This article has been accepted for publication in IEEE Transactions on Aerospace and Electronic Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAES.2025.3553467

Joint DOA Estimation and Distorted Sensor Detection

Xiang-Yu Wang, Xiao-Peng Li, *Member, IEEE*, Huiping Huang, *Member, IEEE*, Hing Cheung So, *Fellow, IEEE*

***Abstract*—The exact knowledge of array manifold is vital for finding the direction-of-arrivals (DOAs) of targets, while the sensor gain and phase uncertainties can degrade the estimation performance. Focusing on the array uncertainty induced by distorted sensors, we present a robust DOA estimation algorithm for uniform linear array, where source enumeration and distorted sensor detection are also accomplished. The received array data in the presence of sensor uncertainties are decomposed into a low-rank matrix and a row-sparse component, corresponding to the perfect array observations and errors, respectively. Rather than tackling these two terms in a separate manner, we review their relationship and jointly optimize the perfect array observations and the sparse gain-phase error vector. When formulating the model, variables with the low-rankness or sparsity property are directly regularized by rank function or ℓ_0-norm, instead of their surrogates. We tackle the resultant problem using block proximal linear method so that closed-form solutions to the subproblems are derived. The subsequent ℓ_0-norm optimization is solved via hard-thresholding operator, where the threshold is adaptively determined by our designed scaled quartile scheme. Such ℓ_0-norm minimization scheme also addresses the tasks of source enumeration and distorted sensor detection. Besides, the convergence of our method is proved. To verify its effectiveness, comprehensive simulations are conducted, demonstrating the superiority of the proposed algorithm over the state-of-the-art methods.**

***Index Terms*—array signal processing, gain-phase uncertainty, DOA estimation, source enumeration, distorted sensor detection, block proximal linear, ℓ_0-norm, convergence.**

NOMENCLATURE

$\mathbf{I}$	Identity matrix
e	Euler's number
M	Array sensor number
M_{distort}	Distorted sensor number
N	Monte Carlo trial number
Q	Source number
T	Snapshot number
j	Imaginary unit
$\{\cdot\}_{k\in\mathbb{N}}$	A sequence indexed by integer k
$(\cdot)^H$	Hermitian transpose
$(\cdot)^T$	Transpose

This work was supported in part by the National Natural Science Foundation of China under Grant 62401373, in part by the Young Innovative Talents Project of Guangdong Provincial Department of Education (Natural Science) under Grant 2023KQNCX063. (Corresponding author: Xiao-Peng Li)

X.-Y. Wang and H. C. 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).

X.-P. 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).

H. Huang is with the Department of Electrical Engineering, Chalmers University of Technology, Sweden (e-mail: huiping@chalmers.se).

I. INTRODUCTION

As a fundamental task in array signal processing, direction-of-arrival (DOA) estimation arises in broad applications, like radar [1], [2] and wireless communications [3], [4]. For a uniform linear array (ULA) of M sensors radiated by Q far-field narrow-band and uncorrelated stationary source signals $\{s_q(t)\}_{q=1}^{Q}$, the received observation vector can be expressed as [5]

$$\bar{\mathbf{x}}(t) = \sum_{q=1}^{Q} \mathbf{a}(\theta_q)\, s_q(t) + \mathbf{n}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t). \quad (1)$$

Here, θ_q is the arrival direction of $s_q(t)$, $\mathbf{s}(t) = [s_1(t), \cdots, s_Q(t)]^T \in \mathbb{C}^Q$, $\mathbf{A} = [\mathbf{a}(\theta_1), \cdots, \mathbf{a}(\theta_Q)] \in \mathbb{C}^{M\times Q}$, and $\mathbf{n}(t) = [n_1(t), \cdots, n_M(t)]^T \in \mathbb{C}^M$. Each steering vector has the form of $\mathbf{a}(\theta_q) = [1, \cdots, e^{\mathrm{j}2\pi(m-1)\sin(\theta_q)d/\lambda}, \cdots, e^{\mathrm{j}2\pi(M-1)\sin(\theta_q)d/\lambda}]^T \in \mathbb{C}^M$, where d is the inter-sensor distance and λ is the wavelength. The zero-mean white Gaussian noise components $n_m(t)$, $m = 1, \cdots, M$, are independent and identically distributed with variance δ^2. Based on (1), the DOA estimation task aims to acquire $\{\theta_q\}_{q=1}^{Q}$ utilizing $\bar{\mathbf{x}}(t)$.

In recent decades, great progress has been achieved in DOA estimation, and various super-resolution algorithms are developed. Subspace-based methods, like multiple signal classification (MUSIC) [6] and estimation of signal parameters via rotational invariance techniques (ESPRIT) [7], investigate the eigenstructure of signal covariance matrix. Assuming that $n_m(t)$ is independent of $\mathbf{s}(t)$, the covariance matrix of $\bar{\mathbf{x}}(t)$ is written as

$$\mathbf{R}_{\bar{\mathbf{x}}} = E\left(\bar{\mathbf{x}}(t)\,\bar{\mathbf{x}}^H(t)\right) = \mathbf{A}\mathbf{R}_{\mathbf{s}}\mathbf{A}^H + \delta^2\mathbf{I}, \quad (2)$$

where $E(\cdot)$ is the expectation operator, and $\mathbf{R}_{\mathbf{s}} = E\left(\mathbf{s}(t)\,\mathbf{s}^H(t)\right)$. Since the source signals are uncorrelated, $\mathbf{R}_{\mathbf{s}}$ is a diagonal matrix. Denoting the eigenvalue decomposition of $\mathbf{R}_{\bar{\mathbf{x}}}$ as $\mathbf{R}_{\bar{\mathbf{x}}} = \mathbf{U}_{\bar{\mathbf{x}}}\mathbf{\Lambda}_{\bar{\mathbf{x}}}\mathbf{U}_{\bar{\mathbf{x}}}^H$, MUSIC divides $\mathbf{U}_{\bar{\mathbf{x}}}$ as the signal subspace $\mathbf{U}_{\mathbf{s}} \in \mathbb{C}^{M\times Q}$ and the noise subspace $\mathbf{U}_{\mathbf{n}} \in \mathbb{C}^{M\times(M-Q)}$, that is,

$$\mathbf{R}_{\bar{\mathbf{x}}} = \mathbf{U}_{\mathbf{s}}\mathbf{\Lambda}_{\mathbf{s}}\mathbf{U}_{\mathbf{s}}^H + \mathbf{U}_{\mathbf{n}}\mathbf{\Lambda}_{\mathbf{n}}\mathbf{U}_{\mathbf{n}}^H, \quad (3)$$

where $\mathbf{\Lambda}_{\mathbf{s}} \in \mathbb{C}^{Q\times Q}$ and $\mathbf{\Lambda}_{\mathbf{n}} \in \mathbb{C}^{(M-Q)\times(M-Q)}$ are diagonal matrices, whose diagonal elements are the Q largest and remaining $M-Q$ eigenvalues, respectively. Apparently, the source signal number Q is a prior knowledge here. As the signal subspace and noise subspace are orthogonal, MUSIC estimates DOAs by searching the Q dominant peaks of the spatial spectrum

$$P(\theta) = \frac{1}{\mathbf{a}^H(\theta)\left(\mathbf{I} - \mathbf{U}_\mathbf{s}\mathbf{U}_\mathbf{s}^H\right)\mathbf{a}(\theta)},\ \theta \in [-90^\circ, 90^\circ]. \quad (4)$$

The basic idea of MUSIC inspires many researchers, resulting in various eigenstructure-based DOA algorithms [8]–[11].

