therefore, the actual noise is usually non-uniform white Gaussian noise. In order to cope with the DOA estimation in the presence of non-uniform noise, many ML-based methods [21]–[23] have been proposed. On the other hand, a new stochastic ML algorithm is proposed in [24] to achieve high precision DOA estimation where the covariance matrix of non-uniform noise is estimated by a modified inverse iteration algorithm. Although these ML-based methods can effectively deal with non-uniform noise, the requirement of joint search on all possible directions results in high computation complexity, which may significantly restrict their practical application. By taking the advantages of SSR technique into consideration, many SSR-based methods [25]–[28] have been introduced to deal with non-uniform noise. An improved SBL method [25], which adopts the modified inverse iteration algorithm [24] to reconstruct noise covariance matrix, is reported to achieve high precision DOA estimation in the presence of non-uniform noise. In addition, some DOA estimation algorithms based on array covariance matrix are proposed [26], [27] to eliminate the influence of non-uniform noise by utilizing the second-order statistical characteristics of the received data. High DOA estimation accuracy can also be achieved in [27] by using an adaptive procedure [11]. However, its performance is restricted by $l_1$-norm optimization and array aperture loss. In order to remove this restriction, a SBL method [28] with the variance of non-uniform noise is estimated by using the least squares (LS) criterion is proposed.

On the other hand, it should be noted that the spatial sparsity of signal is obtained by the discretization grid. SBL based methods can achieve high-precision DOA estimation only when sampling grid is dense enough and all true DOAs are located exactly at the grid points [11], [29]. Unfortunately, it is unrealistic to achieve that all true DOAs fall on the grid points, and there must exist an off-grid error between the true DOA and the grid point closest to it. To deal with the off-grid error problem, some algorithms are proposed [30]–[33] to effectively improve performance. In [31], a sparse Bayesian inference (SBI) method is proposed, in which the off-grid problem is dealt with linear approximation, to achieve high estimation accuracy with the coarse sampling grid. By considering the influence of noise covariance matrix, an improved SBL based DOA estimation method [32] is proposed based on the covariance matrix of received signals. These two methods reported in [31] and [32] achieve high estimation accuracy with high computational complexity. Hence, a root SBL algorithm for off-grid DOA estimation is proposed in [33] to achieve high estimation accuracy with much lower computational complexity. Although the method in [33] balances the computational complexity and estimation accuracy, it ignores the influence of the non-uniform noise. In summary, almost all SSR-based methods consider either the problems of off-grid or nonuniform noise. The DOA estimation method with the coexistence of off-grid error and non-uniform noise is rare. Thus, we focus on solving the DOA estimation problem with the coexistence of off-grid error and non-uniform noise.

In this paper, a robust sparse Bayesian learning method for off-grid DOA estimation with non-uniform noise is proposed to minimize both the off-grid error and the influence of non-uniform noise. Firstly, the signal power of source signal is estimated according to the SBL strategy. Then the modified inverse iteration [24] is adopted to reconstruct the covariance matrix of unknown non-uniform noise. The grid points are refined by solving a polynomial to treat the grid points as the dynamic parameters. Thus, the proposed method can achieve high accuracy off-grid DOA estimation in the presence of unknown non-uniform noise.

*Notation:* The capital and lowercase italic bold letters denote matrices and column vectors, respectively. $(\cdot)^{-1}$, $(\cdot)^H$, $(\cdot)^*$ and $(\cdot)^T$ denote inverse, conjugate transpose, conjugate and transpose operations, respectively. $diag\{\cdot\}$ denotes the diagonalization operation. $\mathrm{tr}(\cdot)$ represents the trace of matrix and $\mathbb{C}^{M\times N}$ denotes a $M \times N$ complex matrix set. $|\cdot|$, $\|\cdot\|_2$ and $\|\cdot\|_F$ stand for the determinant of a square matrix or absolute value, $l_2$ norm and Frobenius norm, respectively. $\mathrm{E}\{\cdot\}$ stands for the mathematical expectation operator.

## II. SPARSE SIGNAL MODEL

Consider a uniform linear array (ULA) consisting of $M$ antennas. The distance between adjacent antennas is $d = \lambda/2$, where $\lambda$ represents the wavelength of signal source. There are $K$ far-field narrow-band signals impinging on the ULA from the direction of $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_K]$, where $\theta_k (k = 1, 2, \ldots, K)$ represents the DOA of the $k$th signal. Then the signal received by the array at the $t$th snapshot can be expressed as [34]

$$\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{s}(t) + \boldsymbol{n}(t), \tag{1}$$

where $\boldsymbol{x}(t) = [x_1(t), x_2(t), \ldots, x_M(t)]^T \in \mathbb{C}^{M\times 1}$, and $x_m(t)$ represents the signal received by the $m$th antennas. $\boldsymbol{A} = [\boldsymbol{a}(\theta_1), \boldsymbol{a}(\theta_2), \ldots, \boldsymbol{a}(\theta_K)] \in \mathbb{C}^{M\times K}$ is steering matrix with $\boldsymbol{a}(\theta_k) = [1, e^{-j2\pi d/\lambda \sin\theta_k}, \ldots, e^{-j2\pi(M-1)d/\lambda \sin\theta_k}]^T \in \mathbb{C}^{M\times 1}$. $\boldsymbol{s}(t) = [s_1(t), s_2(t), \ldots, s_K(t)]^T \in \mathbb{C}^{K\times 1}$ and $\boldsymbol{n}(t) = [n_1(t), n_2(t), \ldots, n_M(t)]^T \in \mathbb{C}^{M\times 1}$ denote the signal source vector and unknown Gaussian white noise vector, respectively. The covariance matrix of $\boldsymbol{n}(t)$ is denoted as

$$\boldsymbol{Q} = \mathrm{E}\{\boldsymbol{n}(t)\boldsymbol{n}(t)^H\} = diag\{\sigma_1^2, \sigma_2^2, \ldots, \sigma_M^2\}, \tag{2}$$

where $\sigma_m^2$ represents the power of the noise received on the $m$th antenna. It should be noticed that if $\boldsymbol{n}(t)$ is unknown non-uniform noise, then $\sigma_m^2$ satisfies $\sigma_1^2 \neq \sigma_2^2 \neq \cdots \neq \sigma_M^2$ and otherwise, $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_M^2$.

In order to achieve the DOA estimation by using sparse representation, the sparse signal model should be constructed firstly. The spacial domain of a range $[-\pi/2, \pi/2]$ can be sampled uniformly to obtain a complete direction vector $\bar{\boldsymbol{\theta}} = [\bar{\theta}_1, \bar{\theta}_2, \ldots, \bar{\theta}_{\bar{K}}]$, where $\bar{K}$, $\bar{K} \gg M > K$, is the number of samples. If the sampling grid is dense enough and all the true

DOAs are on the grid points, the signal model in Eq. (1) can be transformed into a sparse signal model as

$$\boldsymbol{x}(t) = \bar{\boldsymbol{A}}\bar{\boldsymbol{s}}(t) + \boldsymbol{n}(t), \tag{3}$$

where $\bar{\boldsymbol{A}} = [\boldsymbol{a}(\bar{\theta}_1), \boldsymbol{a}(\bar{\theta}_2), \ldots, \boldsymbol{a}(\bar{\theta}_{\bar{K}})]$ is a $M \times \bar{K}$ matrix, known as an overcomplete dictionary and $\boldsymbol{a}(\bar{\theta}_{\bar{k}}) = [1, e^{-j2\pi d/\lambda \sin\bar{\theta}_{\bar{k}}}, \ldots, e^{-j2\pi(M-1)d/\lambda \sin\bar{\theta}_{\bar{k}}}]^T$ with $\bar{k} = 1, 2, \ldots, \bar{K}$. $\bar{\boldsymbol{s}}(t) = [\bar{s}_1(t), \bar{s}_2(t), \ldots, \bar{s}_{\bar{K}}(t)]^T$ is a $\bar{K} \times 1$ dimensional $K$ sparse vector. $\boldsymbol{n}(t)$ is an unknown Gaussian white noise. Collect $T$ snapshots of received signal, the sparse signal model can be expressed in a matrix form

$$\boldsymbol{X} = \bar{\boldsymbol{A}}\bar{\boldsymbol{S}} + \boldsymbol{N}, \tag{4}$$

where $\boldsymbol{X} = [\boldsymbol{x}(1), \boldsymbol{x}(2), \ldots, \boldsymbol{x}(T)]$, $\bar{\boldsymbol{S}} = [\bar{\boldsymbol{s}}(1), \bar{\boldsymbol{s}}(2), \ldots, \bar{\boldsymbol{s}}(T)]$ and $\boldsymbol{N} = [\boldsymbol{n}(1), \boldsymbol{n}(2), \ldots, \boldsymbol{n}(T)]$. By estimating the parameters of each row of $\bar{\boldsymbol{S}}$, the DOA of the target signal can be obtained.

