Robust DOA Estimation With Distorted Sensors

XIANG-YU WANG
City University of Hong Kong, Hong Kong

XIAO-PENG LI, Member, IEEE
Shenzhen University, Shenzhen, China

HUIPING HUANG, Graduate Student Member, IEEE
Chalmers University of Technology, Göteborg, Sweden

HING CHEUNG SO, Fellow, IEEE
City University of Hong Kong, Hong Kong

The distorted sensors in an array system will degrade the signal-to-interference-plus-noise ratio of received signal, resulting in performance deterioration. Without knowing the number of source signals, this article focuses on direction-of-arrival (DOA) estimation for a uniform linear array, in which a small fraction of sensors are distorted. Meanwhile, source enumeration and detection of distorted sensors are realized. We model the array system with distorted sensors by introducing unknown gain and phase errors to the output signals, where the observations corresponding to the distorted sensors are treated as outliers. In this way, we tackle the DOA estimation task under the framework of low-rank and row-sparse matrix decomposition. We directly adopt the rank function and $\ell_{2,0}$-norm to obtain the low-rank and row-sparse matrices, respectively, instead of utilizing their surrogates as in the conventional methods. Therefore, the approximation bias is avoided. In detail, rank and $\ell_{2,0}$-norm optimization is converted to ℓ_0-norm minimization. To solve it, we propose a shifted median absolute deviation-based strategy, achieving adaptive hard-thresholding control. The resultant optimization problem is then handled by proximal block coordinate descent, and the convergences of the objective function value and the solution sequence are proved. Extensive simulation results demonstrate the superior performance of the proposed algorithm in terms of DOA estimation, source number estimation, and distorted sensor detection.

Manuscript received 22 December 2023; revised 11 March 2024; accepted 24 April 2024. Date of publication 30 April 2024; date of current version 11 October 2024.

DOI. No. 10.1109/TAES.2024.3395233

Refereeing of this contribution was handled by W. Liu.

This work was supported in part by the Research Grants Council of the Hong Kong SAR, China, under Grant CityU 11207922 and in part by the Young Innovative Talents Project of Guangdong Provincial Department of Education (Natural Science), China, under Grant 2023KQNCX063.

Authors' addresses: Xiang-Yu Wang and Hing Cheung So are with the Department of Electrical Engineering, City University of Hong Kong, Hong Kong, SAR, China, E-mail: (xwang2286-c@my.cityu.edu.hk; hcso@ee.cityu.edu.hk); Xiao-Peng Li is with the State Key Laboratory of Radio Frequency Heterogeneous Integration, Shenzhen University, Shenzhen 518060, China, E-mail: (x.p.li@szu.edu.cn); Huiping Huang is with the Department of Electrical Engineering, Chalmers University of Technology, 41296 Göteborg, Sweden, E-mail: (huiping@chalmers.se). *(Corresponding author: Xiao-Peng Li.)*

NOMENCLATURE

Symbol	*Meaning*
M	No. of array sensors.
M_{distort}	No. of distorted sensors in the array.
Q	No. of source signals.
θ_q	DOA of the q-th signal.
T	No. of snapshots.
N	No. of Monte Carlo trials.

I. INTRODUCTION

Direction-of-arrival (DOA) estimation is a fundamental task in array signal processing, and has versatile applications in radar [1], sonar [2], and wireless communications [3]. Traditional DOA estimation algorithms are based on discrete Fourier transform [4], [5], of which the performance is limited by the sampling resolution of the received signal, viz. Rayleigh criterion [6]. On the other hand, subspace-based methods, like multiple signal classification (MUSIC) [7] and estimation of signal parameters via rotational invariance techniques (ESPRIT) [8], utilize second-order statistics to achieve super-resolution estimation. In recent years, their extensions and variants [9], [10], [11], [12] are also proposed.

The perfect sensor array assumption is usually adopted in the aforementioned algorithms, and the exact knowledge of the array manifold is vital for their DOA estimation. In practice, there may be a large number of sensors in the array [13], [14], which increases the probability of distorted sensor occurrence. When miscalibrated or distorted sensors exist, gain and phase uncertainties are introduced in the array observations. As a result, the array manifold is distorted, and classical DOA estimation algorithms suffer from performance degradation.

There are several works focusing on array signal processing with distorted or failed sensors. With known positions of uncalibrated sensors, joint DOA and sensor phase/gain error estimation is studied in [15] and [16], where the Cramér–Rao bounds are also analyzed. To deal with array errors in practical 5G communication systems, Pan et al. [17] proposed an in situ calibration framework to reduce the DOA estimation error. Exploiting deep learning, indoor positioning with array impairment is realized in [18]. Besides, Yeo and Lu [19] applied genetic algorithm (GA) to handle array failure correction for digital beamforming. However, as the solution search process of GA is rather random, its convergence speed is a major concern, and the computational cost can be demanding. Algorithms utilizing difference coarray are also suggested [20], [21]. In [20], a virtual array is constructed from the impaired sensor array based on the Khatri–Rao product. Nevertheless, the positions of failed sensors are required. In addition, it cannot work when sensor failure happens at the first

and/or last sensors [22]. To deal with the gain and phase uncertainties, Huang et al. [23] tackled DOA estimation based on block sparse representation, where a block steering matrix dictionary is created. One of its limitations is that the dictionary is constructed with discretized spatial angles. If the DOA does not exactly lie at any one of them, this algorithm may not work well.

More recently, DOA estimation has been addressed based on low-rank and sparse matrix decomposition. In [24], to handle DOA estimation and moving target tracking, the authors decompose the array observations as a low-rank matrix and a sparse matrix, corresponding to received signals emitted by stationary and moving sources, respectively. Then, the low rank and sparsity constraints are realized by minimizing the nuclear norm and ℓ_1-norm with the use of linearized alternating direction method. The same DOA tracking model is solved by Bayesian learning in [25].

As for outlier-resistance DOA estimation, low-rank and sparse matrix decomposition theory is widely adopted. The rank of the noise-free observation matrix equals the number of sources, which is usually smaller than the numbers of array sensors and snapshots. Therefore, the low-rank constraint can be employed to remove the sparse outliers [26], [27], [28]. In [26], the signal subspace is obtained by low-rank factorization of the observation matrix, where the residual fitting error is processed by ℓ_p-norm with $1 \leq p < 2$ to suppress the impulsive noise. Although the performance is satisfactory, solving ℓ_p-norm minimization problem is time-consuming. Alternatively, Liu et al. [27] estimate DOAs in the presence of impulsive noise using robust principal component analysis [29]. Due to the nonconvexity of rank minimization, weakly convex function is developed to tackle the low-rank matrix approximation problem, which is able to produce enhanced low-rankness than nuclear norm [30]. It achieves comparable DOA estimation accuracy with fast computation and no prior knowledge of the source number. However, one concern is that the hyperparameters should be tweaked manually. In contrast, capped Frobenius norm [28] is developed with adaptive parameter adjustment. This sparsity-induced norm is suitable to deal with outliers in complex-valued data. Combining with low-rank matrix factorization, the resultant method outperformed most of the existing robust DOA estimation algorithms.

Considering the sensor gain and phase errors, the output signal of distorted sensors can be seen as non-Gaussian outlier. In typical scenarios, the distorted sensors will not be the majority. As a result, the received signal is the superposition of perfect array observations and row-sparse (formulated by $\ell_{2,0}$-norm) outliers in the matrix format. As mentioned before, the observations of a perfect array correspond to a low-rank matrix. Therefore, DOA estimation with distorted sensors can be formulated as low-rank and row-sparse matrix decomposition (LR^{2}SD), where the array observations are the sum of a low-rank matrix and a row-sparse matrix. In [22], the LR^{2}SD model with nuclear norm and $\ell_{2,1}$-norm regularization is solved by iteratively reweighted least squares (IRLS) [31], which avoids singular value decomposition (SVD) and converges with a linear rate [32], [33]. Then, the low-rank solution is used for DOA estimation based on MUSIC, and the distorted sensor detection is conducted according to the obtained row-sparse matrix. We name the corresponding algorithm LR^{2}SD-IRLS hereafter. Moreover, in [22], alternating direction method of multipliers (ADMM) [34], [35], accelerated proximal gradient (APG) [36], and singular value thresholding (SVT) [37] are employed to solve the LR^{2}SD model, which are referred to as LR^{2}SD-ADMM, LR^{2}SD-APG, and LR^{2}SD-SVT, respectively. For LR^{2}SD-IRLS, however, it adopts convex surrogates of rank function and $\ell_{2,0}$-norm, where the introduced approximation error may degrade the performance.

The rank of a matrix corresponds to the ℓ_0-norm of its singular value vector, viz. the number of nonzero singular values. Low rankness means a sparse distribution of singular values. The matrix $\ell_{2,0}$-norm, measuring the row-sparsity, is the ℓ_0-norm of the vector whose elements are the ℓ_2-norm of matrix row vectors. Therefore, LR^{2}SD model maps to the ℓ_0-norm minimization problem in essence. As usual, the nonconvex and noncontinuous ℓ_0-norm is intractable and replaced by various surrogates [38], [39], [40], [41] which are relatively easier to handle.

By contrast, in this article, we avoid using surrogates and directly utilize ℓ_0-norm to tackle the rank and $\ell_{2,0}$-norm regularization. The desired sparse solution regularized by ℓ_0-norm is acquired by a hard-thresholding operation, where only the large-magnitude elements are retained. This mechanism is also utilized to determine the source number and detect the distorted sensors. Then, to tackle distorted sensors with gain/phase errors, we solve the LR^{2}SD problem via proximal block coordinate descent (BCD) [42], where closed-form solution is acquired by alternating minimization. At last, the DOA is estimated using the low-rank solution based on MUSIC, where source enumeration is also fulfilled. With the row-sparse solution seen as the outliers, distorted sensor detection is accomplished. In addition, we establish the convergence of the proposed algorithm for our nonconvex optimization problem.

The main contributions of our work are summarized as follows:

1) Under the framework of LR^{2}SD, robust DOA estimation with distorted sensors is achieved by rank and $\ell_{2,0}$-norm regularization, which can be converted to ℓ_0-norm minimization. We tackle the ℓ_0-norm minimization via hard-thresholding, where the threshold is adaptively determined based on median absolute deviation (MAD). Due to the nonnegativity of optimized variables, the decision threshold characterized by MAD is further shifted by the median. In this way, we fulfill the enhanced sparsity.
2) The source number and positions of distorted sensors are estimated. Employing the proposed ℓ_0-norm minimization strategy, the large magnitudes of the optimized variables are identified, with which the

source number and the locations of the distorted sensors are determined. Next, with the obtained source number, we estimate the DOAs using the acquired low-rank array observations based on MUSIC.

3) The LR^{2}SD model regularized by rank function and $\ell_{2,0}$-norm is solved by proximal BCD, producing closed-form solutions via alternating minimization. Furthermore, the convergences of the objective function value and solution sequence are proved.
4) Extensive simulations are conducted to exhibit the excellent performance of the proposed algorithm in terms of DOA estimation accuracy, source enumeration, and detection rate of distorted sensors.