Another mainstream finds DOAs in view of compressive sensing [12], which utilizes the spatial sparsity of the source signals. Let $\tilde{\mathbf{A}} = \left[\tilde{\mathbf{a}}(\theta_1), \cdots, \tilde{\mathbf{a}}\left(\theta_{\tilde{Q}}\right)\right]$ be an over-complete steering matrix dictionary, where $\{\theta_{\tilde{q}}\}_{\tilde{q}=1}^{\tilde{Q}}$ is the sampling grid set of all possible source directions, and $\tilde{Q} \gg Q$ represents the grid size. Thus, the received signal can be modeled as [12]

$$\bar{\mathbf{x}}(t) = \tilde{\mathbf{A}}\mathbf{u}(t) + \mathbf{n}(t), \quad (5)$$

and DOA estimation based on compressive sensing (DOA-CS) aims to find a sparse $\mathbf{u}$. In [12], adopting the convex ℓ_1-norm as the sparsity regularizer, DOA-CS utilizing a single time sample is formulated as

$$\mathbf{u}^* = \arg\min_{\mathbf{u}} \left\|\bar{\mathbf{x}} - \tilde{\mathbf{A}}\mathbf{u}\right\|_2^2 + \varpi\|\mathbf{u}\|_1, \quad (6)$$

where $\varpi > 0$ is a penalty parameter, $\|\cdot\|_2$ is the Euclidean norm, and $\|\cdot\|_1$ is the ℓ_1-norm. DOAs correspond to the indices of peaks in $\mathbf{u}^*$. One major concern about (6) is how to construct the dictionary $\tilde{\mathbf{A}}$. If the true directions do not lie on the sampling grids, its performance degrades. While adopting finer grids can improve the estimation accuracy, dense grids lead to a highly coherent matrix, violating the sparse signal recovery condition [13]. To deal with the mismatch of grids and source DOAs, off-grid algorithms are suggested. In [14], the DOAs are not constrained on the grids. A dictionary perturbation caused by grid mismatch is introduced, which is assumed to be Gaussian distributed. However, the Gaussianity may not be satisfied in off-grid DOA estimation. Alternatively, Yang *et al.* [15] assume that the off-grid errors are uniformly distributed and develop a sparse Bayesian algorithm. Utilizing second-order Taylor expansion of the dictionary, [16] estimates the closest grid to the ground truth and their gap. Besides, dictionary learning is also applied to deal with the off-grid error [17], [18], where $\tilde{\mathbf{A}}$ is updated in an iterative manner. In addition, instead of discretizing the spatial domain to construct $\tilde{\mathbf{A}}$, atomic norm minimization approaches are developed [19]–[21]. Treating DOA estimation as a special case of line spectral search [22], continuous dictionary is adopted.

For most existing DOA estimation algorithms, to achieve high-resolution performance, a correctly constructed array manifold is crucial. If there is any bias in the array model, their estimation accuracies decrease. Sensor gain and phase uncertainties are frequently encountered. In reality, sensor manufacturing is not perfect, and the aging rates of sensors within an array may vary. Therefore, the phase delay and gain of each sensor are not the same, leading to gain and phase errors in the array measurement. As such uncertainties exist in each sensor individually, their effects can be described by introducing a diagonal matrix in the array manifold. Then, the received signal with sensor gain and phase uncertainties is [23]

$$\mathbf{x}(t) = \mathbf{\Psi}\left(\mathbf{A}\mathbf{s}(t) + \mathbf{n}(t)\right). \quad (7)$$

The diagonal matrix $\mathbf{\Psi}$ is defined as the sensor uncertainty matrix, and its m-th diagonal element is $\psi_m = g_m e^{\mathrm{j}\phi_m}$, representing the gain and phase of the m-th sensor. Collecting T snapshots, (7) becomes

$$\mathbf{X} = \mathbf{\Psi}\left(\mathbf{A}\mathbf{S} + \mathbf{N}\right), \quad (8)$$

where $\mathbf{X} = [\mathbf{x}(1), \cdots, \mathbf{x}(T)]$, $\mathbf{S} = [\mathbf{s}(1), \cdots, \mathbf{s}(T)]$, and $\mathbf{N} = [\mathbf{n}(1), \cdots, \mathbf{n}(T)]$.

In the presence of array uncertainty, eigenstructure-based methods have also been developed. The covariance matrix of $\mathbf{x}(t)$ is

$$\mathbf{R}_\mathbf{x} = E\left(\mathbf{x}(t)\mathbf{x}^H(t)\right) = \mathbf{\Psi}\mathbf{R}_{\bar{\mathbf{x}}}\mathbf{\Psi}^H \quad (9)$$

For a ULA, $\mathbf{R}_{\bar{\mathbf{x}}}$ is a Toeplitz matrix. Utilizing this property, sensor gains and phases are obtained by solving an equation set [23]. However, compared with the ground truth, the solution is correct up to a rotation. To fix this issue, relative phases between all pairwise sensors should be known. In [24], the authors extend the analysis beyond ULA to the arbitrary geometry case, devising an iterative algorithm to estimate the sensor gains/phases and DOAs simultaneously. Although this algorithm performs well under small gain-phase errors, it converges to sub-optimal solutions when the gain-phase errors are large. To compensate this drawback, Liu *et al.* [25] suggest a method utilizing the eigendecomposition of $\mathbf{R}_\mathbf{x}$, which avoids sub-optimal convergence. However, its high computational complexity and the requirement that the sources should be spatially far away are major concerns. Using the sample covariance matrix, a joint phase and DOA estimation algorithm based on least squares approach is proposed in [26]. Self-calibration is achieved without requiring prior knowledge of the DOA for any of the received sources. Besides, based on ESPRIT, [27] provides both DOAs and unknown sensor gains/phases for partly calibrated ULA in closed-form solutions, where the Cramér-Rao bounds are analyzed.

The gain and phase uncertainties can also be handled under the framework of DOA-CS. According to (5), the received signal in the presence of sensor uncertainty becomes

$$\mathbf{x}(t) = \mathbf{\Psi}\left(\tilde{\mathbf{A}}\mathbf{u}(t) + \mathbf{n}(t)\right). \quad (10)$$

Collecting T snapshots, the matrix format of (10) is:

$$\mathbf{X} = \mathbf{\Psi}\tilde{\mathbf{A}}\mathbf{U} + \mathbf{N}'. \quad (11)$$

Here, $\mathbf{U} = [\mathbf{u}(1), \cdots, \mathbf{u}(T)]$ is a row-sparse matrix and $\mathbf{N}' = \mathbf{\Psi}\mathbf{N}$. Since only the diagonal elements in $\mathbf{\Psi}$ are nonzero, [28] utilizes the sparsity of $\mathbf{\Psi}$, applying ℓ_1-norm and nuclear norm to regularize the row-sparse matrix and uncertainty matrix, respectively, viz.

$$\min \|\mathbf{U}\|_1 + \varpi\|\mathbf{\Psi}\|_*, \quad \text{s.t.} \left\|\mathbf{X} - \mathbf{\Psi}\tilde{\mathbf{A}}\mathbf{U}\right\|_F < \nu, \quad (12)$$

where ν controls the fitting error, $\|\cdot\|_*$ represents the nuclear norm which is the sum of singular values, and $\|\cdot\|_F$ is the Frobenius norm. Similarly, ℓ_1-norm is adopted for uncertainty matrix regularization in [29]. Utilizing errors-in-variables model, Hu *et al.* [30] rewrite (11) as

$$\begin{aligned}\mathbf{X} &= \tilde{\mathbf{A}}\mathbf{U} + (\mathbf{\Psi} - \mathbf{I})\tilde{\mathbf{A}}\mathbf{U} + \mathbf{N}' \\ &= \left(\tilde{\mathbf{A}} + \mathbf{E}\right)\mathbf{U} + \mathbf{N}',\end{aligned} \quad (13)$$

This article has been accepted for publication in IEEE Transactions on Aerospace and Electronic Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAES.2025.3553467

where $\mathbf{E} = (\mathbf{\Psi} - \mathbf{I})\,\tilde{\mathbf{A}}$ is the perturbation and further regularized by the Frobenius norm. The above algorithms are developed with the discretized dictionary. For the continuous dictionary case, a new atomic norm is presented to handle DOAs in the presence of gain-phase uncertainties [31], which is then solved by semidefinite programming.

