Off-Grid Direction-of-Arrival Estimation Using Second-Order Taylor Approximation

Huiping Huang, *Student Member, IEEE*, Hing Cheung So, *Fellow, IEEE* and Abdelhak M. Zoubir, *Fellow, IEEE*

***Abstract*—The problem of off-grid direction-of-arrival (DOA) estimation is investigated in this letter. We develop a grid-based method to jointly estimate the closest angular grids, and the gaps between the estimated grids and the corresponding DOAs. By using a second-order Taylor approximation, the data model under the framework of joint-sparse representation is formulated. We point out an important property of the signals of interest in the model, namely the proportionality relationship, which is empirically demonstrated to be useful in the sense that it increases the probability of the mixing matrix satisfying the block restricted isometry property. Simulation examples demonstrate the effectiveness and superiority of the proposed method against several state-of-the-art approaches.**

***Index Terms*—Block restricted isometry property (RIP), compressive sensing, off-grid DOA estimation, second-order Taylor approximation**

I. Introduction

GRID-BASED methods have gained interest in direction-of-arrival (DOA) estimation in recent years. Such approaches include least absolute shrinkage and selection operator (LASSO) [1], [2] and sparse iterative covariance-based estimation (SPICE) [3]–[6], among others. See [7] for a comprehensive review of grid-based sparse methods for DOA estimation. The advantage of grid-based methods is that they have super-high resolution even in the case when only one single snapshot is available, provided that all the source DOAs align exactly with the preset angular grid. However, this condition may not be satisfied in practice, since the region of interest (ROI) contains infinite angular candidates and hence grid mismatch almost always exists when we split the ROI into a finite number of grids. This is known as the off-grid issue and has attracted a lot of research interest in array signal processing during the past decade, see for example [8]–[19].

Existing solutions to tackle the off-grid problem can be categorized into three groups. The first group uses denser grids or the coarse-to-fine strategy such as [2]. The drawbacks of these methods are twofold. On one hand, denser grids lead to extremely expensive computational complexity; on the other hand, too dense grids may result in weak incoherence among the steering vectors. The second group consists of the so-called *gridless* approach [20]–[25]. Its weakness is that most of these methods are restricted to regularly sampled measurements that can only be taken from a uniform linear array (ULA) [26]. The last group of methods estimates the off-grid bias together with the grids closest to the true DOAs. Representative works include the first-order Taylor approximation [12], [13] and the neighbor-grid based method [16], denoted in this paper as 1st Taylor G-LASSO and Neighbor G-LASSO, respectively.

It is known that in general the first-order Taylor approximation is accurate enough, especially when the grid size is small. However, when the grid size is set not small enough so as to save computational cost, there still exists a large bias. In this situation, a high-order Taylor approximation decreases the approximation error. To this end, we introduce a second-order Taylor approximation in off-grid DOA estimation. We observe in this case the proportionality relationship of the signals of interest. With this, we propose a novel optimization approach which is shown by simulation to produce more accurate DOA estimates in off-grid scenarios.

II. Signal Model

Suppose that a ULA consisting of M sensors receives K far-field narrowband signals from DOAs $\boldsymbol{\phi} = [\phi_1, \phi_2, \cdots, \phi_K]^{\mathrm{T}}$ with $\phi_k \in [-\pi/2, \pi/2]$, where the superscript $\cdot^{\mathrm{T}}$ denotes the transpose operator. The array observation can be modeled as

$$\mathbf{y} = \sum_{k=1}^{K} s_k \mathbf{a}(\phi_k) + \mathbf{n} = \mathbf{A}(\boldsymbol{\phi})\mathbf{s} + \mathbf{n},$$

where s_k is the k-th signal waveform, $\mathbf{s} = [s_1, s_2, \cdots, s_K]^{\mathrm{T}}$ represents the signal vector, and $\mathbf{n} \in \mathbb{C}^M$ is the noise vector with $\mathbb{C}$ being the set of complex numbers. The steering matrix $\mathbf{A}(\boldsymbol{\phi}) = [\mathbf{a}(\phi_1), \mathbf{a}(\phi_2), \cdots, \mathbf{a}(\phi_K)] \in \mathbb{C}^{M\times K}$ has the steering vectors as columns, where

$$\mathbf{a}(\phi_k) = [1, e^{\jmath \frac{2\pi d}{\lambda} \sin(\phi_k)}, \cdots, e^{\jmath \frac{2\pi d}{\lambda} \sin(\phi_k)(M-1)}]^{\mathrm{T}},$$

for $k = 1, 2, \cdots, K$, $\jmath = \sqrt{-1}$, λ is the signal wavelength, and d is the inter-element spacing usually set to be $\lambda/2$.