The rest of this article is organized as follows. Notations and array observation model with distorted sensors are introduced in Section II. The optimization process of our LR^{2}SD model is detailed in Section III, where the DOA estimation, source enumeration, and distorted sensor detection methods are presented. The computational complexity and algorithm convergence analysis are also provided. In Section IV, we include simulation results for algorithm evaluation. Finally, Section V concludes this article.

II. PRELIMINARIES

A. Notations

In this article, scalars are denoted by lowercase or uppercase letters. Vectors and matrices are represented by boldface lowercase and uppercase letters, respectively. For matrices, rank$(\cdot)$, $\|\cdot\|_F$, $(\cdot)^T$, and $(\cdot)^H$ denote matrix rank, Frobenius norm, transpose operator, and Hermitian transpose operator, respectively. For vectors, $\|\cdot\|_0$ and $\|\cdot\|_2$ represent the ℓ_0-norm and ℓ_2-norm, respectively. Given matrix $\mathbf{X} \in \mathbb{C}^{N\times M}$, $\mathbf{X}_{n,:}$, and $\mathbf{X}_{:,m}$ stand for its n-th row and m-th column, respectively. For vector $\mathbf{v}$, its i-th element is denoted as v_i. The $\ell_{2,0}$-norm of $\mathbf{X} \in \mathbb{C}^{N\times M}$ is defined as $\|\mathbf{X}\|_{2,0} = \|[\|\mathbf{X}_{1,:}\|_2, \ldots, \|\mathbf{X}_{n,:}\|_2, \ldots, \|\mathbf{X}_{N,:}\|_2]\|_0$. Operator diag$(\cdot)$ constructs a diagonal matrix, whose diagonal elements are the entries of the input vector. Symbol $\{\cdot^k\}_{k\in\mathbb{N}}$ represents a sequence indexed by integer k. Operators Med$(\cdot)$ and MAD$(\cdot)$ calculate the median and median absolute deviation of a set of entries, respectively. Specifically, j represents the imaginary unit, and $\mathbf{I}$ means the identity matrix. The frequently used symbols in this article are summarized in Table Nomenclature.

B. Ideal Signal Model and MUSIC Algorithm

For a uniform linear array (ULA) of M sensors without distortion, to avoid the phase ambiguity, the intersensor spacing d is set to be equal to half of the wavelength λ. Supposing there are Q far-field narrowband and uncorrelated source signals from distinct directions $\{\theta_q\}_{q=1}^{Q}$ impinging on the array, at time t, we have the following:

$$\bar{\mathbf{x}}(t) = \sum_{q=1}^{Q} \mathbf{a}\left(\theta_q\right) s_q(t) + \mathbf{n}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) \tag{1}$$

where $\bar{\mathbf{x}}(t) \in \mathbb{C}^M$ is the received signal, $\mathbf{a}(\theta_q) \in \mathbb{C}^M$ is the steering vector, and $a_m(\theta_q) = e^{\mathrm{j}2\pi(m-1)\sin(\theta_q)d/\lambda}$. Besides, $\mathbf{A} = [\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_Q)] \in \mathbb{C}^{M\times Q}$ is the steering matrix, $\mathbf{s}(t) \in \mathbb{C}^Q$ is the source signal vector, and $\mathbf{n}(t) \in \mathbb{C}^M$ represents the additive noise vector.

Collecting T snapshots, based on (1), we have the following:

$$\bar{\mathbf{X}} = \mathbf{A}\mathbf{S} + \mathbf{N} \tag{2}$$

where $\bar{\mathbf{X}} = [\bar{\mathbf{x}}(1), \ldots, \bar{\mathbf{x}}(T)] \in \mathbb{C}^{M\times T}$, $\mathbf{S} = [\mathbf{s}(1), \ldots, \mathbf{s}(T)] \in \mathbb{C}^{Q\times T}$, and $\mathbf{N} = [\mathbf{n}(1), \ldots, \mathbf{n}(T)] \in \mathbb{C}^{M\times T}$.

DOA estimation algorithms aim at finding the directions of signals $\{\theta_q\}_{q=1}^{Q}$, among which MUSIC is a well-known one. MUSIC assumes that the noise in (1) is independent and identically distributed white Gaussian process with variance δ^2. With the number of source signals Q as a prior information, it adopts the second-order statistics of the received signal and utilizes the orthogonality of the signal subspace and noise subspace. To be specific, because the signal and noise are uncorrelated, the covariance matrix of $\bar{\mathbf{x}}$ (omitting the time index) is as follows:

$$\mathbf{C}_{\bar{\mathbf{x}}} = E\left(\bar{\mathbf{x}}\bar{\mathbf{x}}^H\right) = \mathbf{A}\mathbf{C}_{\mathbf{s}}\mathbf{A}^H + \delta^2\mathbf{I} \tag{3}$$

where $E(\cdot)$ denotes the expectation operator, and $\mathbf{C}_{\mathbf{s}} = E(\mathbf{s}\mathbf{s}^H) \in \mathbb{R}^{Q\times Q}$. Since the source signals are uncorrelated, $\mathbf{C}_{\mathbf{s}}$ is a diagonal matrix. If we want to differentiate the signals from noise successfully, the diagonal elements of $\mathbf{C}_{\mathbf{s}}$ should be larger than δ^2. Further applying the eigenvalue decomposition (EVD) to $\mathbf{C}_{\bar{\mathbf{x}}}$, we get as follows:

$$\mathbf{C}_{\bar{\mathbf{x}}} = \mathbf{U}_{\mathbf{s}}\mathbf{\Sigma}_{\mathbf{s}}\mathbf{U}_{\mathbf{s}}^H + \delta^2\mathbf{U}_{\mathbf{n}}\mathbf{U}_{\mathbf{n}}^H \tag{4}$$

where the range space spanned by $\mathbf{U}_{\mathbf{s}} \in \mathbb{C}^{M\times Q}$ represents the signal subspace, the subspace spanned by column vectors of $\mathbf{U}_{\mathbf{n}} \in \mathbb{R}^{M\times(M-Q)}$ is named the noise subspace, and $\mathbf{\Sigma}_{\mathbf{s}} \in \mathbb{R}^{Q\times Q}$ is a diagonal matrix whose diagonal elements are the Q largest eigenvalues. It is clear that $\mathbf{U}_{\mathbf{s}}$ in (4) and $\mathbf{A}$ in (3) span the same linear subspace. According to MUSIC, the spatial spectrum is computed as follows:

$$\mathcal{P}(\theta) = \frac{1}{\mathbf{a}^H(\theta)\left(\mathbf{I} - \mathbf{U}_{\mathbf{s}}\mathbf{U}_{\mathbf{s}}^H\right)\mathbf{a}(\theta)},\ \theta \in [-90^\circ, 90^\circ]. \tag{5}$$

Then, DOAs are estimated by peak search of $\mathcal{P}(\theta)$. In practice, we use $\hat{\mathbf{C}}_{\bar{\mathbf{x}}} = 1/T\bar{\mathbf{X}}\bar{\mathbf{X}}^H$ to approximate $\mathbf{C}_{\bar{\mathbf{x}}}$, then perform EVD of $\hat{\mathbf{C}}_{\bar{\mathbf{x}}}$ to obtain $\hat{\mathbf{U}}_{\mathbf{s}}$ and calculate the spatial spectrum.

C. Signal Model With Distorted Sensors

Considering the existence of M_{distort} distorted sensors, which are randomly and sparsely distributed in the array, the array observation becomes as follows:

$$\begin{aligned}\mathbf{X} &= (\mathbf{I} + \mathbf{\Gamma})\mathbf{A}\mathbf{S} + \mathbf{N} \\ &= \mathbf{A}\mathbf{S} + \mathbf{\Gamma}\mathbf{A}\mathbf{S} + \mathbf{N}\end{aligned} \tag{6}$$

where $\mathbf{\Gamma}$ is a diagonal matrix whose diagonal elements γ_m is either 0 or $\alpha_m e^{\mathrm{j}\beta_m}$, denoting perfect sensor or distorted sensor with gain error α_m and phase error β_m [16], [22], [23]. We further denote $\mathbf{L} = \mathbf{A}\mathbf{S}$ and $\mathbf{R} = \mathbf{\Gamma}\mathbf{A}\mathbf{S}$ as the perfect

array observation and outlier matrices, respectively. Then, (6) is rewritten as follows:

$$\mathbf{X} = \mathbf{L} + \mathbf{R} + \mathbf{N}. \tag{7}$$

Apparently, $\text{rank}(\mathbf{L}) = Q < \min(M, T)$, and thus $\mathbf{L}$ is of low rank. From (3) and (4), we have the following:

$$\mathbf{C}_{\mathbf{As}} = E\left(\mathbf{As}\,(\mathbf{As})^H\right) = \mathbf{A}\mathbf{C}_{\mathbf{s}}\mathbf{A}^H \tag{8}$$

and the EVD of $\mathbf{C}_{\mathbf{As}}$ is as follows:

$$\mathbf{C}_{\mathbf{As}} = \mathbf{U}_{\mathbf{s}}\mathbf{\Lambda}_{\mathbf{s}}\mathbf{U}_{\mathbf{s}}^H \tag{9}$$

which means we can use the estimated covariance matrix $\hat{\mathbf{C}}_{\mathbf{As}} = 1/T\mathbf{L}\mathbf{L}^H$ to obtain $\hat{\mathbf{U}}_{\mathbf{s}}$ for DOA estimation. Besides, the rank of $\mathbf{L}$ equals the source signal number. With $\mathbf{L}$, we can also estimate the source number.

For $\mathbf{R}$, as the distorted sensors are sparsely distributed, namely, $M_{\text{distort}} \ll M$, it is row-sparse where the nonzero rows indicate the locations of the distorted sensors.

In this article, we exploit $\mathbf{X}$ to obtain the low-rank $\mathbf{L}$ and row-sparse $\mathbf{R}$ for DOA estimation, source enumeration, and distorted sensor detection.

III. PROPOSED MODEL

A. Algorithm Development

According to (7), the optimization problem is formulated as follows:

$$\begin{aligned} &\min_{\mathbf{L},\mathbf{R}} \ \tau_1 \text{rank}\,(\mathbf{L}) + \tau_2 \|\mathbf{R}\|_{2,0} \\ &\text{s.t. } \mathbf{X} = \mathbf{L} + \mathbf{R} + \mathbf{N} \end{aligned} \tag{10}$$

where $\tau_1 > 0$ and $\tau_2 > 0$ are penalty parameters. Here, independently exploring $\mathbf{L}$ and $\mathbf{R}$ provides us more flexibilities to develop an efficient algorithm, which will be detailed later. Moreover, considering $\mathbf{R} = \mathbf{\Gamma}\mathbf{L}$, the row-sparse regularization term $\|\mathbf{R}\|_{2,0}$ equals $\|\mathbf{\Gamma}\mathbf{L}\|_{2,0}$. For diagonal matrix $\mathbf{\Gamma}$, its diagonal elements are either zero or $\alpha_m e^{\mathrm{j}\beta_m}$, and $\mathbf{L}$ is the noise-free array observation generally with nonzero columns or rows. That is to say, it is $\mathbf{\Gamma}$ that determines term $\|\mathbf{R}\|_{2,0}$, irrelevant with $\mathbf{L}$. Therefore, it is reasonable to deal with $\mathbf{L}$ and $\mathbf{R}$ separately in our algorithm.