Aiming to separate the sensor gain and phase errors, (8) becomes [32]–[34]

$$\begin{aligned}\mathbf{X} &= (\mathbf{I} + \mathbf{\Gamma})\,\mathbf{A}\mathbf{S} + \mathbf{N}^{'} \\ &= \mathbf{A}\mathbf{S} + \mathbf{\Gamma}\mathbf{A}\mathbf{S} + \mathbf{N}^{'} \\ &= \mathbf{Z} + \mathbf{R} + \mathbf{N}^{'}, \end{aligned} \tag{14}$$

where $\mathbf{Z} = \mathbf{A}\mathbf{S}$ and $\mathbf{R} = \mathbf{\Gamma}\mathbf{A}\mathbf{S}$, corresponding to the perfect observations and error matrix, respectively. In (14), the sensor uncertainty matrix is decomposed as $\mathbf{\Psi} = \mathbf{I} + \mathbf{\Gamma}$, and diagonal $\mathbf{\Gamma}$ represents the sensor error matrix, viz. $\mathbf{\Gamma} = \mathrm{diag}\,(\boldsymbol{\gamma})$ with gain-phase error $\gamma_m = \alpha_m e^{\mathrm{j}\beta_m}$. Operator $\mathrm{diag}\,(\cdot)$ produces a diagonal matrix with the input vector elements. Generally, with advancements in manufacturing and decrease in defect rate, the imperfect or distorted sensors are not the majority in an array. Hence, it is reasonable to assume that distorted sensors are sparsely and randomly distributed. That is, the gain-phase error vector $\boldsymbol{\gamma}$ is sparse, resulting in $\mathbf{R}$ a row-sparse matrix. Besides, since $\mathbf{A} \in \mathbb{C}^{M\times Q}$ and $\mathbf{S} \in \mathbb{C}^{Q\times T}$ with $Q < \min(M, T)$ in general, $\mathbf{Z}$ is of low rank [35], [36]. If $\mathbf{Z}$ and $\mathbf{R}$ are obtained from $\mathbf{X}$, DOA estimation and distorted sensor detection can be conducted. To be specific, DOAs are able to be obtained by MUSIC. That is, the covariance matrix is calculated as $\hat{\mathbf{R}}_{\bar{\mathbf{x}}} = \frac{1}{T}\mathbf{Z}\mathbf{Z}^H$. Then, the spatial spectrum is computed according to (3) and (4). As for distorted sensor detection, the positions of distorted sensors correspond to the nonzero rows of $\mathbf{R}$.

Utilizing the properties of $\mathbf{Z}$ and $\mathbf{R}$, [32] estimates DOAs and detects the distorted sensors using low-rank and row-sparse matrix decomposition (LR^{2}SD) scheme, namely,

$$\begin{aligned}(\mathbf{Z}^*, \mathbf{R}^*) = \arg\min_{(\mathbf{Z},\mathbf{R})}\ & \frac{1}{2}\left\|\mathbf{X} - \mathbf{Z} - \mathbf{R}\right\|_F^2 \\ & + \varpi_1 \left\|[\mathbf{Z}, \kappa\mathbf{I}]\right\|_* + \varpi_2 \left\|[\mathbf{R}, \kappa\mathbf{1}]\right\|_{2,1}, \end{aligned} \tag{15}$$

where ϖ_1 and ϖ_2 are two tuning parameters, κ is introduced to smooth the nuclear norm and $\ell_{2,1}$-norm [37], and $\mathbf{1}$ represents the all-one vector. Note that the $\ell_{2,1}$-norm $\|\cdot\|_{2,1}$ calculates the sum of ℓ_2-norms of the row vectors. This problem is then solved by iteratively reweighted least squares (IRLS) [37]. In [32], the authors also employ singular value thresholding (SVT) [38], accelerated proximal gradient (APG) [39], and alternating direction method of multipliers (ADMM) [40], [41] to solve this LR^{2}SD problem. In the following, these algorithms are referred to as LR^{2}SD-IRLS, LR^{2}SD-SVT, LR^{2}SD-APG, and LR^{2}SD-ADMM, respectively. It is shown from (15) that [32] adopts the nuclear norm and $\ell_{2,1}$-norm as the substitutes of rank function and $\ell_{2,0}$-norm, respectively. This loosely convex approximation may deviate the solution from the optimality. To avoid the approximation, rank function and $\ell_{2,0}$-norm are exploited to solve LR^{2}SD [34], where the objective function is optimized using proximal block coordinate descent (BCD) [42]. Rank and $\ell_{2,0}$-norm minimization is boiled down to the ℓ_0-norm minimization, and a shifted median absolute deviation strategy is proposed to solve the subsequent ℓ_0-norm minimization problem. We name the algorithm in [34] as LR^{2}SD-BCD-ℓ_0. Comparing with its competitors, LR^{2}SD-BCD-ℓ_0 performs better.

Revisiting (14), it is found that $\mathbf{R} = \mathbf{\Gamma}\mathbf{Z} = \mathrm{diag}\,(\boldsymbol{\gamma})\,\mathbf{Z}$. That is, DOA estimation and distorted sensor detection are coupled. However, [32] and [34] neglect this connection and tackle $\mathbf{Z}$ and $\mathbf{R}$ separately, which may cause performance loss. To deal with this issue, in this paper, we optimize $\mathbf{Z}$ and $\boldsymbol{\gamma}$, performing joint DOA estimation and distorted sensor detection. To mitigate the convex approximation gap, rank function and ℓ_0-norm are directly employed to regularize the perfect array observations and the sparse gain-phase error vector, respectively, where the rank minimization is further reduced to ℓ_0-norm minimization. As we know, matrix rank equals the ℓ_0-norm of the singular value vector. Furthermore, the rank of $\mathbf{Z}$ is the source number, and the indices of nonzero elements in $\boldsymbol{\gamma}$ correspond to the distorted sensor positions. That is, if the ℓ_0-norm is handled properly, we can also achieve source enumeration and distorted sensor detection. Inspired by this insight, an ℓ_0-norm optimization strategy is presented. The main contributions of our work include:

(1) We devise a joint DOA estimation and distorted sensor detection algorithm without the prior knowledge of the source number. The array measurement with sensor gain-phase uncertainty is decomposed into a low-rank matrix and a row-sparse error matrix. These two components are related by the sensor gain-phase error vector. We leverage this inner relationship and optimize the low-rank component and the sparse error vector, which are regularized by rank function and ℓ_0-norm, respectively. Then, hard-thresholding operator is employed to solve the ℓ_0-norm subproblem, where a scaled quartile method is designed to adaptively determine the threshold. Utilizing this optimization scheme, source enumeration and distorted sensor detection are achieved. With the low-rank solution and estimated source number, we conduct DOA estimation by MUSIC.
(2) The objective function is minimized by the prox-linear method, where the subproblems are formulated by linearizing the objective with respect to (w.r.t.) each variable. By doing so, closed-form solutions to the subproblems are derived. Besides, we prove that the objective function value is non-increasing and the solution sequence converges to a critical point.
(3) The effectiveness of our algorithm is verified by extensive simulations in the aspects of DOA estimation, source enumeration, and distorted sensor detection.

We organize the remainder of this paper as follows. The proposed algorithm is detailed in Section II. Analyses of its computational complexity and convergence behavior are also given. We conduct simulations to evaluate the performance of our method and compare with competing algorithms in Section III. Finally, conclusions are drawn in Section IV.