In practice, it is unrealistic that all true DOAs are located exactly on the grid points. The gap between the grid points and the true DOAs inevitably leads to estimation error, as shown in Fig. 1. In addition, the non-uniform noise $\boldsymbol{N}$ further degrades the estimation performance.

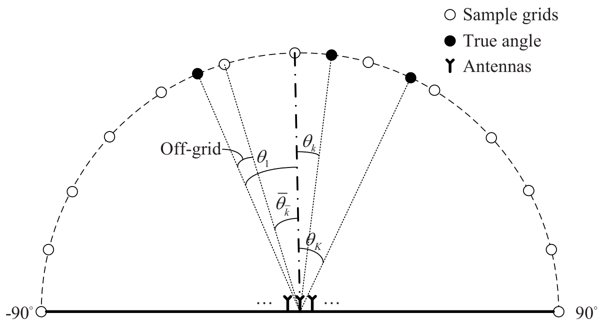



**FIGURE 1.** Schematic diagram of the off-grid sparse model.

## III. DOA ESTIMATION WITH NON-UNIFORM NOISE

In this section, a robust SBL based DOA estimation method is proposed to effectively reduce the effect of off-grid and non-uniform noise. Firstly, a modified inverse iteration method is adopted to reconstruct the noise covariance matrix to eliminate the influence of non-uniform noise. Then by regarding the sampling grid points as the dynamic parameters, an EM algorithm with polynomial root is proposed to refine the grid points.

### A. SPARSE BAYESIAN DERIVATION

According to the SBL strategy, we suppose that $\boldsymbol{s}(t)$ follows a complex Gaussian distribution

$$\bar{\boldsymbol{s}}(t) \sim \mathcal{CN}(0, \boldsymbol{\Delta}), \tag{5}$$

where $t = 1, 2, \ldots, T$ and $\mathcal{CN}(0, \boldsymbol{\Delta})$ denotes a complex Gaussian distribution with zero mean and its variance is $\boldsymbol{\Delta} = diag(\boldsymbol{\delta})$ where $\boldsymbol{\delta} = [\delta_1, \delta_2, \ldots, \delta_{\bar{K}}]^T$ is a hyper-parameters set, and $\delta_{\bar{k}}$ represents the power of signal source from $\bar{\theta}_{\bar{k}}$ ($\bar{k} = 1, 2, \ldots, \bar{K}$). $\delta_{\bar{k}}$ is non-zero only if there is a signal source on $\bar{\theta}_{\bar{k}}$. The prior probability density function of $\bar{\boldsymbol{s}}(t)$ is

$$p(\bar{\boldsymbol{s}}(t)|\boldsymbol{\delta}) = \prod_{\bar{k}=1}^{\bar{K}} (\pi\delta_{\bar{k}})^{-1} \exp\left\{-\frac{(\bar{s}_{\bar{k}}(t))^2}{\delta_{\bar{k}}}\right\}. \tag{6}$$

Hence, we have

$$p(\bar{\boldsymbol{S}}|\boldsymbol{\delta}) = \prod_{t=1}^{T} \mathcal{CN}(\bar{\boldsymbol{s}}(t)|0, \boldsymbol{\Delta}). \tag{7}$$

However, the sparsity of $\bar{\boldsymbol{s}}(t)$ cannot be shown since Eq. (7) contains the unknown hyper-parameter $\boldsymbol{\delta}$. In order to obtain a hierarchical structure, we further assume that $\boldsymbol{\delta}$ obeys the independent Gamma distribution as

$$p(\boldsymbol{\delta}) = \prod_{\bar{k}=1}^{\bar{K}} \text{Gamma}(\delta_{\bar{k}}|a, b), \tag{8}$$

where

$$\text{Gamma}(z|a, b) = \Gamma(z)^{-1} b^a z^{a-1} e^{-bz}, \tag{9}$$

and $\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt$. In order to obtain a broad hyper-prior, $a, b$ is usually set to be the number close to zero, i.e., $a, b \to 0$ [10].

Based on the above assumption and the sparse signal model of Eq. (4), the probability density distribution of the received signal $\boldsymbol{X}$ can be expressed as

$$\begin{aligned} p(\boldsymbol{X}|\bar{\boldsymbol{S}}, \boldsymbol{Q}) &= \mathcal{CN}(\boldsymbol{X}|\bar{\boldsymbol{A}}\bar{\boldsymbol{S}}, \boldsymbol{Q}) = \prod_{t=1}^{T} \mathcal{CN}(\boldsymbol{x}(t)|\bar{\boldsymbol{A}}\bar{\boldsymbol{s}}(t), \boldsymbol{Q}) \\ &= |\pi\boldsymbol{Q}|^{-T} \exp\{-\text{tr}[(\boldsymbol{X} - \bar{\boldsymbol{A}}\bar{\boldsymbol{S}})^H \boldsymbol{Q}^{-1}(\boldsymbol{X} - \bar{\boldsymbol{A}}\bar{\boldsymbol{S}})]\}. \end{aligned} \tag{10}$$

By utilizing the Bayesian theory, the posterior probability density distribution of $\bar{\boldsymbol{S}}$ with respect to $\boldsymbol{X}$ can be derivated as

$$\begin{aligned} p(\bar{\boldsymbol{S}}|\boldsymbol{X}; \boldsymbol{\delta}, \boldsymbol{Q}) &= \frac{p(\boldsymbol{X}|\bar{\boldsymbol{S}}; \boldsymbol{Q})p(\bar{\boldsymbol{S}}|\boldsymbol{\delta})}{\int p(\boldsymbol{X}|\bar{\boldsymbol{S}}; \boldsymbol{Q})p(\bar{\boldsymbol{S}}|\boldsymbol{\delta})d\bar{\boldsymbol{S}}} \\ &= |\pi\boldsymbol{\Sigma}|^{-T} \exp\{-\text{tr}[(\bar{\boldsymbol{S}} - \boldsymbol{\mu})^H \boldsymbol{\Sigma}^{-1}(\bar{\boldsymbol{S}} - \boldsymbol{\mu})]\}, \end{aligned} \tag{11}$$

where

$$\boldsymbol{\mu} = \boldsymbol{\Delta}\bar{\boldsymbol{A}}^H \boldsymbol{\Sigma}_X^{-1} \boldsymbol{X}, \tag{12}$$

$$\boldsymbol{\Sigma} = \boldsymbol{\Delta} - \boldsymbol{\Delta}\bar{\boldsymbol{A}}^H \boldsymbol{\Sigma}_X^{-1} \bar{\boldsymbol{A}}\boldsymbol{\Delta}, \tag{13}$$

$$\boldsymbol{\Sigma}_X = \boldsymbol{Q} + \bar{\boldsymbol{A}}\boldsymbol{\Delta}\bar{\boldsymbol{A}}^H. \tag{14}$$

Since the parameters $\boldsymbol{\Delta}$ and $\boldsymbol{Q}$ can be determined, the posterior estimated value of signal amplitude can be achieved based on Eq. (12). Then the posterior probability density distribution of $\boldsymbol{X}$ with respect to $\boldsymbol{\delta}$ can be expressed as

$$\begin{aligned} p(\boldsymbol{X}|\boldsymbol{\delta}, \boldsymbol{Q}) &= \int p(\boldsymbol{X}|\bar{\boldsymbol{S}}, \boldsymbol{Q})p(\bar{\boldsymbol{S}}|\boldsymbol{\delta})d\bar{\boldsymbol{S}} \\ &= |\pi\boldsymbol{\Sigma}_X|^{-T} \exp\{-\text{tr}(\boldsymbol{X}^H \boldsymbol{\Sigma}_X^{-1} \boldsymbol{X})\}. \end{aligned} \tag{15}$$