In grid-based methods, we formulate the signal model by means of a sparse representation, as

$$\mathbf{y} = \sum_{l=1}^{L} x_l \mathbf{a}(\theta_l) + \mathbf{n} = \mathbf{A}(\boldsymbol{\theta})\mathbf{x} + \mathbf{n},$$

where $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_L]^{\mathrm{T}}$ denotes the angular grid vector with L being the number of grids (in general $L \gg M > K$), $\mathbf{A}(\boldsymbol{\theta}) \in \mathbb{C}^{M\times L}$ stands for the overcomplete dictionary matrix, and $\mathbf{x} = [x_1, x_2, \cdots, x_L]^{\mathrm{T}}$ is a sparse vector whose elements $x_l = s_k$ if $\theta_l = \phi_k$, and $x_l = 0$ otherwise. When the true DOAs do not exactly lie in the preset angular grids, we encounter the off-grid issue. To handle this problem, we propose a method to simultaneously estimate the closest angular grids, and the gaps between the closest grids and the corresponding DOAs, using a second-order Taylor approximation.

The work of H. Huang is supported by the Graduate School CE within the Centre for Computational Engineering at Technische Universität Darmstadt.

H. Huang and A. M. Zoubir are with Technische Universität Darmstadt, 64283 Darmstadt, Germany (e-mail: {h.huang; zoubir}@spg.tu-darmstadt.de).

H. C. So is with Department of Electrical Engineering, City University of Hong Kong, Hong Kong, China (e-mail: hcso@ee.cityu.edu.hk).

III. Proposed Method

A. Second-Order Taylor Approximation

We start by considering a second-order Taylor approximation of the steering vectors. For any ϕ_l, we have

$$\mathbf{a}(\phi_l) \approx \mathbf{a}(\theta_l) + \mathbf{a}'(\theta_l)p_l + \frac{\mathbf{a}''(\theta_l)}{2}p_l^2,$$

where θ_l is the grid closest to ϕ_l, $\mathbf{a}'(\theta_l) = \frac{\mathrm{d}\mathbf{a}(\theta_l)}{\mathrm{d}\theta_l}$, $\mathbf{a}''(\theta_l) = \frac{\mathrm{d}^2\mathbf{a}(\theta_l)}{\mathrm{d}\theta_l^2}$, and $p_l = \phi_l - \theta_l \in [-\delta/2, \delta/2]$ with δ being the grid size. Collecting all the candidates, we have

$$[\mathbf{a}(\phi_1), \cdots, \mathbf{a}(\phi_L)] = \mathbf{A}(\boldsymbol{\theta}) + \mathbf{A}'(\boldsymbol{\theta})\mathrm{diag}\{\mathbf{p}\} + \frac{1}{2}\mathbf{A}''(\boldsymbol{\theta})\mathrm{diag}\{\mathbf{p}\}^2,$$

where $\mathbf{A}'(\boldsymbol{\theta}) = [\mathbf{a}'(\theta_1), \cdots, \mathbf{a}'(\theta_L)] \in \mathbb{C}^{M\times L}$, $\mathbf{A}''(\boldsymbol{\theta}) = [\mathbf{a}''(\theta_1), \cdots, \mathbf{a}''(\theta_L)] \in \mathbb{C}^{M\times L}$, $\mathbf{p} = [p_1, p_2, \cdots, p_L]^{\mathrm{T}}$, and $\mathrm{diag}\{\mathbf{p}\}$ is a diagonal matrix whose main diagonal is $\mathbf{p}$. Hence, the signal model can be approximately written as follows

$$\begin{aligned}\mathbf{y} &\approx \left[\mathbf{A}(\boldsymbol{\theta}) + \mathbf{A}'(\boldsymbol{\theta})\mathrm{diag}\{\mathbf{p}\} + \frac{1}{2}\mathbf{A}''(\boldsymbol{\theta})\mathrm{diag}\{\mathbf{p}\}^2\right]\mathbf{x} + \mathbf{n} \\ &= \left[\mathbf{A}(\boldsymbol{\theta}), \mathbf{A}'(\boldsymbol{\theta}), \frac{1}{2}\mathbf{A}''(\boldsymbol{\theta})\right]\begin{bmatrix}\mathbf{x} \\ \mathrm{diag}\{\mathbf{p}\}\mathbf{x} \\ \mathrm{diag}\{\mathbf{p}\}^2\mathbf{x}\end{bmatrix} + \mathbf{n}, \end{aligned} \tag{1}$$

where the signal of interest $[\mathbf{x}^{\mathrm{T}}, (\mathrm{diag}\{\mathbf{p}\}\mathbf{x})^{\mathrm{T}}, (\mathrm{diag}\{\mathbf{p}\}^2\mathbf{x})^{\mathrm{T}}]^{\mathrm{T}}$ is referred to as the block signal in the sequel.