II. Proposed Model

A. Algorithm Development

Noting that $\mathbf{R} = \text{diag}(\boldsymbol{\gamma})\mathbf{Z}$, (14) can be expressed as:

$$\mathbf{X} = \mathbf{Z} + \text{diag}(\boldsymbol{\gamma})\mathbf{Z} + \mathbf{N}', \tag{16}$$

where $\mathbf{Z}$ is of low rank and $\boldsymbol{\gamma}$ is sparse. Then, under the framework of LR^{2}SD, we formulate:

$$\begin{aligned}\min_{\mathbf{Z},\boldsymbol{\gamma}}\ F(\mathbf{Z},\boldsymbol{\gamma}) = \frac{1}{2}\|\mathbf{X} - \mathbf{Z} - \text{diag}(\boldsymbol{\gamma})\mathbf{Z}\|_F^2 \\ + \lambda_1 \text{rank}(\mathbf{Z}) + \lambda_2 \|\boldsymbol{\gamma}\|_0,\end{aligned} \tag{17}$$

where λ_1 and λ_2 are positive penalty parameters, and $\|\cdot\|_0$ computes the number of nonzero entries. To simplify the following presentation, we define $f(\mathbf{Z},\boldsymbol{\gamma}) = \frac{1}{2}\|\mathbf{X} - \mathbf{Z} - \text{diag}(\boldsymbol{\gamma})\mathbf{Z}\|_F^2$. It is easy to verify that $f(\mathbf{Z},\boldsymbol{\gamma})$ is Lipschitz gradient continuous w.r.t. $\mathbf{Z}$ or $\boldsymbol{\gamma}$, whose Lipschitz gradient constants are denoted as $L_{\mathbf{Z}}$ or $L_{\boldsymbol{\gamma}}$, respectively. We find that (17) is a regularized least squares problem for each variable, which can be efficiently solved by block proximal linear (BPL) method [43], [44] as follows.

1) Update $\mathbf{Z}$*:* In the k-th iteration, linearizing $f(\mathbf{Z},\boldsymbol{\gamma})$ w.r.t. $\mathbf{Z}$, $\mathbf{Z}$ is updated via solving the subproblem:

$$\begin{aligned}\mathbf{Z}^k = \arg\min_{\mathbf{Z}}\ \lambda_1 \text{rank}(\mathbf{Z}) + \frac{\mu_{\mathbf{Z}}^k}{2}\left\|\mathbf{Z} - \mathbf{Z}^{k-1}\right\|_F^2 \\ + \left\langle \nabla_{\mathbf{Z}} f\left(\mathbf{Z}^{k-1},\boldsymbol{\gamma}^{k-1}\right), \mathbf{Z} - \mathbf{Z}^{k-1}\right\rangle,\end{aligned} \tag{18}$$

where $\mu_{\mathbf{Z}}^k$ is the proximal parameter in the k-th iteration, and $\nabla_{\mathbf{Z}} f\left(\mathbf{Z}^{k-1},\boldsymbol{\gamma}^{k-1}\right)$ represents $\nabla_{\mathbf{Z}} f(\mathbf{Z},\boldsymbol{\gamma})\big|_{\mathbf{Z}=\mathbf{Z}^{k-1},\boldsymbol{\gamma}=\boldsymbol{\gamma}^{k-1}}$. Without loss of generality, we set $\mu_{\mathbf{Z}}^k = \rho L_{\mathbf{Z}}^k$ with $\rho > 1$.

To accelerate the iterative process, we update $\mathbf{Z}^k$ based on the extrapolated point $\hat{\mathbf{Z}}^{k-1}$ [43]:

$$\hat{\mathbf{Z}}^{k-1} = \mathbf{Z}^{k-1} + \omega_{\mathbf{Z}}^k\left(\mathbf{Z}^{k-1} - \mathbf{Z}^{k-2}\right). \tag{19}$$

Here, $\omega_{\mathbf{Z}}^k$ is the extrapolation weight. Then, (18) is reformulated as

$$\begin{aligned}\mathbf{Z}^k = \arg\min_{\mathbf{Z}}\ \lambda_1 \text{rank}(\mathbf{Z}) + \frac{\mu_{\mathbf{Z}}^k}{2}\left\|\mathbf{Z} - \hat{\mathbf{Z}}^{k-1}\right\|_F^2 \\ + \left\langle \nabla_{\mathbf{Z}} f\left(\hat{\mathbf{Z}}^{k-1},\boldsymbol{\gamma}^{k-1}\right), \mathbf{Z} - \hat{\mathbf{Z}}^{k-1}\right\rangle.\end{aligned} \tag{20}$$

Ignoring the constant term in (20), we get

$$\begin{aligned}\mathbf{Z}^k = \arg\min_{\mathbf{Z}}\ \lambda_1 \text{rank}(\mathbf{Z}) + \frac{\mu_{\mathbf{Z}}^k}{2}\left\|\mathbf{Z} - \hat{\mathbf{Z}}^{k-1}\right\|_F^2 \\ + \left\langle \nabla_{\mathbf{Z}} f\left(\hat{\mathbf{Z}}^{k-1},\boldsymbol{\gamma}^{k-1}\right), \mathbf{Z}\right\rangle,\end{aligned} \tag{21}$$

which is rewritten as

$$\mathbf{Z}^k = \arg\min_{\mathbf{Z}}\ \lambda_1 \text{rank}(\mathbf{Z}) + \frac{\mu_{\mathbf{Z}}^k}{2}\left\|\mathbf{Z} - \tilde{\mathbf{Z}}^{k-1}\right\|_F^2, \tag{22}$$

where $\tilde{\mathbf{Z}}^{k-1} = \hat{\mathbf{Z}}^{k-1} - \frac{1}{\mu_{\mathbf{Z}}^k}\nabla_{\mathbf{Z}} f\left(\hat{\mathbf{Z}}^{k-1},\boldsymbol{\gamma}^{k-1}\right)$. The updates of $\mu_{\mathbf{Z}}^k$ and $\omega_{\mathbf{Z}}^k$ will be detailed later.

Define the singular value decomposition (SVD) of $\tilde{\mathbf{Z}}^{k-1}$ as $\tilde{\mathbf{Z}}^{k-1} = \mathbf{U}_{\tilde{\mathbf{Z}}}^{k-1}\text{diag}\left(\boldsymbol{\sigma}_{\tilde{\mathbf{Z}}}^{k-1}\right)\left(\mathbf{V}_{\tilde{\mathbf{Z}}}^{k-1}\right)^H$. Employing Von Neumann's trace inequality [45], (22) is equivalent to [46], [47]:

$$\boldsymbol{\sigma}_{\mathbf{Z}}^k = \arg\min_{\boldsymbol{\sigma}_{\mathbf{Z}}}\ \lambda_1\|\boldsymbol{\sigma}_{\mathbf{Z}}\|_0 + \frac{\mu_{\mathbf{Z}}^k}{2}\left\|\boldsymbol{\sigma}_{\mathbf{Z}} - \boldsymbol{\sigma}_{\tilde{\mathbf{Z}}}^{k-1}\right\|_2^2, \tag{23}$$

Algorithm 1 BPL for Solving (17)

Input: Received data $\mathbf{X} \in \mathbb{C}^{M\times T}$, $\zeta_{\mathbf{Z}}^0 = 1000$, $\zeta_{\boldsymbol{\gamma}}^0 = 1000$, $\mu_{\mathbf{Z}}^0 = \mu_{\boldsymbol{\gamma}}^0 = 1$, $\mu_{\mathbf{Z}}^{\min} = \mu_{\boldsymbol{\gamma}}^{\min} = 10^{-3}$, maximum iteration number K, $\epsilon_{\mathbf{Z}}$, $\epsilon_{\boldsymbol{\gamma}}$.