Ignoring the constant term and take the logarithm of Eq. (15), the objective function for optimizing $\boldsymbol{\delta}$ can be expressed as

$$\mathcal{L}(\boldsymbol{\delta}, \boldsymbol{Q}) = g(\hat{\boldsymbol{R}}, \boldsymbol{\Sigma}_X) = -T(\ln|\boldsymbol{\Sigma}_X| + \mathrm{tr}(\boldsymbol{\Sigma}_X^{-1}\boldsymbol{R})), \quad (16)$$

where $\boldsymbol{R} = \mathrm{E}[\boldsymbol{x}(t)\boldsymbol{x}^H(t)]$. Since $\boldsymbol{R}$ is an ideal value that is unrealistic to be obtained, it is approximated by $\hat{\boldsymbol{R}} = \frac{1}{T}\boldsymbol{X}\boldsymbol{X}^H$. By removing the minus sign, the objective function can be rewritten as

$$\mathcal{L}(\boldsymbol{\delta}, \boldsymbol{Q}) = \ln|\boldsymbol{\Sigma}_X| + \mathrm{tr}(\boldsymbol{\Sigma}_X^{-1}\hat{\boldsymbol{R}}). \quad (17)$$

The hyperparameter $\boldsymbol{\delta}$ can be estimated by maximizing Eq. (16) or minimizing the objective function Eq. (17), and the power of the incident signal on the spatial discrete angle set can also be determined. On the other hand, the noise covariance matrix $\boldsymbol{Q}$ can also be determined by optimizing Eq. (17). The following subsection will focus on the estimation of $\boldsymbol{\delta}$ and $\boldsymbol{Q}$. When $\boldsymbol{\delta}$ and $\boldsymbol{Q}$ are estimated, a 1-D search is performed on the power spectrum based on the relationship between $\boldsymbol{\delta}$ and $\bar{\boldsymbol{\theta}}$ to find the $K$ maximum values, which can be used for DOA estimation.

### B. SIGNAL POWER AND NOISE COVARIANCE MATRIX ESTIMATION

In this subsection, the signal power $\boldsymbol{\delta}$ is firstly estimated. Take the partial derivative of the objective function in Eq. (17) with respect to $\boldsymbol{\delta}$ and set it to be zero, we have

$$\frac{\partial \mathcal{L}(\boldsymbol{\delta}, \boldsymbol{Q})}{\partial \boldsymbol{\delta}} = 0. \quad (18)$$

After some mathematical manipulation, the iterative formula of $\boldsymbol{\delta}$ is obtained as

$$\delta_{\bar{k}}^{(p)} = \frac{1}{T}\|(\boldsymbol{\mu}^{(p)})_{\bar{k}\cdot}\|_2^2 + (\boldsymbol{\Sigma}^{(p)})_{\bar{k},\bar{k}}, \quad (19)$$

where $\bar{k} = 1, 2 \ldots, \bar{K}$, $(\cdot)^{(p)}$ represents the $p$th iteration and $\delta_{\bar{k}}^{(p)}$ denotes the estimated value of $\delta_{\bar{k}}$ in the $p$th iteration. $\boldsymbol{\mu}^{(p)}$ and $\boldsymbol{\Sigma}^{(p)}$ are the estimated value of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ in the $p$th iteration, respectively, which can be calculated from Eq. (12) and Eq. (13). $(\boldsymbol{\mu}^{(p)})_{\bar{k}\cdot}$ denotes the $\bar{k}$ row of $\boldsymbol{\mu}^{(p)}$, and $(\boldsymbol{\Sigma}^{(p)})_{\bar{k},\bar{k}}$ denotes the $(\bar{k}, \bar{k})$th element of $\boldsymbol{\Sigma}^{(p)}$. The calculation of Eq. (19) may be abnormal if most elements of $\boldsymbol{\delta}$ are zero during the convergence process. To avoid this situation, Eq. (19) is revised as

$$\delta_{\bar{k}}^{(p)} = \frac{1}{T}\|(\boldsymbol{\mu}^{(p)})_{\bar{k}\cdot}\|_2^2 / \left[1 - \frac{(\boldsymbol{\Sigma}^{(p)})_{\bar{k},\bar{k}}}{\delta_{\bar{k}}^{(p)}}\right] + \tau, \quad (20)$$

where $\tau$ is a very small positive number and is set as $\tau = 10^{-10}$ in this paper [18].

The above estimation of $\boldsymbol{\delta}$ is based on the assumption that the variance of noise (i.e. $\boldsymbol{Q}$) is known. For estimating the variance of non-uniform noise $\boldsymbol{Q}$, it is impossible to optimize the objective function by taking the partial derivative of Eq. (17) directly [18]. Therefore, an inverse iteration procedure [24] is adopted to optimize the objective function Eq. (17) to achieve the estimation of $\boldsymbol{Q}$.

According to the similar operation in [24], we can maximize Eq. (16) to obtain an initial estimate of $\boldsymbol{Q}$ and initialize $\boldsymbol{\theta}$ and $\boldsymbol{\delta}$. Then the gradient vector of $\mathcal{L}(\boldsymbol{\delta}, \boldsymbol{Q})$ in Eq. (16) can be calculated as

$$\begin{aligned}
[\nabla_{\boldsymbol{q}}\mathcal{L}(\boldsymbol{\delta}, \boldsymbol{Q})]_m &= -T\left\{\frac{\partial \ln|\boldsymbol{\Sigma}_X|}{\partial q_m} + \frac{\partial \mathrm{tr}(\boldsymbol{\Sigma}_X^{-1}\hat{\boldsymbol{R}})}{\partial q_m}\right\} \\
&= -T\{\mathrm{tr}(\boldsymbol{\Sigma}_X^{-1}\boldsymbol{E}_{m,m}) - \mathrm{tr}(\boldsymbol{\Sigma}_X^{-1}\hat{\boldsymbol{R}}\boldsymbol{\Sigma}_X^{-1}\boldsymbol{E}_{m,m})\} \\
&= -T\,\mathrm{tr}\{[\boldsymbol{\Sigma}_X^{-1} - \boldsymbol{\Sigma}_X^{-1}\hat{\boldsymbol{R}}\boldsymbol{\Sigma}_X^{-1}]\boldsymbol{E}_{m,m}\} \\
&= -T[\boldsymbol{\Sigma}_X^{-1} - \boldsymbol{\Sigma}_X^{-1}\hat{\boldsymbol{R}}\boldsymbol{\Sigma}_X^{-1}]_{m,m}, \quad (21)
\end{aligned}$$

where $[\nabla_{\boldsymbol{q}}\mathcal{L}(\boldsymbol{\delta}, \boldsymbol{Q})]_m$ denotes the $m$th entry of the gradient vector $\nabla_{\boldsymbol{q}}\mathcal{L}(\boldsymbol{\delta}, \boldsymbol{Q})$. $q_m$ represents the $m$th entry of $\boldsymbol{q} = \mathrm{diag}\{\boldsymbol{Q}\}$ and $\boldsymbol{E}_{i,j}$ is a $M \times M$ matrix with all entries are zero except that the $(i, j)$th entry is 1. Then let $\nabla_{\boldsymbol{q}}\mathcal{L}(\boldsymbol{\delta}, \boldsymbol{Q})$ be 0, we have

$$\mathrm{tr}\{[\boldsymbol{\Sigma}_X^{-1} - \boldsymbol{\Sigma}_X^{-1}\hat{\boldsymbol{R}}\boldsymbol{\Sigma}_X^{-1}]\boldsymbol{E}_{m,m}\} = 0, \quad (22)$$

or

$$[\boldsymbol{\Sigma}_X^{-1} - \boldsymbol{\Sigma}_X^{-1}\hat{\boldsymbol{R}}\boldsymbol{\Sigma}_X^{-1}]_{m,m} = 0, \quad (23)$$

where $m = 1, 2, \ldots, M$. Similar to the procedure of [24], we assume a $M \times M$ diagonal matrix $\hat{\boldsymbol{Q}}$, and $(\hat{\boldsymbol{R}} - \hat{\boldsymbol{Q}}, \hat{\boldsymbol{\Sigma}}_X)$ satisfies Eq. (23). Then we have

$$[\hat{\boldsymbol{\Sigma}}_X^{-1}(\hat{\boldsymbol{\Sigma}}_X - (\hat{\boldsymbol{R}} - \hat{\boldsymbol{Q}}))\hat{\boldsymbol{\Sigma}}_X^{-1}]_{m,m} = 0, \quad (24)$$

where $\hat{\boldsymbol{\Sigma}}_X$ is the initial estimated value of $\boldsymbol{\Sigma}_X$ from Eq. (14) when $\boldsymbol{Q}$ and $\boldsymbol{\Delta}$ are initialized. Similarly, $(\hat{\boldsymbol{R}} - \hat{\boldsymbol{Q}}, \hat{\boldsymbol{\Sigma}}_X)$ also satisfies Eq. (22). Hence, substituting $(\hat{\boldsymbol{R}} - \hat{\boldsymbol{Q}}, \hat{\boldsymbol{\Sigma}}_X)$ into Eq. (22) and after some mathematical manipulation, we have

$$\mathrm{tr}\{[\hat{\boldsymbol{\Sigma}}_X^{-1}\hat{\boldsymbol{R}}\hat{\boldsymbol{\Sigma}}_X^{-1} - \hat{\boldsymbol{\Sigma}}_X^{-1}(\hat{\boldsymbol{\Sigma}}_X + \hat{\boldsymbol{Q}})\hat{\boldsymbol{\Sigma}}_X^{-1}]\boldsymbol{E}_{m,m}\} = 0. \quad (25)$$