B. Properties of the Block Signal

As shown in signal model (1), the unknown block signal is divided into three parts: (i) $\mathbf{x}_1 \triangleq \mathbf{x}$, (ii) $\mathbf{x}_2 \triangleq \mathrm{diag}\{\mathbf{p}\}\mathbf{x}$, and (iii) $\mathbf{x}_3 \triangleq \mathrm{diag}\{\mathbf{p}\}^2\mathbf{x}$. Denote the l-th entries of $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_3$ as $x_{1,l}$, $x_{2,l}$, and $x_{3,l}$, respectively. We notice the following properties of the block signal $[\mathbf{x}_1^{\mathrm{T}}, \mathbf{x}_2^{\mathrm{T}}, \mathbf{x}_3^{\mathrm{T}}]^{\mathrm{T}}$.

- Since $\mathbf{x}$ is a sparse vector as mentioned in Section II, $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_3$ are all sparse and share the same sparsity pattern. This property is known as block-sparsity [12] or joint-sparsity [13].
- It holds that $x_{2,l} = p_l x_{1,l}$ and $x_{3,l} = p_l^2 x_{1,l}$. Due to $-\delta/2 \le p_l \le \delta/2,\ \forall l \in \{1, 2, \cdots, L\}$, it is easy to verify that the following inequalities hold:

$$-\frac{\delta}{2}|\mathbf{x}_1| \le \mathbf{x}_2 \le \frac{\delta}{2}|\mathbf{x}_1|, \tag{2a}$$

$$-\left(\frac{\delta}{2}\right)^2|\mathbf{x}_1| \le \mathbf{x}_3 \le \left(\frac{\delta}{2}\right)^2|\mathbf{x}_1|, \tag{2b}$$

  where $|\mathbf{x}|$ denotes the element-wise absolute value of $\mathbf{x}$, and $\le$ is the element-wise less-than or equal-to symbol.
- It can be seen that $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_3$ satisfy the proportionality relationship, as

$$x_{2,l}^2 = x_{1,l}x_{3,l},\ \forall l \in \{1, 2, \cdots, L\}. \tag{3}$$

C. Problem Formulation Development

Based on the aforementioned relationships among $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_3$, we propose the following minimization problem, as

$$\begin{aligned}\min_{\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3}\quad & g(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \\ \text{s.t.}\quad & \text{(2a), (2b), and (3).}\end{aligned} \tag{4}$$

The cost function in (4) is given by

$$\begin{aligned}g(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \triangleq \frac{1}{2}\left\|\mathbf{y} - \mathbf{A}(\boldsymbol{\theta})\mathbf{x}_1 - \mathbf{A}'(\boldsymbol{\theta})\mathbf{x}_2 - \frac{1}{2}\mathbf{A}''(\boldsymbol{\theta})\mathbf{x}_3\right\|_2^2 + \\ \mu\left\|\left[\mathbf{x}_1^{\mathrm{T}}, \mathbf{x}_2^{\mathrm{T}}, \mathbf{x}_3^{\mathrm{T}}\right]^{\mathrm{T}}\right\|_{2,1},\end{aligned} \tag{5}$$

where μ is a regularization parameter balancing the data fitting and the model sparsity, $\|\cdot\|_2$ denotes the L_2 norm of a vector, and $\|\cdot\|_{2,1}$ is the mixed $L_{2,1}$ norm of a vector, defined as

$$\left\|\left[\mathbf{x}_1^{\mathrm{T}}, \mathbf{x}_2^{\mathrm{T}}, \mathbf{x}_3^{\mathrm{T}}\right]^{\mathrm{T}}\right\|_{2,1} = \sum_{l=1}^{L}\sqrt{|x_{1,l}|^2 + |x_{2,l}|^2 + |x_{3,l}|^2}.$$

Problem (4) is non-convex due to its constraints. Assuming that the signals are positive [13], i.e., $\mathbf{s} > \mathbf{0}$ (and $\mathbf{x}_1 = \mathbf{x} \ge \mathbf{0}$) where $>$ and $\ge$ are element-wise operators, and $\mathbf{0}$ is the all-zeros vector of appropriate length, the first two constraints in (4) become

$$-\frac{\delta}{2}\mathbf{x}_1 \le \mathbf{x}_2 \le \frac{\delta}{2}\mathbf{x}_1,\quad \mathbf{0} \le \mathbf{x}_3 \le \left(\frac{\delta}{2}\right)^2\mathbf{x}_1,\quad \mathbf{x}_1 \ge \mathbf{0}, \tag{6}$$

which is convex. It is worth pointing out that $\mathbf{0} \le \mathbf{x}_3$ in (6) is the result of $x_{3,l} = p_l^2 x_{1,l},\ \forall l \in \{1, 2, \cdots, L\}$ and $\mathbf{x}_1 \ge \mathbf{0}$.