Initialize: Zero matrix $\mathbf{Z}^{-1}, \mathbf{Z}^0 \in \mathbb{C}^{M\times T}$, zero vector $\boldsymbol{\gamma}^{-1}, \boldsymbol{\gamma}^0 \in \mathbb{C}^M$.

for $k = 1$ to K **do**

1) Calculate $\mu_{\mathbf{Z}}^k = \max\left(\left\|\text{diag}\left(\boldsymbol{\gamma}^{k-1}\right) + \mathbf{I}\right\|_F^2, \mu_{\mathbf{Z}}^{\min}\right)$.
2) Calculate $\omega_{\mathbf{Z}}^k = \min\left(1, 0.99999\sqrt{\mu_{\mathbf{Z}}^{k-1}/\mu_{\mathbf{Z}}^k}\right)$.
3) Update $\hat{\mathbf{Z}}^{k-1} = \mathbf{Z}^{k-1} + \omega_{\mathbf{Z}}^k\left(\mathbf{Z}^{k-1} - \mathbf{Z}^{k-2}\right)$ and $\tilde{\mathbf{Z}}^{k-1} = \hat{\mathbf{Z}}^{k-1} - \frac{1}{\mu_{\mathbf{Z}}^k}\nabla_{\mathbf{Z}} f\left(\hat{\mathbf{Z}}^{k-1},\boldsymbol{\gamma}^{k-1}\right)$.
4) Calculate SVD of $\tilde{\mathbf{Z}}^{k-1}$ as $\tilde{\mathbf{Z}}^{k-1} = \mathbf{U}_{\tilde{\mathbf{Z}}}^{k-1}\text{diag}\left(\boldsymbol{\sigma}_{\tilde{\mathbf{Z}}}^{k-1}\right)\left(\mathbf{V}_{\tilde{\mathbf{Z}}}^{k-1}\right)^H$.
5) Update $\zeta_{\mathbf{Z}}^k$ according to (26).
6) Update $\mathbf{Z}^k$ as (27).

  if $F\left(\mathbf{Z}^k,\boldsymbol{\gamma}^{k-1}\right) > F\left(\mathbf{Z}^{k-1},\boldsymbol{\gamma}^{k-1}\right)$ **then**
    Set $\omega_{\mathbf{Z}}^k = 0$ and redo Steps 3) to 6).
  end if

7) Calculate $\mu_{\boldsymbol{\gamma}}^k = \max\left(\left\|\boldsymbol{\Phi}\left(\mathbf{Z}^k\right)\right\|_F^2, \mu_{\boldsymbol{\gamma}}^{\min}\right)$.
8) Update $\omega_{\boldsymbol{\gamma}}^k = \min\left(1, 0.99999\sqrt{\mu_{\boldsymbol{\gamma}}^{k-1}/\mu_{\boldsymbol{\gamma}}^k}\right)$.
9) Update $\hat{\boldsymbol{\gamma}}^{k-1} = \boldsymbol{\gamma}^{k-1} + \omega_{\boldsymbol{\gamma}}^k\left(\boldsymbol{\gamma}^{k-1} - \boldsymbol{\gamma}^{k-2}\right)$ and $\tilde{\boldsymbol{\gamma}}^{k-1} = \hat{\boldsymbol{\gamma}}^{k-1} - \frac{1}{\mu_{\boldsymbol{\gamma}}^k}\nabla_{\boldsymbol{\gamma}} g\left(\mathbf{Z}^k,\hat{\boldsymbol{\gamma}}^{k-1}\right)$.
10) Update $\zeta_{\boldsymbol{\gamma}}^k$ according to (34).
11) Update $\boldsymbol{\gamma}^k$ as (35).

  if $F\left(\mathbf{Z}^k,\boldsymbol{\gamma}^k\right) > F\left(\mathbf{Z}^k,\boldsymbol{\gamma}^{k-1}\right)$ **then**
    Set $\omega_{\boldsymbol{\gamma}}^k = 0$, and redo Steps 9) to 11).
  end if
  if $\frac{\|\mathbf{Z}^k - \mathbf{Z}^{k-1}\|_F}{\|\mathbf{Z}^k\|_F} + \frac{\|\boldsymbol{\gamma}^k - \boldsymbol{\gamma}^{k-1}\|_2}{\|\boldsymbol{\gamma}^k\|_2} \le 10^{-3}$ **then**
    Break
  end if

end for

Output: $\mathbf{Z}^k$, $\boldsymbol{\gamma}^k$, $\zeta_{\mathbf{Z}}^k$, $\zeta_{\boldsymbol{\gamma}}^k$.

where λ_1 penalizes the sparsity of $\boldsymbol{\sigma}_{\mathbf{Z}}^k$. To better control its sparsity, λ_1 is tuned during iterations and denoted as λ_1^k. Then, we further simplify (23) as

$$\boldsymbol{\sigma}_{\mathbf{Z}}^k = \arg\min_{\boldsymbol{\sigma}_{\mathbf{Z}}}\ \left(\zeta_{\mathbf{Z}}^k\right)^2\|\boldsymbol{\sigma}_{\mathbf{Z}}\|_0 + \left\|\boldsymbol{\sigma}_{\mathbf{Z}} - \boldsymbol{\sigma}_{\tilde{\mathbf{Z}}}^{k-1}\right\|_2^2, \tag{24}$$

where $\zeta_{\mathbf{Z}}^k = \sqrt{2\lambda_1^k/\mu_{\mathbf{Z}}^k}$.

For ℓ_0-norm minimization, Lemma II.1 is introduced prior to deriving the solution of (24).

Lemma II.1. *[48] Given the problem:*

$$\mathbf{x}^* = \arg\min_{\mathbf{x}}\ c\|\mathbf{x}\|_0 + \|\mathbf{x} - \mathbf{b}\|_2^2,$$

one of its optimal solutions is calculated as $\mathbf{x}^* = \mathcal{H}_{\sqrt{c}}(\mathbf{b})$. *Herein,* $\mathcal{H}_{\sqrt{c}}(\cdot)$ *is an element-wise operator defined as*

$$\mathcal{H}_{\sqrt{c}}(b) = \begin{cases} b, & |b| \ge \sqrt{c}, \\ 0, & otherwise. \end{cases}$$

Algorithm 2 Joint DOA estimation, source enumeration, and distorted sensor detection

Input: $\mathbf{Z}$, $\boldsymbol{\gamma}$.

1. DOA estimation and source enumeration

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

2) Calculate $\bar{\boldsymbol{\sigma}}_\mathbf{Z} = \boldsymbol{\sigma}_\mathbf{Z}/\sum_{m=1}^{M}\left(\sigma_\mathbf{Z}\right)_m$ and count the number of elements of $\bar{\boldsymbol{\sigma}}_\mathbf{Z}$ which is larger than 0.2, denoted as $\hat{Q}$.

3) Calculate $\hat{\mathbf{R}}_\mathbf{Z} = \frac{1}{T}\mathbf{Z}\mathbf{Z}^H$.

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

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

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

7) Find $\hat{\theta}_{\hat{q}}$ for $\hat{q} \in [1, \hat{Q}]$ via peak search of $\hat{P}\left(\theta\right)$.

2. Failed sensor detection

8) Find the distorted sensor index set $\{i_{\text{distort}}\} = \{i | \left|\gamma_i\right| > 0\}$.

Output: Estimated DOA $\hat{\theta}_{\hat{q}}$ for $\hat{q} \in [1, \hat{Q}]$, the index set of distorted sensors $\{i_{\text{distort}}\}$.