Obviously, Eq. (25) can be rewritten as a matrix form, which is shown as

$$\boldsymbol{U}\hat{\boldsymbol{q}} = \boldsymbol{v}, \quad (26)$$

where $\hat{\boldsymbol{q}}$ is denoted by $\hat{\boldsymbol{q}} = \mathrm{diag}(\hat{\boldsymbol{Q}})$, and

$$\begin{aligned}
[\boldsymbol{U}]_{i,j} &= \mathrm{tr}\{\boldsymbol{E}_{j,j}\hat{\boldsymbol{\Sigma}}_X^{-1}\boldsymbol{E}_{i,i}\hat{\boldsymbol{\Sigma}}_X^{-1}\}, \quad (27) \\
v_i &= \mathrm{tr}\{[\hat{\boldsymbol{\Sigma}}_X^{-1}\hat{\boldsymbol{R}}\hat{\boldsymbol{\Sigma}}_X^{-1} - \hat{\boldsymbol{\Sigma}}_X^{-1}]\boldsymbol{E}_{i,i}\}, \quad (28)
\end{aligned}$$

where $i, j = 1, 2, \ldots, M$. Then, the improving direction $\hat{\boldsymbol{q}}$, which has been proved in [24], can be obtained by solving the linear equation Eq. (26).

The signal power $\boldsymbol{\delta}$ and the covariance matrix of non-uniform noise $\boldsymbol{Q}$ are estimated from the above derivations according to the existing sparse signal model in Eq. (3), then the DOA estimation has already been obtainted. Although the DOA estimation can be realized from the above derivations, the gap between the true DOAs and grid points limits the performance of the DOA estimation. In the following subsection, the problem of off-grid will be dealt with by finding the root of a polynomial.

### C. REFINEMENT OF SAMPLED GRID POINTS

In this subsection, the expectation maximization (EM) algorithm is adopted to refine the grid points. Each iteration of EM algorithm contains E-step and M-step. In E-step, the mathematical expectation operation of the likelihood function is performed, and the M-step is to maximize this expectation. Let us consider the spatial discrete angle set $\bar{\boldsymbol{\theta}}$ as a parameter. According to the EM algorithm, we first perform the mathematical expectation operation to the likelihood function in Eq. (10). Then the objective function can be obtained by ignoring the independent constant terms, which is shown as

$$\begin{aligned}&\mathrm{E}_{p(\bar{\boldsymbol{S}}|\boldsymbol{X};\boldsymbol{\delta},\boldsymbol{Q})}\{\ln(p(\boldsymbol{X}|\bar{\boldsymbol{S}},\boldsymbol{Q}))\}\\&= -\sum_{t=1}^{T}\|\boldsymbol{Q}^{-\frac{1}{2}}(\boldsymbol{x}_t-\bar{\boldsymbol{A}}\boldsymbol{\mu}_t)\|_2^2 - T\,\mathrm{tr}\left((\boldsymbol{Q}^{-\frac{1}{2}}\bar{\boldsymbol{A}})\boldsymbol{\Sigma}(\boldsymbol{Q}^{-\frac{1}{2}}\bar{\boldsymbol{A}})^H\right),\end{aligned} \tag{29}$$

where $\boldsymbol{x}_t = \boldsymbol{x}(t)$ and $\boldsymbol{\mu}_t$ is the $t$th column of $\boldsymbol{\mu}$. To refine $\bar{\boldsymbol{\theta}}$, we maximize the objective function in Eq. (29). Define $\upsilon_{\bar{k}} = e^{-j2\pi d/\lambda \sin\bar{\theta}_{\bar{k}}}$ ($\bar{k} = 1, 2, \ldots, \bar{K}$) and consider the following equations

$$\begin{aligned}&\frac{\partial \sum_t \|\boldsymbol{Q}^{-\frac{1}{2}}(\boldsymbol{x}_t-\bar{\boldsymbol{A}}\boldsymbol{\mu}_t)\|_2^2}{\partial \upsilon_{\bar{k}}}\\&= (\boldsymbol{Q}^{-\frac{1}{2}}\bar{\boldsymbol{a}}'_{\bar{k}})^H\left((\boldsymbol{Q}^{-\frac{1}{2}}\bar{\boldsymbol{a}}_{\bar{k}})\sum_{t=1}^{T}|\boldsymbol{\mu}_{t,\bar{k}}|^2-\sum_{t=1}^{T}\boldsymbol{\mu}^*_{t,\bar{k}}(\boldsymbol{Q}^{-\frac{1}{2}}\boldsymbol{x})_{t-\bar{k}}\right),\end{aligned} \tag{30}$$

$$\begin{aligned}&\frac{\partial \mathrm{tr}((\boldsymbol{Q}^{-\frac{1}{2}}\bar{\boldsymbol{A}})\boldsymbol{\Sigma}(\boldsymbol{Q}^{-\frac{1}{2}}\bar{\boldsymbol{A}}^H))}{\partial \upsilon_{\bar{k}}}\\&= (\boldsymbol{Q}^{-\frac{1}{2}}\bar{\boldsymbol{a}}'_{\bar{k}})^H(\boldsymbol{Q}^{-\frac{1}{2}}\bar{\boldsymbol{A}})\boldsymbol{\varepsilon}_{\bar{k}}\\&= (\boldsymbol{Q}^{-\frac{1}{2}}\bar{\boldsymbol{a}}'_{\bar{k}})^H\left(\varepsilon_{\bar{k},\bar{k}}(\boldsymbol{Q}^{-\frac{1}{2}}\bar{\boldsymbol{a}}_{\bar{k}})+\sum_{i\neq\bar{k}}\varepsilon_{i,\bar{k}}(\boldsymbol{Q}^{-\frac{1}{2}}\bar{\boldsymbol{a}}_i)\right).\end{aligned} \tag{31}$$

Let the derivative of Eq. (29) with respect to $\upsilon_{\bar{k}}$ be 0

$$\begin{aligned}&(\ddot{\boldsymbol{a}}'_{\bar{k}})^H\left(\ddot{\boldsymbol{a}}_{\bar{k}}\sum_{t=1}^{T}(|\boldsymbol{\mu}_{t,\bar{k}}|^2+\varepsilon_{\bar{k},\bar{k}})\right.\\&\qquad\left.+T\sum_{i\neq\bar{k}}\varepsilon_{i,\bar{k}}\ddot{\boldsymbol{a}}_i-\sum_{t=1}^{T}\boldsymbol{\mu}^*_{t,\bar{k}}\ddot{\boldsymbol{x}}_{t-\bar{k}}\right)=0,\end{aligned} \tag{32}$$

where $\ddot{\boldsymbol{a}}_{\bar{k}} = \boldsymbol{Q}^{-\frac{1}{2}}\bar{\boldsymbol{a}}_{\bar{k}}$, $\bar{\boldsymbol{a}}_{\bar{k}}$ is the $\bar{k}$th column of $\bar{\boldsymbol{A}}$, $\ddot{\boldsymbol{a}}'_{\bar{k}} = d\ddot{\boldsymbol{a}}_{\bar{k}}/d\upsilon_{\bar{k}}$, $\ddot{\boldsymbol{x}}_{t-\bar{k}} = \ddot{\boldsymbol{x}}_t - \sum_{i\neq\bar{k}}\boldsymbol{\mu}_{t,\bar{k}}\ddot{\boldsymbol{a}}_i$ and $\ddot{\boldsymbol{x}}_t = \boldsymbol{Q}^{-\frac{1}{2}}\boldsymbol{x}_t$. $\boldsymbol{\mu}_{t,\bar{k}}$ and $\varepsilon_{i,\bar{k}}$ represent the $\bar{k}$th element and the $(i,\bar{k})$th element of $\boldsymbol{\mu}_t$ and $\boldsymbol{\Sigma}$, respectively.