In the sequel, we consider the last constraint in (4), viz. (3). Firstly, we transfer (3) to its equivalent form as in [27]:

$$\left\|\begin{bmatrix}2x_{2,l} \\ x_{1,l} - x_{3,l}\end{bmatrix}\right\|_2 = x_{1,l} + x_{3,l},\ \forall l \in \{1, 2, \cdots, L\}. \tag{7}$$

Then, we introduce an additional variable $\mathbf{z} \in \mathbb{R}^L$ with entries z_l $(l = 1, 2, \cdots, L)$ satisfying

$$0 \le z_l \le \eta,\ \forall l \in \{1, 2, \cdots, L\}, \tag{8}$$

where $\mathbb{R}$ denotes the set of real numbers and η is a small user-defined parameter, and rewrite (7) as

$$\left\|\begin{bmatrix}2x_{2,l} \\ x_{1,l} - x_{3,l}\end{bmatrix}\right\|_2 \le x_{1,l} + x_{3,l} + z_l,\ \forall l \in \{1, 2, \cdots, L\}, \tag{9}$$

which belongs to the set of standard second-order cone (SOC) and hence is convex.

By replacing the constraints (2a) and (2b) with (6) and replacing the constraint (3) with (8) and (9), we finally relax the non-convex problem (4) into a convex one, as

$$\begin{aligned}\min_{\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3,\mathbf{z}}\quad & g(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \\ \text{s.t.}\quad & \text{(6), (8), and (9).}\end{aligned} \tag{10}$$

Remark. *Note that the proposed method is developed for the single-snapshot scenarios. However, it can be easily extended to the case of multiple snapshots.*

To analyze the computational cost, we formulate Problem (10) under the framework of standard SOC programming (SOCP) [28], as

$$\begin{aligned}\min_{\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3,\mathbf{z},\mathbf{t}}\quad & \sum_{l=1}^{L} t_l = \mathbf{1}^{\mathrm{T}}\mathbf{t} \\ \text{s.t.}\quad & \text{(6), (8), and (9),} \\ & \sqrt{|x_{1,l}|^2 + |x_{2,l}|^2 + |x_{3,l}|^2} \le t_l,\ \forall l \in \{1, 2, \cdots, L\}, \\ & \left\|\mathbf{y} - \mathbf{A}(\boldsymbol{\theta})\mathbf{x}_1 - \mathbf{A}'(\boldsymbol{\theta})\mathbf{x}_2 - \frac{1}{2}\mathbf{A}''(\boldsymbol{\theta})\mathbf{x}_3\right\|_2 \le \epsilon,\end{aligned} \tag{11}$$

TABLE I
COMPUTATIONAL COST WITH IMPLEMENTATION OF SOCP

Method	Cost per Iteration	No. of Iterations
LASSO	$\mathcal{O}\big((M+1)L^2\big)$	$\mathcal{O}(1)$
Neighbor G-LASSO	$\mathcal{O}\big(4(M+1)L^2+12L\big)$	$\mathcal{O}(\sqrt{L})$
1st Taylor G-LASSO	$\mathcal{O}\big(4(M+1)L^2+28L\big)$	$\mathcal{O}(\sqrt{L})$
2nd Taylor G-LASSO	$\mathcal{O}\big(9(M+1)L^2+72L\big)$	$\mathcal{O}(\sqrt{L})$

where $\mathbf{1}$ denotes the all-ones vector, $\mathbf{t}=[t_1,t_2,\cdots,t_L]^{\mathrm{T}}$ is an auxiliary variable vector, and ϵ is a tuning parameter related to μ in (10). The computational complexity of Problem (11) with implementation of SOCP is $\mathcal{O}\big(9(M+1)L^2+72L\big)$ per iteration, and the number of iterations is bounded above by $\mathcal{O}(\sqrt{L})$ [28]. The proposed second-order Taylor approximation method is referred to as 2nd Taylor G-LASSO. The computational cost of the 2nd Taylor G-LASSO, as well as those of LASSO [1], [2], Neighbor G-LASSO [16], and 1st Taylor G-LASSO [12], [13], are summarized in Table I.

IV. DISCUSSION

In this section, we discuss the uniqueness of the proposed signal model in (1). To this end, we first introduce the following definition and theorem [29]:

Definition. *An $M\times bL$ block matrix $\mathbf{D}$ is said to have the block restricted isometry property (RIP) with parameter β_K, if for every K block-sparse vector $\mathbf{c}$ of length bL, it holds that*

$$(1-\beta_K)\|\mathbf{c}\|_2^2\le\|\mathbf{D}\mathbf{c}\|_2^2\le(1+\beta_K)\|\mathbf{c}\|_2^2.$$