Based on (10), converting the constraint to a regularization term, we have the following:

$$\begin{aligned} &\min_{\mathbf{L},\mathbf{R}} \ \mathcal{L}\,(\mathbf{L}, \mathbf{R}) \\ &= \min_{\mathbf{L},\mathbf{R}} \ \tau_1 \text{rank}\,(\mathbf{L}) + \tau_2 \|\mathbf{R}\|_{2,0} + \frac{1}{2}\|\mathbf{X} - \mathbf{L} - \mathbf{R}\|_F^2. \end{aligned} \tag{11}$$

For the ease of representation, we denote $\mathcal{H}(\mathbf{L}, \mathbf{R}) = 1/2\|\mathbf{X} - \mathbf{L} - \mathbf{R}\|_F^2$, which has Lipschitz continuous gradient $l_{\mathcal{H}}$.

Adopting proximal BCD, in the k-th iteration, each variable is updated as follows:

$$\mathbf{L}^{k+1} = \arg\min_{\mathbf{L}} \ \mathcal{L}\left(\mathbf{L}, \mathbf{R}^k\right) + \frac{\phi}{2}\left\|\mathbf{L} - \mathbf{L}^k\right\|_F^2 \tag{12}$$

$$\mathbf{R}^{k+1} = \arg\min_{\mathbf{R}} \ \mathcal{L}\left(\mathbf{L}^{k+1}, \mathbf{R}\right) + \frac{\phi}{2}\left\|\mathbf{R} - \mathbf{R}^k\right\|_F^2 \tag{13}$$

where ϕ is the proximal parameter.

1) *Update* $\mathbf{L}$*:* For subproblem (12), we adopt proximal gradient [43] to solve it. The gradient of $\mathcal{H}(\mathbf{L}, \mathbf{R}^k) + \phi/2\|\mathbf{L} - \mathbf{L}^k\|_F^2$ at $\mathbf{L}^k$ is $\nabla_{\mathbf{L}}\mathcal{H}(\mathbf{L}^k, \mathbf{R}^k)$. The gradient descent step-size is set as ζ, which satisfies $0 < \zeta < 1/l_{\mathcal{H}}$. By proximal gradient descent method, (12) is converted as follows:

$$\begin{aligned} \mathbf{L}^{k+1} = \arg\min_{\mathbf{L}} \ &\tau_1 \text{rank}\,(\mathbf{L}) \\ &+ \frac{1}{2\zeta}\left\|\mathbf{L} - \left(\mathbf{L}^k - \zeta\nabla_{\mathbf{L}}\mathcal{H}\left(\mathbf{L}^k, \mathbf{R}^k\right)\right)\right\|_F^2. \end{aligned} \tag{14}$$

Denoting $\tilde{\mathbf{L}}^k = \mathbf{L}^k - \zeta\nabla_{\mathbf{L}}\mathcal{H}(\mathbf{L}^k, \mathbf{R}^k)$ and setting τ_1 as an adaptive parameter during the iteration, (14) is rewritten as follows:

$$\mathbf{L}^{k+1} = \arg\min_{\mathbf{L}} \ \left(\mu_{\mathbf{L}}^{k+1}\right)^2 \text{rank}\,(\mathbf{L}) + \left\|\mathbf{L} - \tilde{\mathbf{L}}^k\right\|_F^2 \tag{15}$$

where $\mu_{\mathbf{L}}^{k+1} = \sqrt{2\zeta\tau_1^{k+1}}$.

We define the SVD of $\tilde{\mathbf{L}}^k$ as $\tilde{\mathbf{L}}^k = \mathbf{U}_{\tilde{\mathbf{L}}^k}\mathbf{\Sigma}_{\tilde{\mathbf{L}}^k}\mathbf{V}_{\tilde{\mathbf{L}}^k}^H = \mathbf{U}_{\tilde{\mathbf{L}}^k}\text{diag}(\boldsymbol{\sigma}_{\tilde{\mathbf{L}}^k})\mathbf{V}_{\tilde{\mathbf{L}}^k}^H$, where $\boldsymbol{\sigma}_{\tilde{\mathbf{L}}^k}$ is the singular value vector of $\tilde{\mathbf{L}}^k$. According to Von Neumann's trace inequality [44], (15) is equivalent to the following problem [41]:

$$\boldsymbol{\sigma}_{\mathbf{L}^{k+1}} = \arg\min_{\boldsymbol{\sigma}_{\mathbf{L}}} \ \left(\mu_{\mathbf{L}}^{k+1}\right)^2 \|\boldsymbol{\sigma}_{\mathbf{L}}\|_0 + \left\|\boldsymbol{\sigma}_{\mathbf{L}} - \boldsymbol{\sigma}_{\tilde{\mathbf{L}}^k}\right\|_2^2 \tag{16}$$

We introduce Lemma III.1 to handle (16).

LEMMA III.1 (SEE [45] AND [46]) An optimal solution of problem

$$\min_{\mathbf{x}} \ \mu\|\mathbf{x}\|_0 + \|\mathbf{x} - \mathbf{y}\|_2^2$$

is denoted as $\mathbf{x}^*$, whose entry is given as follows:

$$x^* = \mathcal{T}_{\sqrt{\mu}}(y) = \begin{cases} y, & y \geq \sqrt{\mu} \\ 0, & \text{otherwise} \end{cases}$$

where $\mathcal{T}_{\sqrt{\mu}}(y)$ is a thresholding operator.

According to Lemma III.1, the solution of (16) is $\boldsymbol{\sigma}_{\mathbf{L}^{k+1}} = \mathcal{T}_{\mu_{\mathbf{L}}^{k+1}}(\boldsymbol{\sigma}_{\tilde{\mathbf{L}}^k})$. That is to say, $\mu_{\mathbf{L}}^{k+1}$ controls the sparsity of $\boldsymbol{\sigma}_{\mathbf{L}^{k+1}}$ or the rank of $\mathbf{L}^{k+1}$. We determine $\mu_{\mathbf{L}}^{k+1}$ by shifted normalized MAD, viz.

$$\begin{aligned} \mu_{\mathbf{L}}^{k+1} &= \min\left(\mu_{\mathbf{L}}^k, \tilde{\mu}_{\mathbf{L}}^{k+1}\right) \\ \tilde{\mu}_{\mathbf{L}}^{k+1} &= \text{Med}\left(\boldsymbol{\sigma}_{\tilde{\mathbf{L}}^k}\right) + \epsilon_{\mathbf{L}} \times 1.4826 \times \text{MAD}\left(\boldsymbol{\sigma}_{\tilde{\mathbf{L}}^k}\right). \end{aligned} \tag{17}$$

Here, $\left(\text{MAD}\left(\boldsymbol{\sigma}_{\tilde{\mathbf{L}}^k}\right) = \text{Med}\left(\left|\boldsymbol{\sigma}_{\tilde{\mathbf{L}}^k} - \text{Med}\left(\boldsymbol{\sigma}_{\tilde{\mathbf{L}}^k}\right)\right|\right)\right)$, and 1.4826 is the normalization scalar [47]. Because singular values are nonnegative, we shift the normalized MAD by $\text{Med}(\boldsymbol{\sigma}_{\tilde{\mathbf{L}}^k})$ for easier parameter adjustment. Furthermore, $\epsilon_{\mathbf{L}}$ is introduced to control the confidence interval, which differentiates between the large outlier values and small negligible elements. The choice of $\epsilon_{\mathbf{L}}$ is discussed in Section IV-A.

Besides, the nonincreasing property of sequence $\{\mu_{\mathbf{L}}^k\}_{k\in\mathbb{N}}$ is helpful for convergence analysis, which will be detailed in Appendix A. Note that all appendixes are presented in the Supplementary Material.

Then, the solution of (14) is given as follows:

$$\mathbf{L}^{k+1} = \mathbf{U}_{\tilde{\mathbf{L}}^k} \mathrm{diag}\left(\boldsymbol{\sigma}_{\mathbf{L}^{k+1}}\right) \mathbf{V}_{\tilde{\mathbf{L}}^k}^H. \tag{18}$$

For the k-th iteration, the estimated source number $\hat{Q}^{k+1}$ equals the number of elements in $\boldsymbol{\sigma}_{\mathbf{L}^{k+1}}$ which are larger than $\mu_{\mathbf{L}}^{k+1}$.

2) *Update* $\mathbf{R}$*:* For subproblem (13), it is further simplified as follows:

$$\begin{aligned}\mathbf{R}^{k+1} &= \arg\min_{\mathbf{R}} \tau_2 \|\mathbf{R}\|_{2,0} + \frac{1+\phi}{2} \left\|\mathbf{R} - \tilde{\mathbf{R}}^k\right\|_F^2 \\ &= \arg\min_{\mathbf{R}} \left(\mu_{\mathbf{R}}^{k+1}\right)^2 \|\mathbf{R}\|_{2,0} + \left\|\mathbf{R} - \tilde{\mathbf{R}}^k\right\|_F^2 \end{aligned} \tag{19}$$

where $\tilde{\mathbf{R}}^k = \mathbf{X} - \mathbf{L}^{k+1} + /\phi\mathbf{R}^k 1 + \phi$, and $\mu_{\mathbf{R}}^{k+1} = \sqrt{2\tau_2^{k+1}/(1+\phi)}$, similar to $\mu_{\mathbf{L}}^{k+1}$. According to the definition of $\ell_{2,0}$-norm, (19) is converted to

$$\mathbf{r}^{k+1} = \arg\min_{\mathbf{r}} \left(\mu_{\mathbf{R}}^{k+1}\right)^2 \|\mathbf{r}\|_0 + \left\|\mathbf{r} - \tilde{\mathbf{r}}^k\right\|_2^2. \tag{20}$$

Here, $\mathbf{r} = [\|\mathbf{R}_{1,:}\|_2, \ldots, \|\mathbf{R}_{M,:}\|_2]$, and $\tilde{\mathbf{r}}^k$ is defined similarly.