In order to convert Eq. (32) into a polynomial form, we have the following definition

$$\Phi^{(\bar{k})} = \sum_{t=1}^{T}(|\boldsymbol{\mu}_{t,\bar{k}}|^2+\varepsilon_{\bar{k},\bar{k}}), \tag{33}$$

$$\Psi^{(\bar{k})} = T\sum_{i\neq\bar{k}}\varepsilon_{i,\bar{k}}\ddot{\boldsymbol{a}}_i-\sum_{t=1}^{T}\boldsymbol{\mu}^*_{t,\bar{k}}\ddot{\boldsymbol{x}}_{t-\bar{k}}. \tag{34}$$

Then Eq. (32) is rewritten as

$$[\upsilon_{\bar{k}}, 1, \upsilon_{\bar{k}}^{-1}, \ldots, \upsilon_{\bar{k}}^{-(M-2)}]\begin{bmatrix}\frac{M(M-1)}{2}\Phi^{(\bar{k})}\\ \Psi_2^{(\bar{k})}\\ 2\Psi_3^{(\bar{k})}\\ \vdots\\ (M-1)\Psi_M^{(\bar{k})}\end{bmatrix}=0, \tag{35}$$

where $\Psi_m^{(\bar{k})}$ is the $m$th element of $\Psi^{(\bar{k})}$. Obviously, after solving Eq. (35), we need to select only one root out of the $M-1$ roots to refine the grid point. According to the characteristics of $\upsilon_{\bar{k}}$, the root with the absolute value closest to 1 is selected to refine the grid point. With $\upsilon_{\bar{k}*}$ representing the selected root, the refined grid point is expressed as

$$\bar{\theta}_{\bar{k}*}^{REF} = \arcsin\left(-\frac{\lambda}{2\pi d}\cdot \mathrm{angle}(\upsilon_{\bar{k}*})\right). \tag{36}$$

In addition, it should be noticed that if the estimated DOA is almost lie on the original grid, the refined operation will inevitably increase the estimation error. To avoid this phenomenon, a further condition is set to determine whether the point needs to be refined. For example, we can use $(\bar{\theta}_{\bar{k}*-1}+\bar{\theta}_{\bar{k}*})/2 \leq \bar{\theta}_{\bar{k}*}^{REF} \leq (\bar{\theta}_{\bar{k}*}+\bar{\theta}_{\bar{k}*+1})/2$ to determine whether the grid refinement needs to be performed.

Until now, a robust SBL method for off-grid DOA estimation with non-uniform noise has been proposed. The algorithm can be implemented by the iteration procedure as given in algorithm 1.

**Algorithm 1** Robust SBL for Off-Grid DOA Estimation With Non-Uniform Noise

**Input:** The received signal $\boldsymbol{X}$;
**Initialization:** $\boldsymbol{Q}$, $\boldsymbol{\delta}$;
**while** $\sim$ Converge **do**
    Calculate $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ by Eq. (12) and Eq. (13);
    Update $\boldsymbol{\delta}$ according to Eq. (19);
    Calculate improving direction $\hat{\boldsymbol{q}}$ according to Eq. (26);
    **while** $g(\hat{\boldsymbol{R}}, \hat{\boldsymbol{\Sigma}}_X + t\hat{\boldsymbol{Q}}) < g(\hat{\boldsymbol{R}}, \hat{\boldsymbol{\Sigma}}_X) + \alpha t\nabla_{\boldsymbol{q}} g(\hat{\boldsymbol{R}}, \hat{\boldsymbol{\Sigma}}_X)^T\hat{\boldsymbol{q}}$ or $\boldsymbol{Q}+t\hat{\boldsymbol{Q}} < 0$ **do**
        $t = \beta t$;
    **end while**
    $\boldsymbol{Q} \leftarrow \boldsymbol{Q}+t\hat{\boldsymbol{Q}}$;
    Update $\bar{\boldsymbol{\theta}}$ according to Eq. (35) and Eq. (36).
**end while**
**Output:** $\boldsymbol{\mu}$ and $\bar{\boldsymbol{\theta}}$;
Perform a 1-D spectrum search on new $\bar{\boldsymbol{\theta}}$ to find the $K$ maximum values to achieve DOA estimation.

*Remark 1:* In this article, $\boldsymbol{Q}$ is initialized as $\boldsymbol{Q} = (\sigma^2)^{(0)}\boldsymbol{I}_M$, where $(\sigma^2)^{(0)} = \mathrm{tr}\{(\boldsymbol{I}_M - \bar{\boldsymbol{A}}\bar{\boldsymbol{A}}^H)\hat{\boldsymbol{R}}\}/(M-K)$, and $\boldsymbol{I}_M$ is an $M\times M$ identity matrix. $\boldsymbol{\delta}$ is initialized as $(\delta_{\bar{k}})^{(0)} = \frac{1}{T}\|(\boldsymbol{\mu}^{(0)})_{\bar{k}.}\|_2^2$, where $\boldsymbol{\mu}^{(0)} = \bar{\boldsymbol{A}}^H(\bar{\boldsymbol{A}}\bar{\boldsymbol{A}}^H)^{-1}\boldsymbol{X}$. And in algorithm 1, $\alpha$ and $\beta$ are two constants which satisfy $0 < \alpha < 0.5$ and $0 < \beta < 1$, respectively [24].

*Remark 2:* The main problems solved in this article are in step 6 and step 11. In order to achieve DOA estimation with the coexistence of non-uniform noise and off-grid, the problem of non-uniform noise and off-grid are handled iteratively in the proposed method. Firstly, in step 6, based on the specified structure of covariance matrix of non-uniform noise $\boldsymbol{Q}$, the modified inverse iteration algorithm [24] is utilized to accurately reconstruct the diagonal matrix $\boldsymbol{Q}$. In each iteration of this modified inverse iteration algorithm, the estimated $\boldsymbol{Q}$ becomes closer to the actual $\boldsymbol{Q}$. Then the sample grid point is treated as a parameter which is refined by solving a polynomial. The non-uniform noise covariance matrix can be accurately reconstructed, and the discrete grid points are iteratively refined, which enable our algorithm to achieve satisfactory performance with the coexisting of non-uniform noise and off-grid.

*Remark 3:* Our proposed method is mainly to solve the DOA estimation problem with the coexistence of off-grid error and non-uniform noise. When the noise is uniform white Gaussian noise, our method still works, but its performance may not be better than that of the OGSBI and root SBL, which has been proved in the subsequent section.

## IV. SIMULATION RESULTS

A number of simulations under different conditions are carried out in this section to verify the robustness of the proposed method for off-grid and non-uniform noise. The sparse iterative covariance-based estimation (SPICE) method [26], covariance sparse-aware DOA estimation method in [27], OGSBI [31] and the root SBL method [33] are adopted to be compared with our proposed method. In addition, Cramér-Rao Lower Bound (CRLB) [27] is also utilized to measure the performance of these algorithms. In order to improve the estimation accuracy, an adaptive process [11] is introduced for root mean square error analysis in [27], while we do not use this adaptive process for comparison in order to prove the effectiveness of our method more clearly. All simulations are based on a ULA consisting of $M = 10$ antennas and the distance between adjacent antennas is half-wavelength. Suppose that there are two far-field narrow-band signal sources impinging on the ULA from $\theta_1 = -11.3^\circ$ and $\theta_2 = 15.6^\circ$, respectively. Without special instructions, the covariance matrix of non-uniform noise is modeled as $\boldsymbol{Q} = \mathrm{diag}\{20, 2, 1.5, 0.5, 8, 0.7, 1.1, 3, 6, 3\}$, and the spatial domain range $[-90^\circ, 90^\circ]$ is discretized by $1^\circ$. For more intuitive analysis of performance, root mean square error (RMSE) is introduced as

$$\mathrm{RMSE} = \frac{1}{K}\sum_{k=1}^{K}\sqrt{\frac{1}{\xi}\sum_{i=1}^{\xi}\left(\hat{\theta}_{i,k} - \theta_k\right)^2}, \tag{37}$$

where $\xi = 100$ is the total number of Monte Carlo trials and $\hat{\theta}_{i,k}$ is the estimated result of $\theta_k$ in the $i$th Monte Carlo simulation.