Theorem. *Let $\mathbf{y}=\mathbf{D}\mathbf{c}_0$ be measurements of a K block-sparse vector $\mathbf{c}_0$. If $\mathbf{D}$ satisfies the block RIP with parameter $\beta_{2K}<1$, then there exists a unique block-sparse vector $\mathbf{c}$ satisfying $\mathbf{y}=\mathbf{D}\mathbf{c}$; and further, if $\mathbf{D}$ satisfies the block RIP with parameter $\beta_{2K}<\sqrt{2}-1$, then the convex optimization problem: $\min_{\mathbf{c}}\|\mathbf{c}\|_{2,1}$ s.t. $\mathbf{y}=\mathbf{D}\mathbf{c}$, has a unique solution and the solution is equal to $\mathbf{c}_0$.*

Define $\mathbf{c}_0=[\mathbf{x}^{\mathrm{T}},(\mathrm{diag}\{\mathbf{p}\}\mathbf{x})^{\mathrm{T}},(\mathrm{diag}\{\mathbf{p}\}^2\mathbf{x})^{\mathrm{T}}]^{\mathrm{T}}$ and $\mathbf{D}=[\mathbf{A}(\boldsymbol{\theta}),\mathbf{A}'(\boldsymbol{\theta}),\frac{1}{2}\mathbf{A}''(\boldsymbol{\theta})]$. In the absence of noise, our proposed model in (1) can be rewritten as: $\mathbf{y}=\mathbf{D}\mathbf{c}_0$. Without loss of generality, we denote $\bar{\mathbf{D}}$ as the column-normalized matrix structured from $\mathbf{D}$. Our goal is to check whether $\bar{\mathbf{D}}$ satisfies the block RIP with parameter $\beta_{2K}<1$ and $\beta_{2K}<\sqrt{2}-1$. Note that determining the RIP parameter, i.e., β_{2K}, of a given matrix is in general an NP-hard problem [30]. In what follows, we introduce a Monte Carlo test to check the condition of the block RIP of $\bar{\mathbf{D}}$.

According to the definition, if $\bar{\mathbf{D}}$ has the block RIP with parameter β_{2K}, then for any $2K$ block-sparse vector $\mathbf{c}$ of length bL, it holds that

$$(1-\beta_{2K})\|\mathbf{c}\|_2^2\le\|\bar{\mathbf{D}}\mathbf{c}\|_2^2\le(1+\beta_{2K})\|\mathbf{c}\|_2^2. \tag{12}$$

Note that, for any $2K$ block-sparse vector $\mathbf{c}$, we can write its unit-norm vector as $\bar{\mathbf{c}}=\mathbf{c}/\|\mathbf{c}\|_2$, such that $\|\bar{\mathbf{c}}\|_2=1$. As a result, (12) becomes:

$$(1-\beta_{2K})\le\frac{\|\bar{\mathbf{D}}\mathbf{c}\|_2^2}{\|\mathbf{c}\|_2^2}=\|\bar{\mathbf{D}}\bar{\mathbf{c}}\|_2^2\le(1+\beta_{2K}).$$

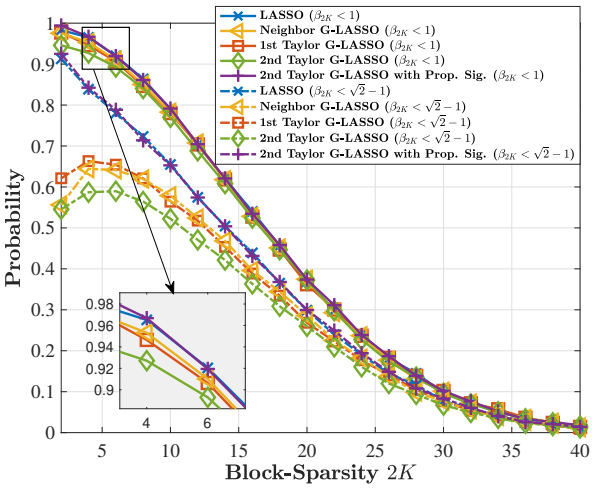

Fig. 1. Empirical probabilities of $\{\beta_{2K}<1\}$ and $\{\beta_{2K}<\sqrt{2}-1\}$ versus block-sparsity $2K$ with 10^4 Monte Carlo runs, $M=8$, and $L=201$.