According to Lemma II.1, the solution of (24) is given by

$$\boldsymbol{\sigma}_\mathbf{Z}^k = \mathcal{H}_{\zeta_\mathbf{Z}^k}\left(\boldsymbol{\sigma}_{\tilde{\mathbf{Z}}}^{k-1}\right), \tag{25}$$

viz, the elements whose values smaller than $\zeta_\mathbf{Z}^k$ are negligible in the updated $\boldsymbol{\sigma}_\mathbf{Z}^k$.

We propose a scaled quartile strategy to determine $\zeta_\mathbf{Z}^k$ adaptively, viz.

$$\zeta_\mathbf{Z}^k = \min\left(\zeta_\mathbf{Z}^{k-1}, \tilde{\zeta}_\mathbf{Z}^k\right),\ \tilde{\zeta}_\mathbf{Z}^k = \epsilon_\mathbf{Z} \times Q_3\left(\boldsymbol{\sigma}_{\tilde{\mathbf{Z}}}^{k-1}\right), \tag{26}$$

where $Q_3\left(\cdot\right)$ calculates the third quartile of input elements. The non-increasing property of sequence $\left\{\zeta_\mathbf{Z}^k\right\}_{k\in\mathbb{N}}$ avoids the objective increasing when adjusting $\zeta_\mathbf{Z}^k$. The hyperparameter $\epsilon_\mathbf{Z}$ is introduced to scale the third quartile.

Then, $\mathbf{Z}^k$ is updated as

$$\mathbf{Z}^k = \mathbf{U}_{\tilde{\mathbf{Z}}}^{k-1}\mathrm{diag}\left(\boldsymbol{\sigma}_\mathbf{Z}^k\right)\left(\mathbf{V}_{\tilde{\mathbf{Z}}}^{k-1}\right)^H. \tag{27}$$

2) Update $\boldsymbol{\gamma}$: $\boldsymbol{\gamma}$-subproblem is

$$\min_{\boldsymbol{\gamma}}\ \frac{1}{2}\left\|\mathbf{X} - \mathbf{Z} - \mathrm{diag}\left(\boldsymbol{\gamma}\right)\mathbf{Z}\right\|_F^2 + \lambda_2\left\|\boldsymbol{\gamma}\right\|_0. \tag{28}$$

Vectorizing $\mathbf{X} - \mathbf{Z} - \mathrm{diag}\left(\boldsymbol{\gamma}\right)\mathbf{Z}$, (28) becomes:

$$\min_{\boldsymbol{\gamma}}\ \frac{1}{2}\left\|\boldsymbol{\Phi}\left(\mathbf{Z}\right)\boldsymbol{\gamma} - \bar{\boldsymbol{\gamma}}\right\|_2^2 + \lambda_2\left\|\boldsymbol{\gamma}\right\|_0, \tag{29}$$

where $\bar{\boldsymbol{\gamma}} = \mathrm{vec}\left(\mathbf{X} - \mathbf{Z}\right)$ and $\mathrm{vec}\left(\cdot\right)$ constructs a vector by stacking the matrix columns. For $\boldsymbol{\Phi}$, we utilize $\mathrm{vec}\left(\mathbf{EFG}\right) = \left(\mathbf{G}^T \odot \mathbf{E}\right)\mathbf{f}$ with $\mathbf{F} = \mathrm{diag}\left(\mathbf{f}\right)$ and $\odot$ being the Khatri-Rao product. Hence, $\boldsymbol{\Phi}\left(\mathbf{Z}\right) = \mathbf{Z}^T \odot \mathbf{I}$. In the following, $\frac{1}{2}\left\|\boldsymbol{\Phi}\left(\mathbf{Z}\right)\boldsymbol{\gamma} - \bar{\boldsymbol{\gamma}}\right\|_2^2$ is denoted as $g\left(\mathbf{Z}, \boldsymbol{\gamma}\right)$.

Similar to (18), (20), and (22), we linearize $g\left(\mathbf{Z}, \boldsymbol{\gamma}\right)$ w.r.t. $\boldsymbol{\gamma}$ and utilize the extrapolation strategy to update $\boldsymbol{\gamma}$, resulting in

$$\begin{aligned}\boldsymbol{\gamma}^k = &\arg\min_{\boldsymbol{\gamma}}\ \lambda_2\left\|\boldsymbol{\gamma}\right\|_0 + \frac{\mu_{\boldsymbol{\gamma}}^k}{2}\left\|\boldsymbol{\gamma} - \hat{\boldsymbol{\gamma}}^{k-1}\right\|_2^2 \\ &+ \left\langle \nabla_{\boldsymbol{\gamma}} g\left(\mathbf{Z}^k, \hat{\boldsymbol{\gamma}}^{k-1}\right), \boldsymbol{\gamma} - \hat{\boldsymbol{\gamma}}^{k-1}\right\rangle, \end{aligned}\tag{30}$$

$$\Rightarrow$$

$$\begin{aligned}\boldsymbol{\gamma}^k = &\arg\min_{\boldsymbol{\gamma}}\ \lambda_2\left\|\boldsymbol{\gamma}\right\|_0 + \frac{\mu_{\boldsymbol{\gamma}}^k}{2}\left\|\boldsymbol{\gamma} - \hat{\boldsymbol{\gamma}}^{k-1}\right\|_2^2 \\ &+ \left\langle \nabla_{\boldsymbol{\gamma}} g\left(\mathbf{Z}^k, \hat{\boldsymbol{\gamma}}^{k-1}\right), \boldsymbol{\gamma}\right\rangle, \end{aligned}\tag{31}$$

where $\hat{\boldsymbol{\gamma}}^{k-1} = \boldsymbol{\gamma}^{k-1} + \omega_{\boldsymbol{\gamma}}^k\left(\boldsymbol{\gamma}^{k-1} - \boldsymbol{\gamma}^{k-2}\right)$ and $\mu_{\boldsymbol{\gamma}}^k = \rho L_{\boldsymbol{\gamma}}^k$. Simplifying (31) produces

$$\boldsymbol{\gamma}^k = \arg\min_{\boldsymbol{\gamma}}\ \lambda_2\left\|\boldsymbol{\gamma}\right\|_0 + \frac{\mu_{\boldsymbol{\gamma}}^k}{2}\left\|\boldsymbol{\gamma} - \tilde{\boldsymbol{\gamma}}^{k-1}\right\|_2^2. \tag{32}$$

Here, $\tilde{\boldsymbol{\gamma}}^{k-1} = \hat{\boldsymbol{\gamma}}^{k-1} - \frac{1}{\mu_{\boldsymbol{\gamma}}^k}\nabla_{\boldsymbol{\gamma}} g\left(\mathbf{Z}^k, \hat{\boldsymbol{\gamma}}^{k-1}\right)$.

In (32), λ_2 controls the sparsity level of $\boldsymbol{\gamma}$. Similar to (23), we adjust λ_2 during iterations and rewrite (32) as

$$\boldsymbol{\gamma}^k = \arg\min_{\boldsymbol{\gamma}}\ \left(\zeta_{\boldsymbol{\gamma}}^k\right)^2\left\|\boldsymbol{\gamma}\right\|_0 + \left\|\boldsymbol{\gamma} - \tilde{\boldsymbol{\gamma}}^{k-1}\right\|_2^2. \tag{33}$$

where $\zeta_{\boldsymbol{\gamma}}^k = \sqrt{2\lambda_2^k/\mu_{\boldsymbol{\gamma}}^k}$ and is computed as

$$\zeta_{\boldsymbol{\gamma}}^k = \min\left(\zeta_{\boldsymbol{\gamma}}^{k-1}, \tilde{\zeta}_{\boldsymbol{\gamma}}^k\right),\ \tilde{\zeta}_{\boldsymbol{\gamma}}^k = \epsilon_{\boldsymbol{\gamma}} \times Q_3\left(\left|\tilde{\boldsymbol{\gamma}}^{k-1}\right|\right). \tag{34}$$

Hence, the solution of (33) is

$$\boldsymbol{\gamma}^k = \mathcal{H}_{\zeta_{\boldsymbol{\gamma}}^k}\left(\tilde{\boldsymbol{\gamma}}^{k-1}\right). \tag{35}$$

The procedure for solving (17) is summarized in Algorithm 1. The updates of the proximal terms $\mu_\mathbf{Z}^k$, $\mu_{\boldsymbol{\gamma}}^k$ and extrapolation weights $\omega_\mathbf{Z}^k$, $\omega_{\boldsymbol{\gamma}}^k$ are also included, which are similar to strategies in [49].