First of all, the spatial spectrum of five methods under non-uniform noise condition are compared with each other as shown in Fig. 2, and the estimation results of 50 independent simulations of the root SBL and the proposed method are also compared as shown in Fig. 3, where SNR = 0dB and the snapshot number $T = 400$. The estimated results of Fig. 2 is given in Table 1. As seen from Fig. 2, the proposed method can realize DOA estimation effectively in the coexistence of non-uniform noise and off-grid case. It can be seen from Fig. 3 that the estimated result of the proposed method is closer to the true DOA and more stable than the root SBL

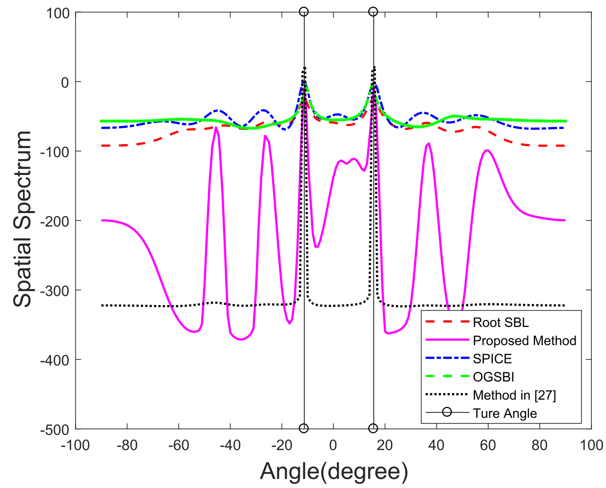



**FIGURE 2.** Comparison of spatial spectrum of five methods.

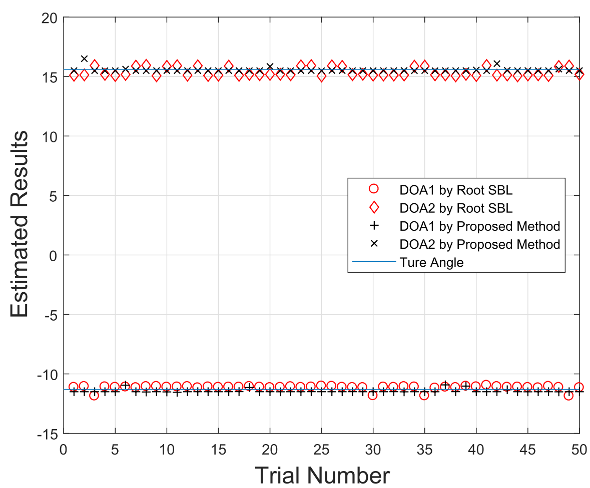



**FIGURE 3.** Comparison of estimated results of 50 independent simulations.

**TABLE 1.** Comparison of estimation results of five methods.

| Methods | DOA 1 | DOA 2 |
|---|---|---|
| Ture DOA | $-11.3000^\circ$ | $15.6000^\circ$ |
| Proposed method | $-11.2392^\circ$ | $15.6910^\circ$ |
| Root SBL | $-11.1272^\circ$ | $15.2034^\circ$ |
| OGSBI | $-10.8815^\circ$ | $14.9798^\circ$ |
| SPICE | $-11.0000^\circ$ | $16.0000^\circ$ |
| Method in [27] | $-11.0000^\circ$ | $16.0000^\circ$ |

method which only considers off-grid with uniform noise. Thus, the proposed method can effectively mitigate the influence of non-uniform noise on DOA estimation.

Then, the performance of the proposed method is tested in uniform noise case, i.e., $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_M^2$, where snapshot number is set as $T = 400$. Fig. 4 shows the result of the probability of successful detection (PSD) versus SNR in the uniform noise case. If both of the two estimated results are satisfied $|\hat{\theta}_{i,k} - \theta_k| \leq 0.5°$, the detection is successful, and the subsequent non-uniform noise scenario is identical as well. It can be seen from Fig. 4 that, although the PSD of the proposed method is not as high as the OGSBI and root SBL, it still works normally in the uniform noise case. On the other hand, it can be seen that the PSD of the proposed method reaches 100% much sooner than the SPICE and the method in [27]. The relationship between RMSE and SNR is given in Fig. 5, in which the RMSE of the proposed method, OGSBI and root SBL is decreasing with the increasing of SNR. While the other two methods decline slowly with the increasing of SNR. The main reason for these results of Fig. 4 and Fig. 5 is that the proposed method, OGSBI and root SBL can effectively reduce the off-grid error. Thus, the proposed method can not only accurately reconstruct the covariance matrix of the uniform noise, but also effectively reduce the off-grid error.

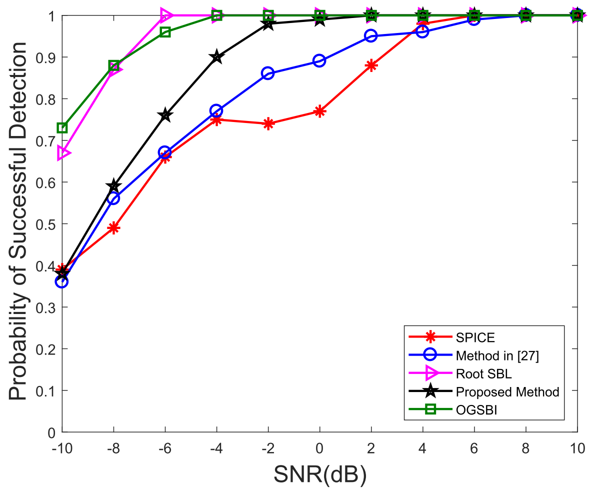



**FIGURE 4. Probability of successful detection versus SNR in the uniform noise case.**

Now, we consider the non-uniform noise scenario $\boldsymbol{Q} = \text{diag}\{20, 2, 1.5, 0.5, 8, 0.7, 1.1, 3, 6, 3\}$. The PSD versus SNR is shown in Fig. 6 with the snapshot number $T = 800$. It can be seen that, when the SNR is relatively low, the PSD of OGSBI and root SBL is lower than that of the other methods, This is because that when the SNR is low, the non-uniform noise ignored by OGSBI and root SBL seriously affects their estimation performance. On the other hand, since the SPICE and the method in [27] only consider non-uniform noise and ignore the off-grid error, their PSD are lower than that of the proposed method. Obviously, the proposed method keeps the highest PSD at all SNRs and reaches 100% sooner than the other four methods, which means that the proposed algorithm can reduce the influence of off-grid error and non-uniform noise simultaneously.

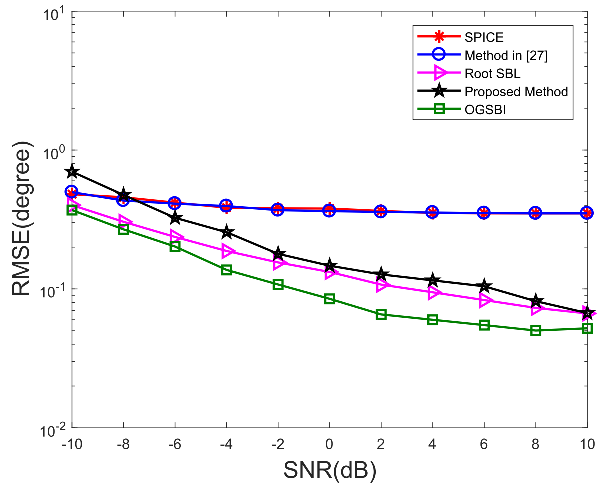



**FIGURE 5. RMSE versus SNR in the uniform noise case.**

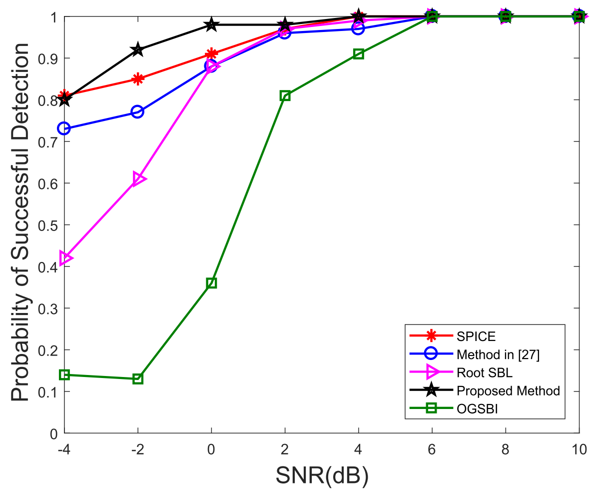



**FIGURE 6. Probability of successful detection versus SNR in the non-uniform noise case.**

Fig. 7 shows the RMSE versus SNR in non-uniform noise case with $T = 800$, and Fig. 8 depicts the RMSE versus snapshot number in non-uniform noise case with SNR $=$ 0dB. It can be obviously observed that the proposed method achieves the lowest RMSE in the relatively low SNR, and when SNR $>$ 6dB the RMSE of OGSBI is lower than the proposed method. This is because in the case of high SNR, the non-uniform noise will not have a significant impact on the estimated performance, and our proposed method mainly aims at non-uniform noise. Since the off-grid error is not considered, the performance of SPICE and the method in [27] are not significantly improved. On the other hand, although the OGSBI and the root SBL can effectively reduce the off-grid error, their performance are still seriously affected due to the influence of non-uniform noise. Hence, the OGSBI shows