By the same way as solving (16), the solution of (20) is given by $\mathbf{r}^{k+1} = \mathcal{T}_{\mu_{\mathbf{R}}^{k+1}}(\tilde{\mathbf{r}}^k)$, where $\mu_{\mathbf{R}}^{k+1}$ is determined as follows:

$$\begin{aligned}\mu_{\mathbf{R}}^{k+1} &= \min\left(\mu_{\mathbf{R}}^k, \tilde{\mu}_{\mathbf{R}}^{k+1}\right) \\ \tilde{\mu}_{\mathbf{R}}^{k+1} &= \mathrm{Med}\left(\tilde{\mathbf{r}}^k\right) + \epsilon_{\mathbf{R}} \times 1.4826 \times \mathrm{MAD}\left(\tilde{\mathbf{r}}^k\right) \\ \mathrm{MAD}\left(\tilde{\mathbf{r}}^k\right) &= \mathrm{Med}\left(\left|\tilde{\mathbf{r}}^k - \mathrm{Med}\left(\tilde{\mathbf{r}}^k\right)\right|\right). \end{aligned} \tag{21}$$

The choice of $\epsilon_{\mathbf{R}}$ will be discussed in Section IV-A.

After obtaining $\mathbf{r}^{k+1}$, we find the index set $\Omega = \{m : r_m^{k+1} > 0\}$. Then, the solution of (19) is obtained as follows:

$$\mathbf{R}_{m,:}^{k+1} = \begin{cases} \mathbf{0}, & \text{otherwise} \\ \tilde{\mathbf{R}}_{m,:}^k, & m \in \Omega. \end{cases} \tag{22}$$

The procedure for solving (11) is summarized in Algorithm 1. After obtaining the outputs of Algorithm 1, viz. $\mathbf{L}$, $\mathbf{R}$, $\mu_{\mathbf{L}}$, and $\mu_{\mathbf{R}}$, we perform DOA estimation, source enumeration, and distorted sensor detection according to Algorithm 2.

Algorithm 1: Proximal BCD for Solving (11).

Input: Array observation $\mathbf{X} \in \mathbb{C}^{M\times T}$, $\mu_{\mathbf{L}}^0 = 1000$, $\mu_{\mathbf{R}}^0 = 1000$, maximum iteration number K, $\phi = 10^{-5}$, $\epsilon_{\mathbf{L}}$, $\epsilon_{\mathbf{R}}$, ζ.

Initialize: Random matrix $\mathbf{L}^0 \in \mathbb{C}^{M\times T}$, zero matrix $\mathbf{R}^0 \in \mathbb{C}^{M\times T}$.

for $k = 0$ to K **do**

1) Calculate $\tilde{\mathbf{L}}^k$ and its SVD $\tilde{\mathbf{L}}^k = \mathbf{U}_{\tilde{\mathbf{L}}^k}\mathrm{diag}(\boldsymbol{\sigma}_{\tilde{\mathbf{L}}^k})\mathbf{V}_{\tilde{\mathbf{L}}^k}^H$.

2) Update $\mu_{\mathbf{L}}^{k+1}$ according to (17).

3) Update $\boldsymbol{\sigma}_{\mathbf{L}^{k+1}} = \mathcal{T}_{\mu_{\mathbf{L}}^{k+1}}(\boldsymbol{\sigma}_{\tilde{\mathbf{L}}^k})$.

4) Update $\mathbf{L}^{k+1}$ according to (18).

5) Calculate $\tilde{\mathbf{R}}^k$ and $\tilde{\mathbf{r}}^k$.

6) Update $\mu_{\mathbf{R}}^{k+1}$ according to (21).

7) Update $\mathbf{r}^{k+1} = \mathcal{T}_{\mu_{\mathbf{R}}^{k+1}}(\tilde{\mathbf{r}}^k)$.

8) Update $\mathbf{R}^{k+1}$ according to (22).

if $\frac{\|\mathbf{L}^{k+1}-\mathbf{L}^k\|_F}{\|\mathbf{L}^k\|_F} + \frac{\|\mathbf{R}^{k+1}-\mathbf{R}^k\|_F}{\|\mathbf{R}^k\|_F} \le 10^{-3}$ **then**

break

end if

end for

Output: $\mathbf{L}^{k+1}$, $\mathbf{R}^{k+1}$, $\mu_{\mathbf{L}}^{k+1}$, $\mu_{\mathbf{R}}^{k+1}$.

Algorithm 2: DOA Estimation, Source Enumeration, and Distorted Sensor Detection.

Input: $\mathbf{L}$, $\mathbf{R}$, $\mu_{\mathbf{L}}$, $\mu_{\mathbf{R}}$.

1. Source enumeration

1) Calculate the SVD of $\mathbf{L}$, viz. $\mathbf{L} = \mathbf{U}_{\mathbf{L}}\mathrm{diag}(\boldsymbol{\sigma}_{\mathbf{L}})\mathbf{V}_{\mathbf{L}}^H$.

2) Calculate the number of elements in $\boldsymbol{\sigma}_{\mathbf{L}}$, which are larger than $\mu_{\mathbf{L}}$, denoted as $\hat{Q}$.

2. DOA estimation

3) Calculate the estimated correlation matrix $\hat{\mathbf{C}}_{\mathbf{As}} = \frac{1}{T}\mathbf{L}\mathbf{L}^H$.

4) Calculate the SVD of $\hat{\mathbf{C}}_{\mathbf{As}}$, viz. $\hat{\mathbf{C}}_{\mathbf{As}} = \mathbf{U}\mathrm{diag}(\boldsymbol{\sigma})\mathbf{V}^H$.

5) Construct signal subspace $\hat{\mathbf{U}}_{\mathbf{s}} = [\mathbf{U}_{:,1}, \ldots, \mathbf{U}_{:,\hat{Q}}]$.

6) Calculate spectrum $\mathcal{P}(\theta) = \frac{1}{\mathbf{a}^H(\theta)(\mathbf{I}-\hat{\mathbf{U}}_{\mathbf{s}}\hat{\mathbf{U}}_{\mathbf{s}}^H)\mathbf{a}(\theta)}$.

7) Find $\hat{\theta}_{\hat{q}}$ for $\hat{q} = 1, \ldots, \hat{Q}$ via peak search of $\mathcal{P}(\theta)$.

3. Distorted sensor detection

8) Construct $\mathbf{r} = [\|\mathbf{R}_{1,:}\|_2, \ldots, \mathbf{R}_{M,:}]$.

9) Find the distorted sensor index set $\{m_{\mathrm{distort}}\} = \{m : r_m > \mu_{\mathbf{R}}\}$.

Output: Estimated $\{\hat{\theta}_{\hat{q}}\}_{\hat{q}=1}^{\hat{Q}}$ with estimated source number $\hat{Q}$, the index set of detected distorted sensors $\{m_{\mathrm{distort}}\}$.

B. Computational Complexity Analysis

In the k-th iteration, the main computational cost of Algorithm 1 is spent on the SVD of $\tilde{\mathbf{L}}^k$, the calculation of $\mathbf{L}^{k+1}$, and the update of $\mu_{\mathbf{L}}^{k+1}$ and $\mu_{\mathbf{R}}^{k+1}$. The SVD of $\tilde{\mathbf{L}}^k$ consumes at most $\mathcal{O}(TM^2)$ [48] (assuming $M \le T$). The calculation of $\mathbf{L}^{k+1}$ according to (18) costs $\mathcal{O}(\hat{Q}^k TM)$. The update of $\mu_{\mathbf{L}}^{k+1}$ and $\mu_{\mathbf{R}}^{k+1}$ mainly involves the median operation, which consumes $\mathcal{O}(M\log(M))$. In summary, the computational complexity for one iteration is $\mathcal{O}(TM^2)$.

C. Convergence Analysis

We first give Definitions III.1, III.2, and III.3 which will be used in later proofs.

DEFINITION III.1 *(Subdifferentials [49]):* Let $\Psi : \mathbb{R}^n \to (-\infty, +\infty]$ be a proper and lower semicontinuous function. For $\mathbf{x} \in \mathrm{dom}\Psi$, the Fréchet subdifferential of Ψ at $\mathbf{x}$, denoted as $\hat{\partial}\Psi(\mathbf{x})$, is the set of vectors $\mathbf{v} \in \mathbb{R}^n$ which satisfies

$$\lim_{\mathbf{y}\neq\mathbf{x}} \inf_{\mathbf{y}\to\mathbf{x}} \frac{\Psi(\mathbf{y}) - \Psi(\mathbf{x}) - \langle \mathbf{v}, \mathbf{y}-\mathbf{x}\rangle}{\|\mathbf{y}-\mathbf{x}\|} \ge 0.$$

The limiting-subdifferential/subdifferential of Ψ at $\mathbf{x}$ is defined as follows:

$$\partial\Psi(\mathbf{x}) := \left\{\mathbf{v} : \exists\, \mathbf{x}^k \to \mathbf{x}, \Psi\left(\mathbf{x}^k\right) \to \Psi(\mathbf{x})\right.$$

$$\text{and } \mathbf{v}^k \in \hat{\partial}\Psi(\mathbf{x}) \to \mathbf{v} \text{ when } k \to \infty\Big\}.$$

If $\mathbf{x}$ is a local minimizer of Φ, then $0 \in \partial\Psi(\mathbf{x})$.

DEFINITION III.2 *(Distance between an element and a set):* The distance from $\mathbf{x}$ to set S is defined as follows:

$$\operatorname{dist}(\mathbf{x}, S) := \inf\{\|\mathbf{y} - \mathbf{x}\| : \mathbf{y} \in S\}.$$

When S is the empty set, $\operatorname{dist}(\mathbf{x}, S) = \infty$ for all $\mathbf{x}$.

Let $0 < \eta < \infty$, Φ_η denotes the class of all concave and continuous functions $\varphi : [0, \eta) \to \mathbb{R}_+$, which satisfies the following conditions:

1) $\varphi(0) = 0$;
2) φ is C^1 on $(0, \eta)$ and continuous at 0;
3) $\forall\, x \in (0, \eta)$, $\varphi'(x) > 0$.

DEFINITION III.3 *(Kurdyka–Łojasiewicz (KŁ) property [50]):* Let $\Psi : \mathbb{R}^n \to (-\infty, +\infty]$ be proper and lower semicontinuous. The function Ψ is said to have KŁ property at $\bar{\mathbf{x}} \in \operatorname{dom}\partial\Psi := \{\mathbf{x} \in \mathbb{R}^n : \partial\Psi(\mathbf{x}) \neq \emptyset\}$ if there exist $0 < \eta < \infty$, a neighborhood U of $\bar{\mathbf{x}}$, and a function $\varphi \in \Phi_\eta$, such that for $\mathbf{x} \in U \cap \{\mathbf{x} : \Psi(\bar{\mathbf{x}}) < \Psi(\mathbf{x}) < \Psi(\bar{\mathbf{x}}) + \eta\}$, it holds that

$$\varphi'(\Psi(\mathbf{x}) - \Psi(\bar{\mathbf{x}}))\operatorname{dist}(0, \partial\Psi(\mathbf{x})) \geq 1.$$

If Ψ satisfies the KŁ property for all $\mathbf{x} \in \operatorname{dom}\partial\Psi$, then Ψ is a KŁ function.

Next, the objective function value convergence is given by Theorem III.1.