Based on the above inequalities, the parameter β_{2K} can be calculated as

$$\beta_{2K}=\max\left\{\|\bar{\mathbf{D}}\bar{\mathbf{c}}\|_2^2-1\,,1-\|\bar{\mathbf{D}}\bar{\mathbf{c}}\|_2^2\right\}. \tag{13}$$

We randomly generate a unit-norm $2K$ block-sparse vector $\bar{\mathbf{c}}$, and calculate β_{2K} using (13). By repeatedly performing the above steps for 10^4 Monte Carlo runs, we estimate the empirical probabilities of $\{\beta_{2K}<1\}$ and $\{\beta_{2K}<\sqrt{2}-1\}$. The empirical probabilities versus block-sparsity $2K$ are presented in Fig. 1, with $M=8$, $L=201$, and $b=1$ for LASSO, $b=2$ for Neighbor G-LASSO and 1st Taylor G-LASSO, and $b=3$ for 2nd Taylor G-LASSO. It is seen that when the block-sparsity is small (less than 8), the probabilities of $\{\beta_{2K}<1\}$ of all the tested methods are high (greater than 0.9), and their probabilities of $\{\beta_{2K}<\sqrt{2}-1\}$ are larger than 0.5. Note that in Fig. 1, the plot of 2nd Taylor G-LASSO with proportional signals (abbreviated as "Prop. Sig." in the figure), i.e., (3), has the highest probability. This reveals that the proportionality relationship of the block signal contains useful information in the sense that it increases the probabilities of $\{\beta_{2K}<1\}$ and $\{\beta_{2K}<\sqrt{2}-1\}$.

V. SIMULATION

We evaluate the DOA estimation performance of the proposed 2nd Taylor G-LASSO, compared with LASSO [1], [2], Neighbor G-LASSO [16], and 1st Taylor G-LASSO [12], [13]. We adopt the root-mean squared error (RMSE) and the empirical probability of correct detection (PCD) as performance metrics, defined as [31]:

$$\mathrm{RMSE}=10\log_{10}\left(\sqrt{\frac{1}{KQ}\sum_{k=1}^{K}\sum_{q=1}^{Q}\Big(\sin(\widehat{\phi}_{k,q})-\sin(\phi_k)\Big)^2}\right)$$

and $\mathrm{PCD}=Q_{\mathrm{suc}}/Q$, respectively, where $\widehat{\phi}_{k,q}$ denotes the DOA estimate of the k-th signal in the q-th Monte Carlo run, Q is the

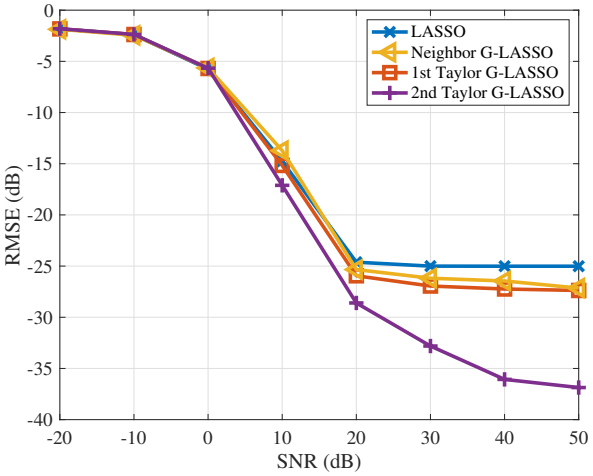

Fig. 2. RMSE versus SNR with $M = 16$ sensors, $K = 2$ sources, $L = 201$ angular grids, and grid size $\delta = \pi/200$.

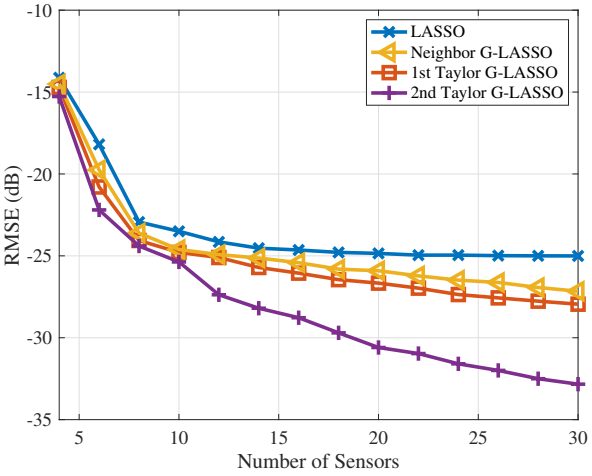

Fig. 4. RMSE versus number of sensors with SNR $= 20$ dB, $K = 2$ sources, $L = 201$ angular grids, and grid size $\delta = \pi/200$.

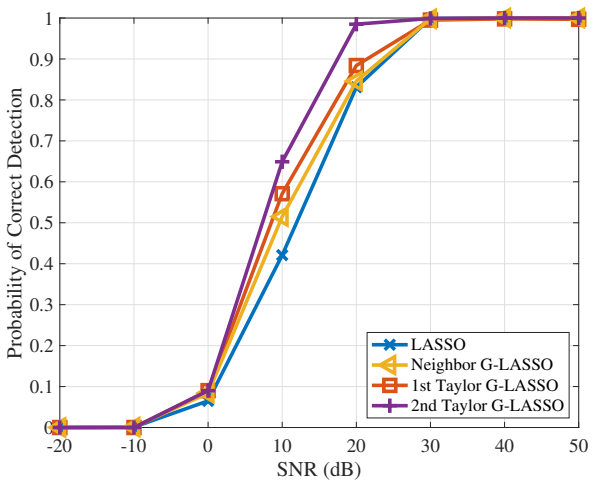

Fig. 3. PCD versus SNR with $M = 16$ sensors, $K = 2$ sources, $L = 201$ angular grids, and grid size $\delta = \pi/200$.