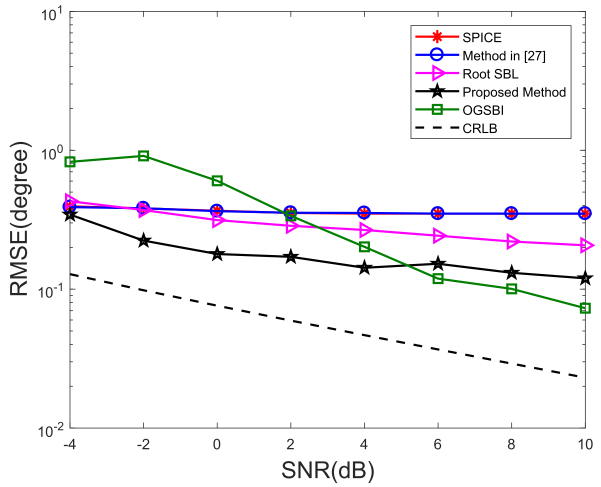



**FIGURE 7.** **RMSE versus SNR in the non-uniform noise case.**

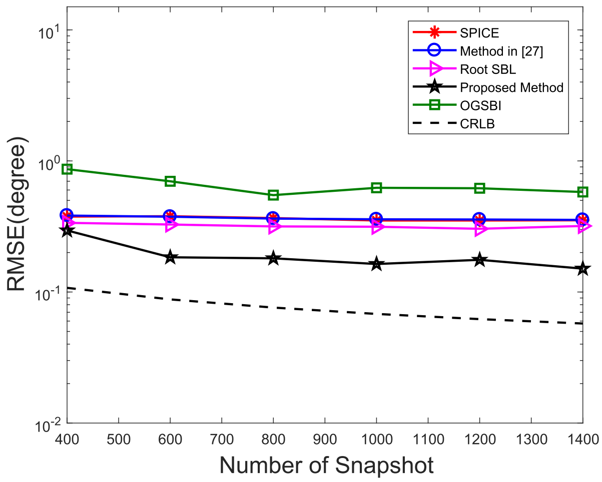



**FIGURE 8.** **RMSE versus snapshot in the non-uniform noise case.**

the worst performance in Fig. 8 and the proposed method maintains the most superior performance.

Fig. 9 shows the comparison of RMSE for different methods at different grid intervals, where SNR $=$ 0dB and the snapshot number is $T = 800$. In different grid intervals, the proposed method can still maintain superior performance over the other four methods, which should give the credit to the update process to the sampling grid in the proposed method. Conversely, the performance of SPICE and the method in [27] are worse, because they are both based on the true DOA overlapping with grid points. In addition, although the OGSBI and the root SBL also have the process of reducing the off-grid error, the appearance of non-uniform noise makes their performance inferior to the proposed method.

Finally, in order to further illustrate the robustness and effectiveness of the proposed method, the RMSE versus the worst noise power ratio (WNPR) is given in Fig. 10, where WNPR is defined as WNPR $= \sigma_{max}^2/\sigma_{min}^2$ [10]. $\sigma_{max}^2$ represents the maximum noise power and $\sigma_{min}^2$ is the minimum

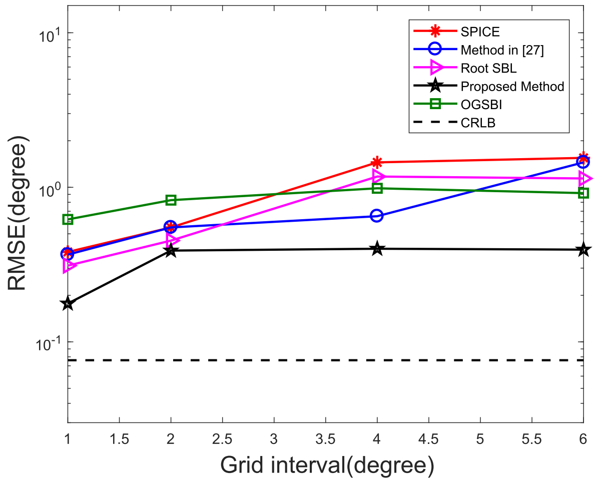



**FIGURE 9.** **RMSE versus grid interval in the non-uniform noise case.**

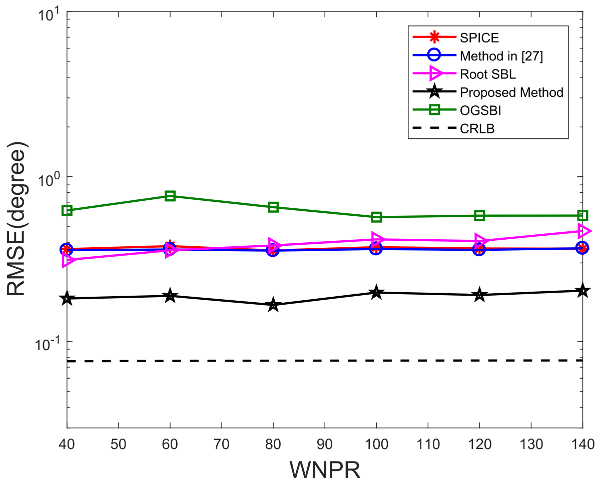



**FIGURE 10.** **RMSE versus WNPR in the non-uniform noise case.**

noise power. As seen from Fig. 10, in the case of SNR $=$ 0dB and $T = 800$, the RMSE of OGSBI is worst and that of the root SBL grows with the increasing of WNPR, and finally tends to fail to work. In contrast, the SPICE, the method in [27] and the proposed method are all robust to WNPR. And since we adopt a coarser grid, the SPICE and method in [27] are less affected by WNPR. Although the RMSE of the proposed method has some fluctuations, it is still acceptable and generally stable. Moreover, the proposed method has the lowest RMSE because of its robustness to off-grid error simultaneously, which leads to superior performance than the other methods.

## V. CONCLUSION

In this paper, we have proposed a robust SBL method for off-grid DOA estimation with non-uniform noise. In our proposed method, both the off-grid error and the influence of non-uniform noise can be effectively minimized. Based on the accurate reconstruction for non-uniform noise covariance matrix and iterative updating of the grid, the proposed method is robust to both off-grid and non-uniform noise. Extensive simulation results have proved that our method can achieve DOA estimation accurately with coexistence of off-grid error and non-uniform noise.

## REFERENCES


[1] H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," *IEEE Signal Process. Mag.*, vol. 13, no. 4, pp. 67–94, Jul. 1996.

[2] L. Wan, G. Han, H. Wang, L. Shu, N. Feng, and B. Peng, "Wearable sensor localization considering mixed distributed sources in health monitoring systems," *Sensors*, vol. 16, no. 3, p. 368, Mar. 2016.

[3] X. Wang, M. Hang, C. Shen, and D. Meng, "Robust vehicle localization exploiting two based stations cooperation: A MIMO radar perspective," *IEEE Access*, vol. 6, pp. 48747–48755, 2018.

[4] N. Hu, Z. Ye, D. Xu, and S. Cao, "A sparse recovery algorithm for DOA estimation using weighted subspace fitting," *Signal Process.*, vol. 92, no. 10, pp. 2566–2570, Oct. 2012.

[5] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propag.*, vol. AP-34, no. 3, pp. 276–280, Mar. 1986.

[6] R. Roy and T. Kailath, "Esprit-estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 7, pp. 984–995, Jul. 1989.

[7] N. E. Hurt, "Maximum likelihood estimation and MUSIC in array localization signal processing: A review," *Multidimensional Syst. Signal Process.*, vol. 1, no. 3, pp. 279–325, Sep. 1990.