THEOREM III.1 Setting $\mathcal{L}(\mathbf{L}^k, \mu_{\mathbf{L}}^k, \mathbf{R}^k, \mu_{\mathbf{R}}^k)$ as the objective function generated by Algorithm 1, the following statements hold:

1) Sequence $\{\mathcal{L}(\mathbf{L}^k, \mu_{\mathbf{L}}^k, \mathbf{R}^k, \mu_{\mathbf{R}}^k)\}_{k\in\mathbb{N}}$ is nonincreasing;
2) $\mathcal{L}(\mathbf{L}^k, \mu_{\mathbf{L}}^k, \mathbf{R}^k, \mu_{\mathbf{R}}^k)$ is lower bounded.

Therefore, $\{\mathcal{L}(\mathbf{L}^k, \mu_{\mathbf{L}}^k, \mathbf{R}^k, \mu_{\mathbf{R}}^k)\}_{k\in\mathbb{N}}$ converges.

The solution sequence convergence is established in Theorem III.2.

THEOREM III.2 We introduce $\mathbf{Z} := (\mathbf{L}, \mathbf{R})$. For the objective function $\mathcal{L}(\mathbf{Z})$, we assume $\mathcal{L}(\mathbf{Z}) \to \infty$ if and only if $\mathbf{Z} \to \infty$. Furthermore, $\{\mathbf{Z}^k\}_{k\in\mathbb{N}}$ is assumed to be bounded. The following statements hold:

1) There exists a subsequence $\{\mathbf{Z}^{k_p}\}_{p\in\mathbb{N}}$, which converges to a critical point of $\mathcal{L}(\mathbf{Z})$;
2) The objective function $\mathcal{L}(\mathbf{Z})$ satisfies KŁ property, and sequence $\{\mathbf{Z}^k\}_{k\in\mathbb{N}}$ generated by Algorithm 1 converges to a critical point of $\mathcal{L}(\mathbf{Z})$.

The proofs of Theorems III.1 and III.2 are found in Appendixes A and B, respectively.

IV. SIMULATIONS

In this section, we evaluate the performance of the proposed algorithm under various scenarios, and conduct 100 Monte Carlo trials for each scenario. The DOA estimation performance is measured by the root-mean-squared error (RMSE) defined as follows:

$$\text{RMSE} = \sqrt{\frac{1}{NQ}\sum_{n=1}^{N}\sum_{q=1}^{Q}\left(\hat{\theta}_{q,n} - \theta_q\right)^2} \tag{23}$$

where N is the total number of Monte Carlo trials and $\hat{\theta}_{q,n}$ denotes the estimated DOA of q-th source in the n-th trial. For a trial, if $|\hat{\theta}_q - \theta_q| \leq 0.3^\circ$ for $q = 1, \ldots, Q$, it is considered as successful estimation. Then, we introduce another metric, resolution probability, which is defined as the ratio N_{succ}/N, and N_{succ} means the number of successful Monte Carlo trials.

To evaluate the performance of distorted sensor detection, we adopt the probability of distorted sensor detection, which is calculated as $1/N\sum_{n=1}^{N} M_{\text{detect},n}/M_{\text{distort}}$. Here, $M_{\text{detect},n}$ denotes the number of correctly detected distorted sensors in the n-th trial.

In each trial, when generating the array measurements, the positions of the distorted sensors are randomly chosen. The gain error α_m and phase error β_m are sampled from uniform distributions on [0.5, 1.5] and $[-90^\circ, 90^\circ]$, respectively.

A. Investigation of Parameter Selection

When solving the ℓ_0-norm minimization problems (16) and (20), the tunable parameters $\epsilon_{\mathbf{L}}$ in (17) and $\epsilon_{\mathbf{R}}$ in (21) control the sparsity of $\boldsymbol{\sigma}_{\mathbf{L}}$ and $\mathbf{r}$, respectively. Before comparison, we investigate the impacts of $\epsilon_{\mathbf{L}}$ and $\epsilon_{\mathbf{R}}$ on the DOA estimation accuracy. The gradient descent step-size ζ is assigned as 0.03. We set $M = 20$, $M_{\text{distort}} = 3$, $T = 100$, and $Q = 3$ from directions -25°, -10°, and 5°. The RMSEs of DOA estimation at SNR $= -6$ dB, 0 dB, and 6 dB are shown in Fig. 1. It is clear that the proposed method achieves a low RMSE around $\epsilon_{\mathbf{R}} = 2$, whereas the RMSE is not sensitive to the change of $\epsilon_{\mathbf{L}}$. Under the same settings, the probability of distorted sensor detection versus $\epsilon_{\mathbf{L}}$ and $\epsilon_{\mathbf{R}}$ are displayed in Fig. 2. Under different SNRs, we conclude that the change of $\epsilon_{\mathbf{L}}$ does not make an obvious difference in the results. On the other hand, the probability of distorted sensor detection slowly decreases as $\epsilon_{\mathbf{R}}$ increases. Therefore, we set $\epsilon_{\mathbf{R}}$ in the range [1, 3] in the following.

As for the choice of $\epsilon_{\mathbf{L}}$, it directly controls the rank of $\mathbf{L}$, viz. the estimated source signal number. Then, we further investigate the impact of $\epsilon_{\mathbf{L}}$ on the performance of source enumeration. We set $\epsilon_{\mathbf{R}} = 2$, $\zeta = 0.03$, $M = 20$, $M_{\text{distort}} = 3$, $Q = 3$ ($\theta_1 = -25^\circ, \theta_2 = -10^\circ, \theta_3 = 5^\circ$), and $T = 100$. When SNR varies from -9 dB to 9 dB, for different $\epsilon_{\mathbf{L}}$, the number of incorrect source enumeration trials is plotted in Fig. 3. We see that at $\epsilon_{\mathbf{L}} = 3$, our algorithm correctly determines the source number under all SNRs. In the simulation, we choose $\epsilon_{\mathbf{L}}$ in range [1.5, 3] to ensure the correct source enumeration for all situations.

For the gradient descent step-size ζ, we prefer $\zeta \in [0.02, 0.03]$, which is explained in Appendix C.

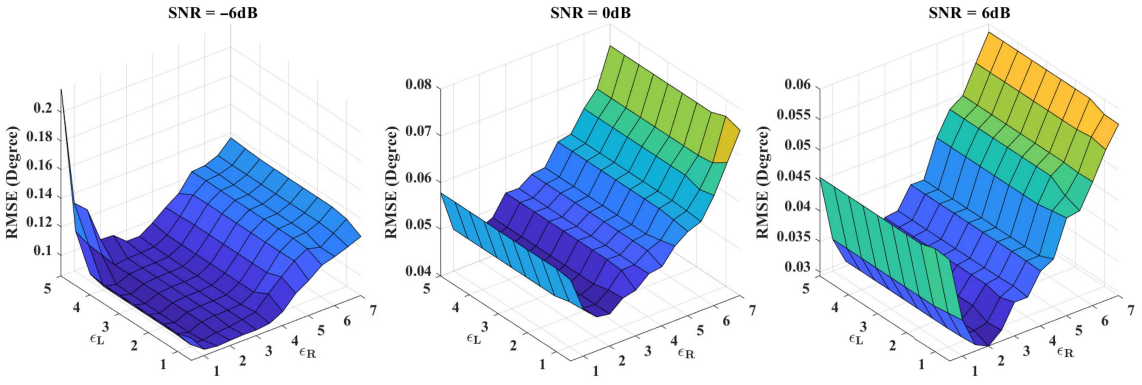

Fig. 1. RMSE versus $\epsilon_{\mathbf{L}}$ and $\epsilon_{\mathbf{R}}$ at SNR $= -6$ dB, 0 dB, and 6 dB, when $\zeta = 0.03$, $M = 20$, $M_{\text{distort}} = 3$, $Q = 3$ ($\theta_1 = -25^\circ, \theta_2 = -10^\circ, \theta_3 = 5^\circ$), and $T = 100$.

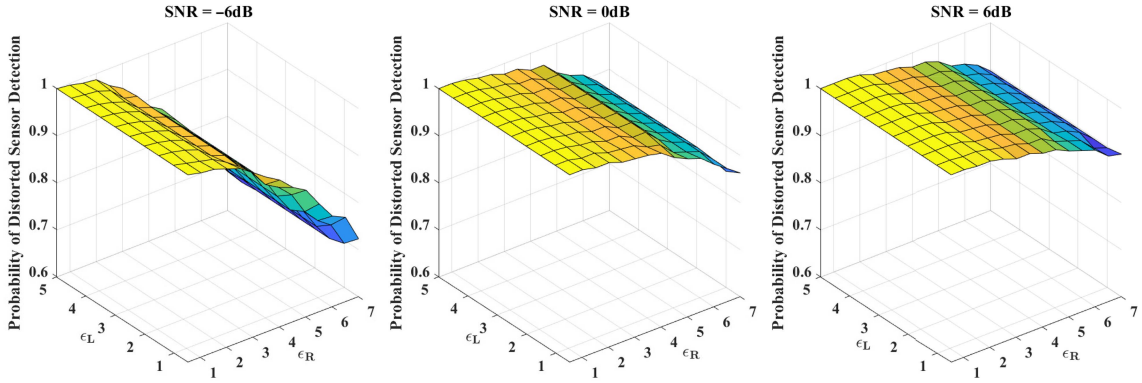

Fig. 2. Probability of distorted sensor detection versus $\epsilon_{\mathbf{L}}$ and $\epsilon_{\mathbf{R}}$ at SNR $= -6$ dB, 0 dB, and 6 dB, when $\zeta = 0.03$, $M = 20$, $M_{\text{distort}} = 3$, $Q = 3$ ($\theta_1 = -25^\circ, \theta_2 = -10^\circ, \theta_3 = 5^\circ$), and $T = 100$.

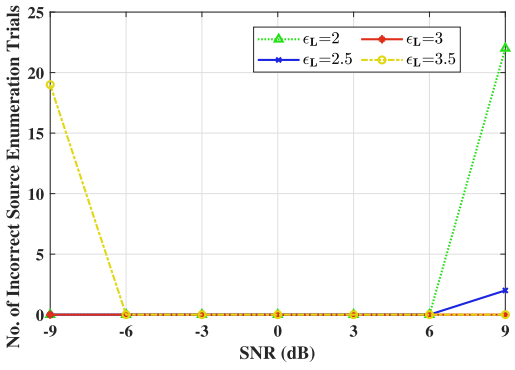

Fig. 3. Number of incorrect source enumeration trials under different SNRs and $\epsilon_{\mathbf{L}}$. Here, $\epsilon_{\mathbf{R}} = 2$, $\zeta = 0.03$, $M = 20$, $M_{\text{distort}} = 3$, $Q = 3$ ($\theta_1 = -25^\circ, \theta_2 = -10^\circ, \theta_3 = 5^\circ$), and $T = 100$.

B. Performance Comparison

We compare the proposed method with Capon [51], ESPRIT [8], MUSIC [7], LR^{2}SD-ADMM, LR^{2}SD-SVT, LR^{2}SD-APG, and LR^{2}SD-IRLS [22] for various situations. The hyperparameters of the compared algorithms are adjusted according to their authors' suggestions. Results for the array without distorted sensors by MUSIC are also included as the benchmark, denoted as MUSIC-Perfect.