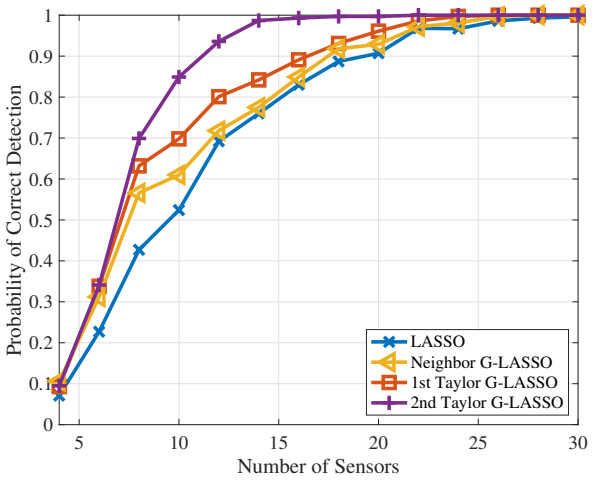

Fig. 5. PCD versus number of sensors with SNR $= 20$ dB, $K = 2$ sources, $L = 201$ angular grids, and grid size $\delta = \pi/200$.

total number of Monte Carlo trials, and Q_{suc} is the number of trials where the DOA estimates $\left\{\widehat{\phi}_k \,\middle|\, k = 1, 2, \cdots, K\right\}$ satisfy

$$\max_k \left\{\left|\widehat{\phi}_k - \phi_k\right|\right\} \leq \delta/2.$$

In the first experiment, a ULA of $M = 16$ omnidirectional sensors is considered to receive $K = 2$ signals from directions $\boldsymbol{\phi} = [0.1825, 0.9177]^{\mathrm{T}}$. The angular grid size is set to be $\delta = \pi/200$, and the number of grids is $L = 201$. Two parameters utilized in (10) are given as $\eta = 10^{-5}$ and $\mu = \sigma\sqrt{M \log_{10} M}$ [32] with σ denoting the standard deviation of the noise vector, which is assumed to be known *a priori* in our simulations. $Q = 100$ Monte Carlo trials are performed. The results of RMSE versus SNR and PCD versus SNR are plotted in Figs. 2 and 3, respectively. It is seen that, in the large SNR region, 2nd Taylor G-LASSO has significantly lower RMSE compared with the other grid-based approaches, and the PCD of 2nd Taylor G-LASSO is higher than those of the other tested methods.

In the second experiment, we fix SNR to 20 dB and change the number of sensors from 4 to 30, while the remaining parameters are the same as those in the first experiment. The RMSE and PCD are depicted in Figs. 4 and 5, respectively. The results exhibit again better performance of the proposed 2nd Taylor G-LASSO than the other competitors.

VI. Conclusion

We have investigated the off-grid DOA estimation problem and have proposed a method using the second-order Taylor approximation. By exploring the properties of the block signal, we have added the proportionality relationship to our optimization problem. A Monte Carlo test has shown the usefulness of such proportionality relationship in the sense that it increases the probabilities of $\{\beta_{2K} < 1\}$ and $\{\beta_{2K} < \sqrt{2}-1\}$. Numerical results have demonstrated that the proposed method outperforms several existing off-grid DOA estimation approaches.

REFERENCES

[1] R. Tibshirani, "Regression shrinkage and selection via the LASSO," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.

[2] D. Malioutov, M. Çetin, and A. S. Willsky, "A sparse signal reconstruction perspective for source localization with sensor arrays," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 3010–3022, August 2005.

[3] P. Stoica, P. Babu, and J. Li, "New method of sparse parameter estimation in separable models and its use for spectral analysis of irregularly sampled data," *IEEE Transactions on Signal Processing*, vol. 59, no. 1, pp. 35–47, January 2011.

[4] ——, "SPICE: A sparse covariance-based estimation method for array processing," *IEEE Transactions on Signal Processing*, vol. 59, no. 2, pp. 629–638, February 2011.

[5] P. Stoica and P. Babu, "SPICE and LIKES: Two hyperparameter-free methods for sparse-parameter estimation," *Signal Processing*, vol. 92, no. 7, pp. 1580–1590, July 2012.

[6] P. Babu and P. Stoica, "Connection between SPICE and square-root LASSO for sparse parameter estimation," *Signal Processing*, vol. 95, pp. 10–14, February 2014.