After obtaining $\mathbf{Z}$ and $\boldsymbol{\gamma}$, we perform DOA estimation, source enumeration, and failed sensor detection according to Algorithm 2. As mentioned before, the source number should be the ℓ_0-norm of $\boldsymbol{\sigma}_\mathbf{Z}$. However, due to the non-convexity of ℓ_0-norm, if we strictly set the nonzero element number equal to the source number, the solution may converge to a suboptimal point as iteration proceeds. Therefore, we retain slightly more nonzero elements when updating $\boldsymbol{\sigma}_\mathbf{Z}^k$, which is achieved by tuning $\epsilon_\mathbf{Z}$. When performing source enumeration after stopping the iterations, $\boldsymbol{\sigma}_\mathbf{Z}$ is normalized and only elements larger than a threshold, which is set as 0.2 here, are counted. More details are included in Algorithm 2. The influences of different $\epsilon_\mathbf{Z}$ and $\epsilon_{\boldsymbol{\gamma}}$ values on the source enumeration and distorted sensor detection performance are investigated in Section III-A.

B. Computational Complexity

We assume $M \leq T$ in the analysis. At each iteration, when formulating $\mathbf{Z}$- and $\boldsymbol{\gamma}$-subproblems, we mainly need to compute $\left\|\mathrm{diag}\left(\boldsymbol{\gamma}^{k-1}\right) + \mathbf{I}\right\|_F^2$, $\nabla_\mathbf{Z} f\left(\hat{\mathbf{Z}}^{k-1}, \boldsymbol{\gamma}^{k-1}\right)$, $\left\|\boldsymbol{\Phi}\left(\mathbf{Z}^k\right)\right\|_F^2$, and $\nabla_{\boldsymbol{\gamma}} g\left(\mathbf{Z}^k, \hat{\boldsymbol{\gamma}}^{k-1}\right)$, which consumes $\mathcal{O}\left(TM^2\right)$. For the update of $\mathbf{Z}^k$, the main cost is spent on the SVD, leading to $\mathcal{O}\left(TM^2\right)$ [50]. Updating $\zeta_\mathbf{Z}^k$ and $\zeta_{\boldsymbol{\gamma}}^k$ needs to calculate quartiles, corresponding to $\mathcal{O}\left(M\log(M)\right)$. Totally, Algorithm 1 costs $\mathcal{O}\left(TM^2\right)$ per iteration.

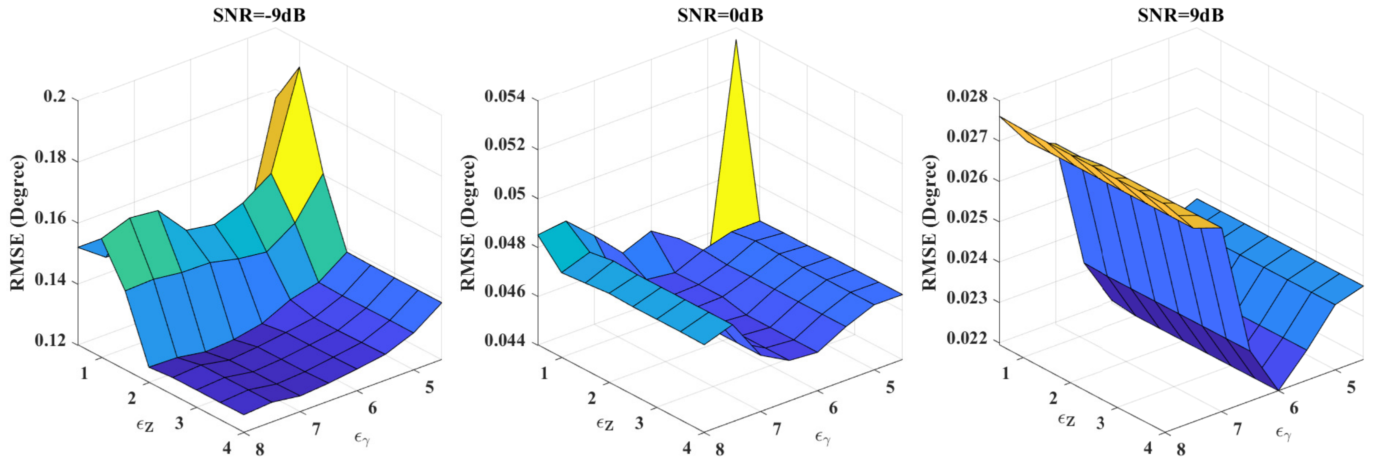

Fig. 1: RMSE versus $\epsilon_{\mathbf{Z}}$ and ϵ_{γ} at SNR = -9 dB, 0 dB and 9 dB, when $M = 20$, $M_{\text{distort}} = 3$, $Q = 3$ ($\theta_1 = -25°, \theta_2 = -10°, \theta_3 = 5°$), $T = 100$.

C. Convergence Analysis

The convergence behavior is summarized in Theorems II.1 and II.2.

Theorem II.1. *Denote* $\mathbf{L} := (\mathbf{Z}, \boldsymbol{\gamma})$. *The objective sequence* $\left\{F\left(\mathbf{L}^k\right)\right\}_{k\in\mathbb{N}}$ *generated by Algorithm 1 is non-increasing and lower-bounded.*

Theorem II.2. *Assume* $F\left(\mathbf{L}^k\right) \to \infty$ *if and only if* $\|\mathbf{L}\|_F \to \infty$ *and* $\mathbf{L}^0$ *is sufficiently close to the limit point of sequence* $\left\{\mathbf{L}^k\right\}_{k\in\mathbb{N}}$. *Then,* $\left\{\mathbf{L}^k\right\}_{k\in\mathbb{N}}$ *converges to a critical point of* $F\left(\mathbf{L}\right)$.

The proofs of Theorems II.1 and II.2 are provided in Appendices A and B, respectively.

III. Experimental Results

In this section, we conduct extensive simulations to evaluate the performance of the proposed algorithm. The number of Monte Carlo trials for each experimental setting is 100. In each trial, for an M-sensor array, M_{distort} sensors are randomly chosen as the distorted sensors, of which the gain error α_m and phase error β_m are sampled from uniform distributions on $[0.3, 2]$ and $[-90°, 90°]$, respectively.

For performance metrics, we adopt the root-mean squared error (RMSE) to measure DOA estimation accuracy, which is defined as

$$\text{RMSE} = \sqrt{\frac{1}{NQ}\sum_{n=1}^{N}\sum_{q=1}^{Q}\left(\hat{\theta}_{q,n} - \theta_q\right)^2}, \tag{36}$$

where $\hat{\theta}_{q,n}$ represents the estimated DOA of the q-th signal in the n-th trial, and N means the total Monte Carlo trial number. At each independent run, if $\left|\hat{\theta}_{q,n} - \theta_q\right| < 0.3°$ for any q, we classify it as a successful trial. To further assess DOA estimation, resolution probability is also investigated and calculated as N_{succ}/N. Here, N_{succ} denotes the number of successful trials. For distorted sensor detection, we use probability of distorted sensor detection [32], [34] computed as $\frac{1}{N}\sum_{n=1}^{N} M_{\text{detect},n}/M_{\text{distort}}$, where $M_{\text{detect},n}$ means the correctly detected distorted sensor number in the n-th trial.