We first compare the runtime of LR^{2}SD-ADMM, LR^{2}SD-SVT, LR^{2}SD-APG, LR^{2}SD-IRLS, and the proposed method, where the snapshot number varies from 100 to 500. We set $M = 20$, $M_{\text{distort}} = 3$, SNR $= 0$ dB, and $Q = 3$ from directions -25°, -10°, and 5°. The simulations are implemented in MATLAB R2023b on a computer with a 2.9 GHz i7-10700 CPU and 32 GB RAM. The results are shown in Fig. 4. As we see, for all algorithms, the runtime increases with the snapshot number. LR^{2}SD-IRLS involves the shortest runtime for all snapshot numbers. Our algorithm comes the second. It is because LR^{2}SD-ADMM, LR^{2}SD-SVT, LR^{2}SD-APG, and our algorithm need to compute SVD in each iteration, which is computationally demanding, whereas LR^{2}SD-IRLS avoids calculating SVD and thus is the fastest. The computational complexities of these algorithms are listed in Table I. Our algorithm has the same complexity order as LR^{2}SD-SVT, LR^{2}SD-ADMM, and LR^{2}SD-APG. When $M < T$, LR^{2}SD-IRLS fulfills the smallest complexity. Overall, the runtime of the proposed algorithm is acceptable comparing with LR^{2}SD-ADMM, LR^{2}SD-SVT, and LR^{2}SD-APG. Moreover, the runtime gap between LR^{2}SD-IRLS and the proposed algorithm is ignorable.

In the following, we test the performance of the selected algorithms for various scenarios.

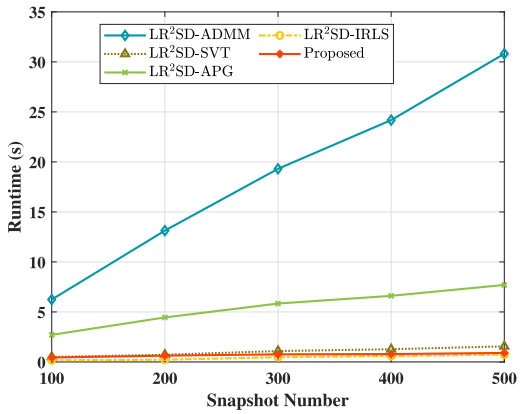

Fig. 4. Runtime of different algorithms versus snapshot number at SNR $= 0$ dB, $M = 20$, $M_{\text{distort}} = 3$, and $Q = 3$ ($\theta_1 = -25^\circ, \theta_2 = -10^\circ, \theta_3 = 5^\circ$).

TABLE I
Computational Complexities of Different Algorithms Per Iteration

Algorithm	Computational complexity
LR^{2}SD-SVT	$\mathcal{O}\left(TM^2\right)$
LR^{2}SD-ADMM	$\mathcal{O}\left(TM^2\right)$
LR^{2}SD-APG	$\mathcal{O}\left(TM^2\right)$
LR^{2}SD-IRLS	$\mathcal{O}\left(M^3\right)$
Proposed	$\mathcal{O}\left(TM^2\right)$

1) *SNR:* We consider a ULA with 20 array sensors, three of which are randomly chosen as distorted sensors. The number of snapshots is set as 100. There are three signal sources from directions -25°, -10°, and 5°. The SNR varies from -9 dB to 9 dB, the RMSEs, resolution probabilities, and probability of distorted sensor detection are plotted in Fig. 5(a), (b), and (c), respectively. We see that our method achieves lower RMSE than other algorithms, and the gap increases with SNR. The resolution probability is also higher than the competing methods. Furthermore, we also find that the DOA estimation accuracy of the proposed method is close to the perfect array result. Our method directly minimizes the ℓ_0-norm and $\ell_{2,0}$-norm. However, other LR^{2}SD based algorithms adopt ℓ_1-norm and $\ell_{2,1}$-norm to approximate the nonconvex ℓ_0-norm and $\ell_{2,0}$-norm, where the approximation gap may cause a relatively low DOA estimation accuracy. ESPRIT, Capon, and MUSIC do not consider the existence of distorted sensors, and their poor performances are predictable.

   As for the probability of distorted sensor detection, our method is slightly worse than other algorithms, for example, at SNR $= 0$ dB and 3 dB. Nonetheless, the correct detection rates are still high, and the gaps are not obvious. Since we adopt different detection strategy from others, using other strategies to improve the correct detection rate is promising, and we leave it as a future work.
2) *Sensor Number:* We test these algorithms for different numbers of array sensors $M = 10 : 5 : 30$ at SNR $= 0$ dB. Here, $M_{\text{distort}} = 3$, $T = 100$, $Q = 3$ with $\theta_1 = -25^\circ$, $\theta_2 = -10^\circ$, and $\theta_3 = 5^\circ$. The results are shown in Fig. 6. In terms of RMSE and resolution probability, our method outperforms the others. For the probability of distorted sensor detection, our method is comparable with the competing algorithms.
3) *Distorted Sensor Number:* When SNR is 0 dB, the performance of the selected algorithms with different numbers of distorted sensors is investigated. We set $M = 20$, $T = 100$, $Q = 3$ with $\theta_1 = -25^\circ$, $\theta_2 = -10^\circ$, and $\theta_3 = 5^\circ$. The results are depicted in Fig. 7. We see that RMSE of the proposed method increases slightly with increasing number of distorted sensors, and the resolution probability even remains as 1. These results showcase the superiority of our algorithm again. As for the distorted sensor detection, the result of our method is not inferior to others.
4) *Source Number:* The DOA estimation and distorted sensor detection accuracies for different numbers of sources are examined. For other parameters, SNR $= 0$ dB, $M = 20$, $M_{\text{distort}} = 3$, and $T = 100$. The simulation results for two sources (from directions -10°, and 5°), three sources (from directions -25°, -10°, and 5°), and four sources (from directions -25°, -10°, 5°, and 20°) are shown in Fig. 8. We see that more sources cause higher RMSEs, and lower resolution probabilities. The proposed algorithm performs better than the others with respect to the DOA estimation accuracy, whereas for distorted sensor detection, the correct detection rates are slightly smaller than the compared algorithms.
5) *Snapshot Number:* Next, we test different algorithms with number of snapshots ranging from 50 to 350 in step-size 50. The other parameters are set as SNR $= 0$ dB, $M = 20$, $M_{\text{distort}} = 3$, $Q = 3$ with $\theta_1 = -25^\circ$, $\theta_2 = -10^\circ$, and $\theta_3 = 5^\circ$. The results are shown in Fig. 9. In general, the large number of snapshots is helpful for DOA estimation and produce a low RMSE and high-resolution probability. Among these algorithms, our method performs the best. The resolution probability even remains as 1 for all numbers of snapshots. The probability of distorted sensor detection is also higher than other algorithms for most numbers of snapshots.
6) *Angular Separation:* We set SNR $= 0$ dB, $M = 20$, $M_{\text{distort}} = 3$, and $T = 100$. As for Q, there are two source signals. One is from direction 0°, whereas the direction of the other varies from 1° to 8° with step-size of 1°. The results are plotted in Fig. 10. We have to admit that most of the compared algorithms outperform the proposed method when the angular separations of the two sources are 1° and 2°. However, with regard to RMSE and resolution probability, our method surpasses the others when the angular separation is larger than 2°. As for the distorted sensor detection, our method also performs

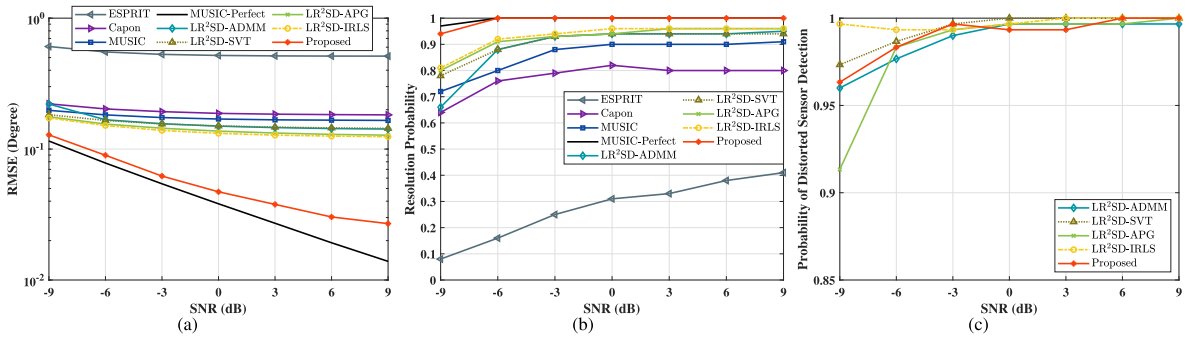

Fig. 5. Performance versus SNR. Here, $M = 20$, $M_{\text{distort}} = 3$, $T = 100$, $Q = 3$ with $\theta_1 = -25^\circ$, $\theta_2 = -10^\circ$, and $\theta_3 = 5^\circ$. (a) RMSE versus SNR. (b) Resolution probability versus SNR. (c) Probability of distorted sensor detection versus SNR.

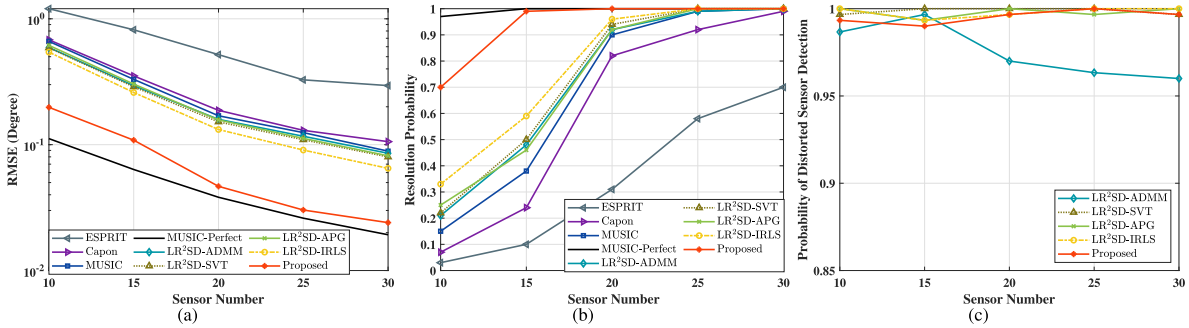

Fig. 6. Performance versus sensor number. We set SNR $= 0$ dB, $M_{\text{distort}} = 3$, $T = 100$, $Q = 3$ with $\theta_1 = -25^\circ$, $\theta_2 = -10^\circ$, and $\theta_3 = 5^\circ$. (a) RMSE versus sensor number. (b) Resolution probability versus sensor number. (c) Probability of distorted sensor detection versus sensor number.