[7] Z. Yang, J. Li, P. Stoica, and L. Xie, "Chapter 11 - Sparse methods for direction-of-arrival estimation," in *Academic Press Library in Signal Processing, Volume 7*, R. Chellappa and S. Theodoridis, Eds. Academic Press, 2018, pp. 509–581.

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

[9] Z. Yang, C. Zhang, and L. Xie, "Robustly stable signal recovery in compressed sensing with structured matrix perturbation," *IEEE Transactions on Signal Processing*, vol. 60, no. 9, pp. 4658–4671, September 2012.

[10] Z. Yang, L. Xie, and C. Zhang, "Off-grid direction of arrival estimation using sparse Bayesian inference," *IEEE Transactions on Signal Processing*, vol. 61, no. 1, pp. 38–43, January 2013.

[11] M. F. Duarte and R. G. Baraniuk, "Spectral compressive sensing," *Applied and Computational Harmonic Analysis*, vol. 35, no. 1, pp. 111–129, July 2013.

[12] R. Jagannath and K. Hari, "Block sparse estimator for grid matching in single snapshot DoA estimation," *IEEE Signal Processing Letters*, vol. 20, no. 11, pp. 1038–1041, November 2013.

[13] Z. Tan, P. Yang, and A. Nehorai, "Joint sparse recovery method for compressed sensing with structured dictionary mismatches," *IEEE Transactions on Signal Processing*, vol. 62, no. 19, pp. 4997–5008, October 2014.

[14] A. C. Walewski, C. Steffens, and M. Pesavento, "Off-grid parameter estimation based on joint sparse regularization," in *Proceedings of International ITG Conference on Systems, Communications and Coding (SCC)*, Hamburg, Germany, February 2017, pp. 1–6.

[15] 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, November 2017.

[16] A. Abtahi, S. Gazor, and F. Marvasti, "Off-grid localization in MIMO radars using sparsity," *IEEE Signal Processing Letters*, vol. 25, no. 2, pp. 313–317, February 2018.

[17] X. Wu, W.-P. Zhu, J. Yan, and Z. Zhang, "Two sparse-based methods for off-grid direction-of-arrival estimation," *Signal Processing*, vol. 142, pp. 87–95, January 2018.

[18] C. Zhou, Y. Gu, Z. Shi, and Y. D. Zhang, "Off-grid direction-of-arrival estimation using coprime array interpolation," *IEEE Signal Processing Letters*, vol. 25, no. 11, pp. 1710–1714, November 2018.

[19] Z. Wan and W. Liu, "Non-coherent DOA estimation of off-grid signals with uniform circular arrays," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toronto, Canada, June 2021, pp. 4370–4374.

[20] 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, November 2013.

[21] 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, June 2015.

[22] J. Steinwandt, F. Roemer, C. Steffens, M. Haardt, and M. Pesavento, "Gridless super-resolution direction finding for strictly non-circular sources based on atomic norm minimization," in *Proceedings of 50th Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, Pacific Grove, USA, November 2016, pp. 1518–1522.

[23] C. Steffens, W. Suleiman, A. Sorg, and M. Pesavento, "Gridless compressed sensing under shift-invariant sampling," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, New Orleans, USA, March 2017, pp. 4735–4739.

[24] M. Wagner, P. Gerstoft, and Y. Park, "Gridless DOA estimation via. alternating projections," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brighton, UK, May 2019, pp. 4215–4219.

[25] J. Zhang, D. Rakhimov, and M. Haardt, "Gridless channel estimation for hybrid mmWave MIMO systems via tensor-ESPRIT algorithms in DFT beamspace," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 3, pp. 816–831, April 2021.

[26] M. Wagner, Y. Park, and P. Gerstoft, "Gridless DOA estimation and root-MUSIC for non-uniform linear arrays," *IEEE Transactions on Signal Processing*, vol. 69, pp. 2144–2157, March 2021.

[27] J. Park and S. Boyd, "General heuristics for nonconvex quadratically constrained quadratic programming," *arXiv: Optimization and Control*, 2017.

[28] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, "Applications of second-order cone programming," *Linear Algebra and Its Applications*, vol. 284, no. 1, pp. 193–228, November 1998.

[29] Y. C. Eldar and M. Mishali, "Robust recovery of signals from a structured union of subspaces," *IEEE Transactions on Information Theory*, vol. 55, no. 11, pp. 5302–5316, November 2009.

[30] Y. C. Eldar, P. Kuppinger, and H. Bölcskei, "Block-sparse signals: Uncertainty relations and efficient recovery," *IEEE Transactions on Signal Processing*, vol. 58, no. 6, pp. 3042–3054, June 2010.

[31] C. Steffens and M. Pesavento, "Block- and rank-sparse recovery for direction finding in partly calibrated arrays," *IEEE Transactions on Signal Processing*, vol. 66, no. 2, pp. 384–399, January 2018.

[32] 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, December 2013.