[8] A. G. Jaffer, "Maximum likelihood direction finding of stochastic sources: A separable solution," in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, vol. 5, Apr. 1988, pp. 2893–2896.

[9] P. Stoica and K. Sharman, "Maximum likelihood methods for direction-of-arrival estimation," *IEEE Trans. Acoust., Speech Signal Process.*, vol. 38, no. 7, pp. 1132–1143, Jul. 1990.

[10] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J. Mach. Learn. Res.*, vol. 1, pp. 211–244, Sep. 2001.

[11] D. Malioutov, M. Çetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3010–3022, Aug. 2005.

[12] X. Xu, X. Wei, and Z. Ye, "DOA estimation based on sparse signal recovery utilizing weighted $\ell_1$-norm penalty," *IEEE Trans. Signal Process.*, vol. 19, no. 3, pp. 155–158, Mar. 2012.

[13] X. Wei, Y. Yuan, and Q. Ling, "DOA estimation using a greedy block coordinate descent algorithm," *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6382–6394, Dec. 2012.

[14] L. Wan, X. Kong, and F. Xia, "Joint range-Doppler-angle estimation for intelligent tracking of moving aerial targets," *IEEE Internet Things J.*, vol. 5, no. 3, pp. 1625–1636, Jun. 2017.

[15] X. Wang, L. Wang, X. Li, and G. Bi, "Nuclear norm minimization framework for DOA estimation in MIMO radar," *Signal Process.*, vol. 135, pp. 147–152, Jun. 2017.

[16] J. Liu, X. Wang, and W. Zhou, "Covariance vector sparsity-aware DOA estimation for monostatic MIMO radar with unknown mutual coupling," *Signal Process.*, vol. 119, pp. 21–27, Feb. 2016.

[17] X. Wang, W. Wang, J. Liu, X. Li, and J. Wang, "A sparse representation scheme for angle estimation in monostatic MIMO radar," *Signal Process.*, vol. 104, pp. 258–263, Nov. 2014.

[18] Z.-M. Liu, Z.-T. Huang, and Y.-Y. Zhou, "An efficient maximum likelihood method for direction-of-arrival estimation via sparse Bayesian learning," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 1–11, Oct. 2012.

[19] M. Pesavento and A. B. Gershman, "Maximum-likelihood direction-of-arrival estimation in the presence of unknown nonuniform noise," *IEEE Trans. Signal Process.*, vol. 49, no. 7, pp. 1310–1324, Jul. 2001.

[20] A. B. Gershman, A. L. Matveyev, and J. F. Bohme, "Maximum likelihood estimation of signal power in sensor array in the presence of unknown noise field," *IEE Proc. Radar, Sonar Navigat.*, vol. 142, no. 5, pp. 218–224, Oct. 1995.

[21] D. Madurasinghe, "A new DOA estimator in nonuniform noise," *IEEE Signal Process. Lett.*, vol. 12, no. 4, pp. 337–339, Apr. 2005.

[22] Y. Wu, C. Hou, G. Liao, and Q. Guo, "Direction-of-arrival estimation in the presence of unknown nonuniform noise fields," *IEEE J. Ocean. Eng.*, vol. 31, no. 2, pp. 504–510, Apr. 2006.

[23] B. Liao, G.-S. Liao, and J. Wen, "A method for DOA estimation in the presence of unknown nonuniform noise," *J. Electromagn. Waves Appl.*, vol. 22, nos. 14–15, pp. 2113–2123, Dec. 2008.

[24] C. E. Chen, F. Lorenzelli, R. E. Hudson, and K. Yao, "Stochastic maximum-likelihood DOA estimation in the presence of unknown nonuniform noise," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3038–3044, Jul. 2008.

[25] P. Yang, Z. Liu, and W.-L. Jiang, "Improved sparse Bayesian learning method for direction-of-arrival estimation in non-uniform noise," *J. Electromagn. Waves Appl.*, vol. 28, no. 5, pp. 563–573, Feb. 2014.

[26] P. Stoica, P. Babu, and J. Li, "SPICE: A sparse covariance-based estimation method for array processing," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 629–638, Feb. 2011.

[27] Z.-Q. He, Z.-P. Shi, and L. Huang, "Covariance sparsity-aware DOA estimation for nonuniform noise," *Digit. Signal Process.*, vol. 28, no. 1, pp. 75–81, May 2014.

[28] X. Wang, M. Huang, and G. Bi, "Sparse Bayesian learning for DOA estimation in MIMO radar with unknown nonuniform noise," in *Proc. CIE Int. Conf. Radar (RADAR)*, Guangzhou, China, Oct. 2016, pp. 1–5.

[29] L. Wan, G. Han, L. Shu, S. Chan, and N. Feng, "PD source diagnosis and localization in industrial high-voltage insulation system via multimodal joint sparse representation," *IEEE Trans. Ind. Electron.*, vol. 63, no. 4, pp. 2506–2516, Apr. 2016.

[30] H. Zhu, G. Leus, and G. B. Giannakis, "Sparsity-cognizant total least-squares for perturbed compressive sampling," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2002–2016, May 2011.

[31] Z. Yang, L. Xie, and C. Zhang, "Off-grid direction of arrival estimation using sparse Bayesian inference," *IEEE Trans. Signal Process.*, vol. 61, no. 1, pp. 38–43, Jan. 2013.

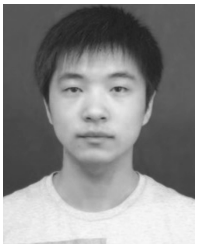
[32] Y. Zhang, Z. Ye, X. Xu, and N. Hu, "Off-grid DOA estimation using array covariance matrix and block-sparse Bayesian learning," *Signal Process.*, vol. 98, no. 1, pp. 197–201, May 2014.

[33] J. Dai, X. Bao, W. Xu, and C. Chang, "Root sparse Bayesian learning for off-grid DOA estimation," *IEEE Signal Process. Lett.*, vol. 24, no. 1, pp. 46–50, Jan. 2017.

[34] L. Wan, G. Han, L. Shu, S. Chan, and T. Zhu, "The application of DOA estimation approach in patient tracking systems with high patient density," *IEEE Trans. Ind. Informat.*, vol. 12, no. 6, pp. 2353–2364, Dec. 2016.


**HUAFEI WANG** was born in Jiangxi in 1995. He received the B.S. degree from Hainan University, Haikou, China, in 2017, where he is currently pursuing the M.S. degree in information and communication engineering. His research interests are array signal processing and MIMO radars.

**XIANPENG WANG** was born in Hainan in 1986. He received the M.S. and Ph.D. degrees from the College of Automation from Harbin Engineering University, Harbin, China, in 2012 and 2015, respectively. He was a full-time Research Fellow with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, from 2015 to 2016. He is currently a Professor with the College of Information Science and Technology, Hainan University. His major research interests include communication systems, array signal processing, radar signal processing, and compressed sensing and its applications. He has authored or published over 50 papers in his research-related journals and international conference proceedings. He has served as a reviewer for over 20 journals.



**LIANGTIAN WAN** (M'15) received the B.S. and Ph.D. degrees from the College of Information and Communication Engineering, Harbin Engineering University, Harbin, China, in 2011 and 2015, respectively. From 2015 to 2017, he was a Research Fellow with the School of Electrical and Electrical Engineering, Nanyang Technological University, Singapore. He is currently an Associate Professor with the School of Software, Dalian University of Technology, China. He has published over 40 scientific papers in international journals and conferences. His research interests include social network analysis and mining, big data, array signal processing, wireless sensor networks, and compressive sensing and its applications. He is currently an Associate Editor of the IEEE Access.

**MENGXING HUANG** received the Ph.D. degree from Northwestern Polytechnical University in 2007. He then joined the Research Institute of Information Technology, Tsinghua University, as a Post-Doctoral Researcher. In 2009, he joined Hainan University. He is currently a Professor and the Ph.D. Supervisor of computer science and technology, and the Dean of the College of Information Science and Technology. He is also the Executive Vice-President of the Hainan Province Institute of Smart City and the Leader of the Service Science and Technology Team, Hainan University. He has published over 60 academic papers as the first or corresponding author, and two monographs and two translations. He held 12 patents of invention and owns three software copyright. He received one Second Class and one Third Class Prize from The Hainan Provincial Scientific and Technological Progress. His current research interests include signal processing for sensor systems, big data, and intelligent information processing.