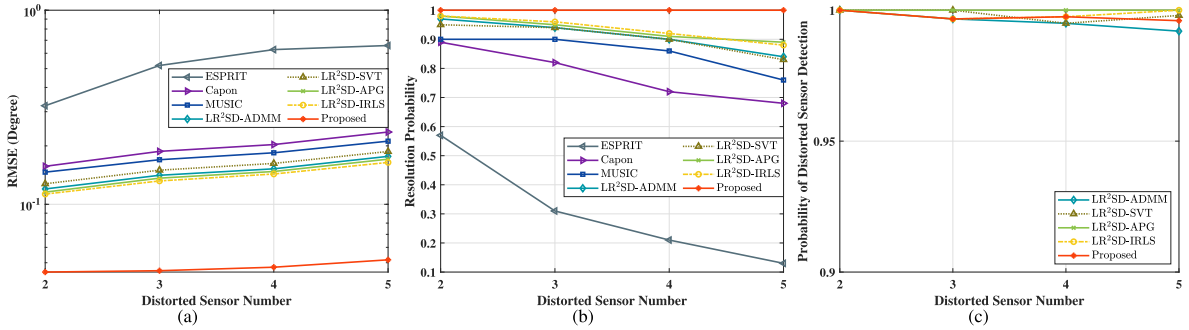

Fig. 7. Performance versus distorted sensor number. For other settings, SNR $= 0$ dB, $M = 20$, $T = 100$, $Q = 3$ with $\theta_1 = -25^\circ$, $\theta_2 = -10^\circ$, and $\theta_3 = 5^\circ$. (a) RMSE versus distorted sensor number. (b) Resolution probability versus distorted sensor number. (c) Probability of distorted sensor detection versus distorted sensor number.

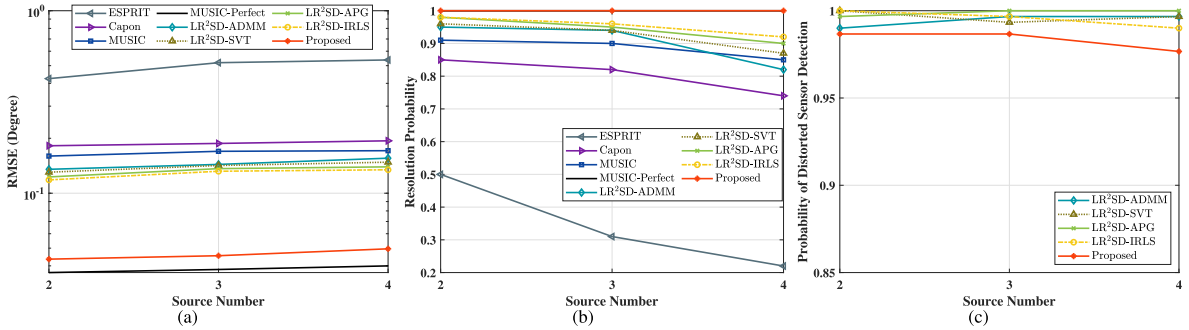

Fig. 8. Performance versus source number. Here, SNR $= 0$ dB, $M = 20$, $M_{\text{distort}} = 3$, $T = 100$. For $Q = 2$, $\theta_1 = -10^\circ$, and $\theta_3 = 5^\circ$. For $Q = 3$, $\theta_1 = -25^\circ$, $\theta_2 = -10^\circ$, and $\theta_3 = 5^\circ$. For $Q = 4$, $\theta_1 = -25^\circ$, $\theta_2 = -10^\circ$, $\theta_3 = 5^\circ$, and $\theta_4 = 20^\circ$. (a) RMSE versus source number. (b) Resolution probability versus source number. (c) Probability of distorted sensor detection versus source number.

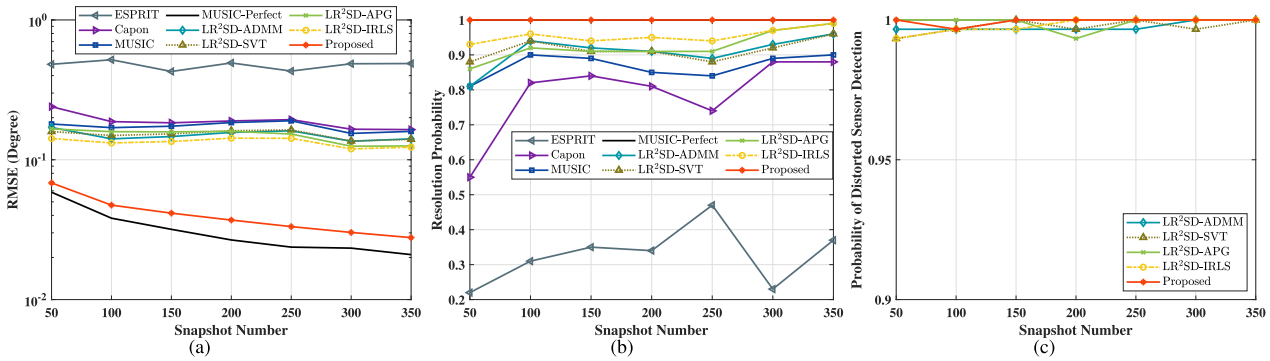

Fig. 9. Performance versus snapshot number. We set SNR $= 0$ dB, $M = 20$, $M_{\text{distort}} = 3$, $Q = 3$ with $\theta_1 = -25^\circ$, $\theta_2 = -10^\circ$, and $\theta_3 = 5^\circ$. (a) RMSE versus snapshot number. (b) Resolution probability versus snapshot number. (c) Probability of distorted sensor detection versus snapshot number.

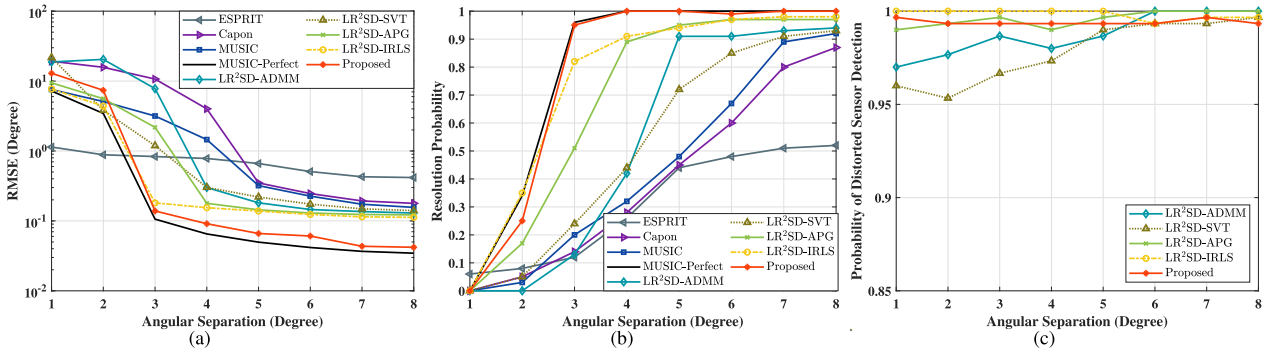

Fig. 10. Performance versus source separation angle. We set SNR $= 0$ dB, $M = 20$, $M_{\text{distort}} = 3$, $T = 100$, $Q = 2$ with $\theta_1 = 0^\circ$, and θ_2 ranges from 1° to 8°. (a) RMSE versus source separation angle. (b) Resolution probability versus source separation angle. (c) Probability of distorted sensor detection versus source separation angle.

well, and the probability of distorted sensor detection is relatively stable with the increasing of angular separation.

V. CONCLUSION

A robust DOA estimation algorithm for an array system with distorted sensors is developed in this article. The array measurements with sensor gain and phase uncertainties are decomposed into the low-rank and row-sparse outlier matrices, which is formulated as rank and $\ell_{2,0}$-norm minimization problem and further converted to ℓ_0-norm minimization. We then directly handle ℓ_0-norm problem via hard-thresholding, where the adaptive decision threshold is determined by a shifted MAD. The solutions are acquired by proximal BCD with guaranteed convergence. With the utilized ℓ_0-norm minimization strategy, we also achieve source number estimation and distorted sensor detection. Our method achieves the highest DOA estimation accuracy among all competing algorithms. With respect to the correct detection rate of distorted sensors, the proposed algorithm is slightly worse than other algorithms.

For our method, distorted sensor detection is based on the vector whose elements are the ℓ_2-norm of rows in the obtained row-sparse matrix. The indices of elements larger than the threshold indicate positions of distorted sensors. A possible future work is to enhance the distorted sensor detection performance via considering other outlier decision rules.

REFERENCES

[1] J. Shi, F. Wen, and T. Liu, "Nested MIMO radar: Coarrays, tensor modeling, and angle estimation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 57, no. 1, pp. 573–585, Feb. 2021.

[2] Y. I. Wu and K. T. Wong, "Acoustic near-field source-localization by two passive anchor-nodes," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 1, pp. 159–169, Jan. 2012.

[3] A. Gorcin and H. Arslan, "A two-antenna single RF front-end DOA estimation system for wireless communications signals," *IEEE Trans. Antennas Propag.*, vol. 62, no. 10, pp. 5321–5333, Oct. 2014.

[4] 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.

[5] S. Burintramart, T. K. Sarkar, Y. Zhang, and M. S.-Palma, "Non-conventional least squares optimization for DOA estimation," *IEEE Trans. Antennas Propag.*, vol. 55, no. 3, pp. 707–714, Mar. 2007.

[6] W. Rueckner and C. Papaliolios, "How to beat the rayleigh resolution limit: A lecture demonstration," *Amer. J. Phys.*, vol. 70, no. 6, pp. 587–594, May 2002.

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

[8] 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.

[9] M. D. Zoltowski, G. M. Kautz, and S. D. Silverstein, "Beamspace root-MUSIC," *IEEE Trans. Signal Process.*, vol. 41, no. 1, p. 344, Jan. 1993.
[10] N. Yuen and B. Friedlander, "Asymptotic performance analysis of ESPRIT, higher order ESPRIT, and virtual ESPRIT algorithms," *IEEE Trans. Signal Process.*, vol. 44, no. 10, pp. 2537–2550, Oct. 1996.
[11] E. D. D. Claudio, R. Parisi, and G. Jacovitti, "Space time MUSIC: Consistent signal subspace estimation for wideband sensor arrays," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2685–2699, May 2018.
[12] J. Lee, J. Park, and J. Chun, "Weighted two-dimensional root MUSIC for joint angle-doppler estimation with MIMO radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 55, no. 3, pp. 1474–1482, Jun. 2019.
[13] L. Gan, W. Jiang, Q. Chen, X. Li, Z. Zhou, and S. Gong, "Method to estimate antenna mode radar cross section of large-scale array antennas," *IEEE Trans. Antennas Propag.*, vol. 69, no. 10, pp. 7029–7034, Oct. 2021.
[14] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 501–513, Apr. 2016.
[15] A. Weiss and B. Friedlander, "DOA and steering vector estimation using a partially calibrated array," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 3, pp. 1047–1057, Jul. 1996.
[16] B. Liao and S. C. Chan, "Direction finding with partly calibrated uniform linear arrays," *IEEE Trans. Antennas Propag.*, vol. 60, no. 2, pp. 922–929, Feb. 2012.
[17] M. Pan et al., "In situ calibration of antenna arrays for positioning with 5G networks," *IEEE Trans. Microw. Theory Tech.*, vol. 71, no. 10, pp. 4600–4613, Oct. 2023.
[18] S. Liu, H. Wang, M. Pan, P. Liu, Y. Ma, and Y. Huang, "5G NR monostatic positioning with array impairments: Data-and-model-driven framework and experiment results," in *Proc. 29th ACM MobiCom Workshop Integr. Sens. Commun. Syst.*, 2023, pp. 1–6.
[19] B.-K. Yeo and Y. Lu, "Array failure correction with a genetic algorithm," *IEEE Trans. Antennas Propag.*, vol. 47, no. 5, pp. 823–828, May 1999.
[20] C. Zhu, W.-Q. Wang, H. Chen, and H. C. So, "Impaired sensor diagnosis, beamforming, and DOA estimation with difference co-array processing," *IEEE Sens. J.*, vol. 15, no. 7, pp. 3773–3780, Jul. 2015.
[21] C.-L. Liu and P. P. Vaidyanathan, "Robustness of difference coarrays of sparse arrays to sensor failures–Part I: A theory motivated by coarray MUSIC," *IEEE Trans. Signal Process.*, vol. 67, no. 12, pp. 3213–3226, Jun. 2019.
[22] H. Huang, Q. Liu, H. C. So, and A. M. Zoubir, "Low-rank and row-sparse decomposition for joint DOA estimation and distorted sensor detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 59, no. 4, pp. 4763–4773, Feb. 2023.
[23] H. Huang, M. Fauß, and A. M. Zoubir, "Block sparsity-based DOA estimation with sensor gain and phase uncertainties," in *Proc. 27th Eur. Signal Process. Conf.*, 2019, pp. 1–5.
[24] B. Lin, J. Liu, M. Xie, and J. Zhu, "Direction-of-arrival tracking via low-rank plus sparse matrix decomposition," *IEEE Antennas Wirel. Propag. Lett.*, vol. 14, pp. 1302–1305, 2015.
[25] A. Das, "A Bayesian sparse-plus-low-rank matrix decomposition method for direction-of-arrival tracking," *IEEE Sens. J.*, vol. 17, no. 15, pp. 4894–4902, Aug. 2017.
[26] W.-J. Zeng, H. C. So, and L. Huang, "ℓ_p-MUSIC: Robust direction-of-arrival estimator for impulsive noise environments," *IEEE Trans. Signal Process.*, vol. 61, no. 17, pp. 4296–4308, Sep. 2013.
[27] Q. Liu, Y. Gu, and H. C. So, "DOA estimation in impulsive noise via low-rank matrix approximation and weakly convex optimization," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 55, no. 6, pp. 3603–3616, Apr. 2019.
[28] X. P. Li, Z. Liu, Z.-L. Shi, and H. C. So, "MUSIC with capped Frobenius norm: Efficient robust direction-of-arrival estimator," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 59, no. 6, pp. 8090–8103, Dec. 2023.
[29] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?," *J. ACM*, vol. 58, no. 3, pp. 1–37, May 2011.
[30] C. Zhang, H. Li, C. Chen, Y. Qian, and X. Zhou, "Enhanced group sparse regularized nonconvex regression for face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 5, pp. 2438–2452, May 2022.
[31] C. Lu, Z. Lin, and S. Yan, "Smoothed low rank and sparse matrix recovery by iteratively reweighted least squares minimization," *IEEE Trans. Image Process.*, vol. 24, no. 2, pp. 646–654, Feb. 2015.
[32] D. Ba, B. Babadi, P. L. Purdon, and E. N. Brown, "Convergence and stability of iteratively re-weighted least squares algorithms," *IEEE Trans. Signal Process.*, vol. 62, no. 1, pp. 183–195, Jan. 2014.
[33] I. Daubechies, R. DeVore, M. Fornasier, and C. S. Güntürk, "Iteratively reweighted least squares minimization for sparse recovery," *Commun. Pure Appl. Math.*, vol. 63, no. 1, pp. 1–38, Oct. 2009.
[34] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jul. 2011.
[35] Z. Lin, M. Chen, and Y. Ma, "The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices," Sep. 2010, *arXiv:1009.5055*.
[36] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imag. Sci.*, vol. 2, no. 1, pp. 183–202, Mar. 2009.
[37] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM J. Optim.*, vol. 20, no. 4, pp. 1956–1982, Mar. 2010.
[38] L. Chen and Y. Gu, "The convergence guarantees of a non-convex approach for sparse recovery," *IEEE Trans. Signal Process.*, vol. 62, no. 15, pp. 3754–3767, Jun. 2014.
[39] L. Chen, X. Jiang, X. Liu, and Z. Zhou, "Logarithmic norm regularized low-rank factorization for matrix and tensor completion," *IEEE Trans. Image Process.*, vol. 30, pp. 3434–3449, 2021.
[40] F. Nie, H. Wang, H. Huang, and C. Ding, "Joint Schatten p-norm and ℓ_p-norm robust matrix completion for missing value recovery," *Knowl. Inf. Syst.*, vol. 42, no. 3, pp. 525–544, Mar. 2015.
[41] X.-Y. Wang, X. P. Li, and H. C. So, "Truncated quadratic norm minimization for bilinear factorization based matrix completion," *Signal Process.*, vol. 214, Jan. 2024, Art. no. 109219.
[42] F. Wen, R. Ying, P. Liu, and T.-K. Truong, "Nonconvex regularized robust PCA using the proximal block coordinate descent algorithm," *IEEE Trans. Signal Process.*, vol. 67, no. 20, pp. 5402–5416, Oct. 2019.
[43] N. Parikh and S. Boyd, "Proximal algorithms," *Found. Trends Optim.*, vol. 1, no. 3, pp. 127–239, 2014.
[44] L. Mirsky, "A trace inequality of John von Neumann," *Monatshefte für Math.*, vol. 79, no. 4, pp. 303–306, 1975.
[45] X. P. Li, Z.-L. Shi, C.-S. Leung, and H. C. So, "Sparse index tracking with k-sparsity or ε-deviation constraint via ℓ_0-norm minimization," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 12, pp. 10 930–10 943, Dec. 2023.
[46] X. P. Li, Z.-L. Shi, Q. Liu, and H. C. So, "Fast robust matrix completion via entry-wise ℓ_0-norm minimization," *IEEE Trans. Cybern.*, vol. 53, no. 11, pp. 7199–7212, Nov. 2023.
[47] A. M. Zoubir, V. Koivunen, E. Ollila, and M. Muma, *Robust Statistics for Signal Processing*. Cambridge, U.K.: Cambridge Univ. Press, 2018.
[48] X. Li, S. Wang, and Y. Cai, "Tutorial: Complexity analysis of singular value decomposition and its variants," Jun. 2019, arXiv:1906.12085.
[49] F. Giannessi, "Variational analysis and generalized differentiation," *J. Optim. Theory Appl.*, vol. 131, no. 2, Nov. 2006, Art. no. 313.
[50] H. Attouch, J. Bolte, P. Redont, and A. Soubeyran, "Proximal alternating minimization and projection methods for nonconvex problems: An approach based on the Kurdyka-Łojasiewicz inequality," *Math. Oper. Res.*, vol. 35, no. 2, pp. 438–457, May 2010.
[51] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, no. 8, pp. 1408–1418, Aug. 1969.

Xiang-Yu Wang received the B.E. degree in electronic engineering from Northwestern Polytechnical University, Xi'an, China, in 2021. He is currently working toward the Ph.D. degree in electronic engineering with the Department of Electrical Engineering, City University of Hong Kong.

His research interests include array signal processing and tensor analysis.

Xiao-Peng Li (Member, IEEE) received the B.Eng. degree as an outstanding graduate in electronic science and technology from Yanshan University, Qinhuangdao, China, in 2015, and the M.Sc. degree with distinction in electronic information engineering, and the Ph.D. degree in electrical engineering from the City University of Hong Kong, Hong Kong SAR, China, in 2018 and 2022, respectively.

From 2018 to 2019, he was a Research Assistant with the Department of Information Engineering, Shenzhen University, Shenzhen, China. From 2022 to 2023, he was a Postdoctoral Fellow with the Department of Electrical Engineering, City University of Hong Kong. He is currently an Assistant Professor with the College of Electronics and Information Engineering, Shenzhen University. His research interests include robust signal processing, sparse recovery, matrix processing, tensor processing, optimization methods, machine learning, and their applications in various areas of engineering, including target estimation, image recovery, video restoration, hyperspectral unmixing, and stock market analysis.

Huiping Huang (Graduate Student Member, IEEE) received the B.Eng. and M.Sc. degrees from Shenzhen University, Shenzhen, China, in 2015 and 2018, respectively, and the Ph.D. degree from Darmstadt University of Technology, Darmstadt, Germany, in 2023, all in electrical and electronic engineering.

He is a Marie Curie postdoctoral Researcher with the Communication Systems Group, Chalmers University of Technology, Gothenburg, Sweden. His research interests include signal processing and wireless communications, with main focus on direction-of-arrival estimation, source localization, channel estimation, robust adaptive beamforming, sparse array design, compressed sensing, and optimization theory with applications to radar, sonar, navigation, microphone array processing, etc.

Hing Cheung So (Fellow, IEEE) was born in Hong Kong. He received the B.Eng. degree in electronic engineering from the City University of Hong Kong, Hong Kong, in 1990, and the Ph.D. degree in electronic engineering from the Chinese University of Hong Kong, Hong Kong, in 1995, respectively.

From 1990 to 1991, he was an Electronic Engineer with the Research and Development Division, Everex Systems Engineering Ltd., Hong Kong. During 1995 to 1996, he was a Postdoctoral Fellow with the Chinese University of Hong Kong. From 1996 to 1999, he was a Research Assistant Professor with the Department of Electronic Engineering, City University of Hong Kong, where he is currently a Professor. His research interests include detection and estimation, fast and adaptive algorithms, multidimensional harmonic retrieval, robust signal processing, source localization, and sparse approximation.

Dr. So has been on the editorial boards of IEEE SIGNAL PROCESSING MAGAZINE (2014–2017), IEEE TRANSACTIONS ON SIGNAL PROCESSING (2010–2014), *Signal Processing* (2010–), and *Digital Signal Processing* (2011–). He was also Lead Guest Editor for IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING, special issue on "Advances in Time/Frequency Modulated Array Signal Processing" in 2017. In addition, he was an elected Member in Signal Processing Theory and Methods Technical Committee (2011–2016) of the IEEE Signal Processing Society where he was Chair in the awards subcommittee (2015–2016).