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	$\mathcal{O}\left(M^3\right)$
LR^{2}SD-BCD-ℓ_0	$\mathcal{O}\left(TM^2\right)$
LR^{2}SD-Entangle-IRLS	$\mathcal{O}\left(M^3\right)$
Proposed	$\mathcal{O}\left(TM^2\right)$

A. Choice of Hyperparamters

As shown in Algorithm 1, $\epsilon_{\mathbf{Z}}$ and ϵ_{γ} are involved in threshold computation and vital for sparsity control. Therefore, we investigate the influences of their choices on the performance under different signal-to-noise ratios (SNRs), where the array sensor number is 20 (three of which are distorted), snapshot number is 100, and source number is 3 from directions $-25°$, $-10°$, $5°$.

The RMSE versus $\epsilon_{\mathbf{Z}}$ and ϵ_{γ} is plotted in Fig. 1. We see that for low SNRs (i.e., SNR = -9 dB and 0 dB), RMSE is more sensitive to the change of $\epsilon_{\mathbf{Z}}$. Whereas for SNR = 9 dB, the RMSE change induced by $\epsilon_{\mathbf{Z}}$ is small. For the proposed algorithm, DOA estimation performance is associated with the solution $\mathbf{Z}$. The parameter $\epsilon_{\mathbf{Z}}$ is involved in threshold determination when updating the singular value vector of $\mathbf{Z}$. Therefore, selection of $\epsilon_{\mathbf{Z}}$ is crucial for DOA estimation accuracy, especially when the Gaussian noise is strong. As we know, $\mathbf{Z}$ is a low-rank matrix, and $\boldsymbol{\sigma}_{\mathbf{Z}}^k$ should be a sparse vector. For SNR = -9 dB and 0 dB, the element magnitudes of $\boldsymbol{\sigma}_{\tilde{\mathbf{Z}}}^{k-1}$ are disturbed by strong noise. A small $\epsilon_{\mathbf{Z}}$ may not wipe the nonzero elements induced by noise, which results in many nonzero singular values in $\boldsymbol{\sigma}_{\mathbf{Z}}^k$ and overestimates the source number. Hence, DOA estimation performance degrades. We have to increase $\epsilon_{\mathbf{Z}}$ to handle this scenario. When the Gaussian noise is weak, i.e., SNR = 9 dB, the magnitude disturbance in $\boldsymbol{\sigma}_{\tilde{\mathbf{Z}}}^{k-1}$ is negligible. Even a small $\epsilon_{\mathbf{Z}}$ can effectively eliminate the mild magnitude disturbance. On the

This article has been accepted for publication in IEEE Transactions on Aerospace and Electronic Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAES.2025.3553467

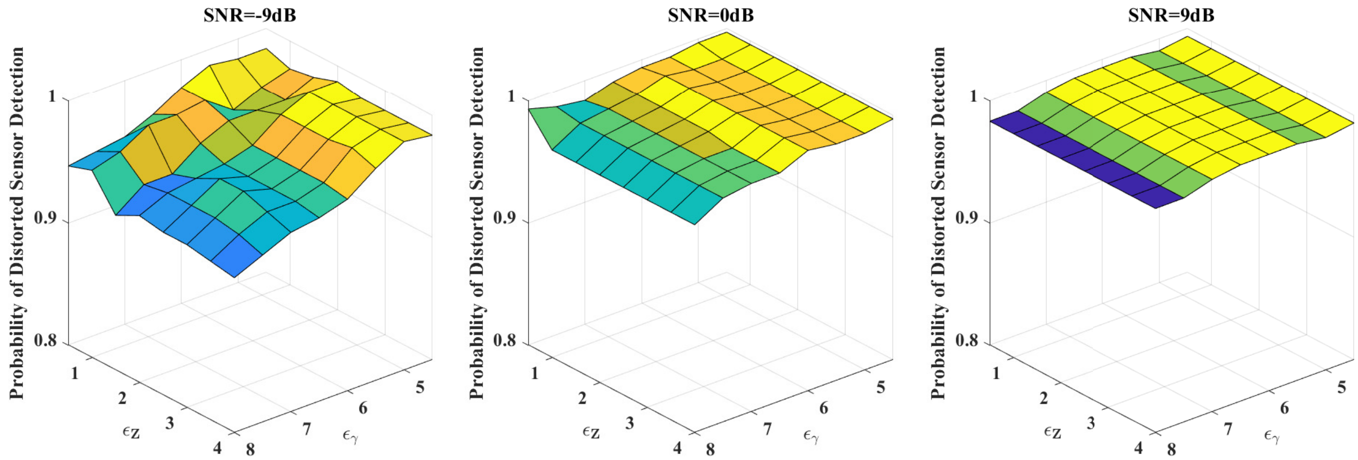

Fig. 2: Probability of distorted sensor detection versus $\epsilon_{\mathbf{Z}}$ and ϵ_γ at SNR = -9 dB, 0 dB and 9 dB, when $M = 20$, $M_{\text{distort}} = 3$, $Q = 3$ ($\theta_1 = -25°, \theta_2 = -10°, \theta_3 = 5°$), $T = 100$.

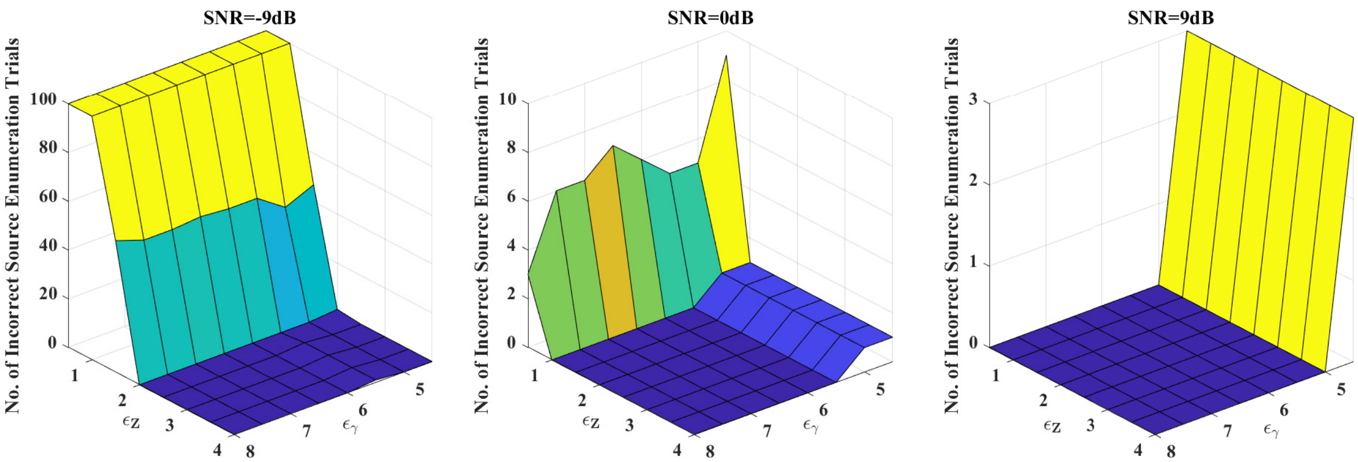

Fig. 3: Number of incorrect source enumeration trials versus $\epsilon_{\mathbf{Z}}$ and ϵ_γ at SNR = -9 dB, 0 dB and 9 dB, when $M = 20$, $M_{\text{distort}} = 3$, $Q = 3$ ($\theta_1 = -25°, \theta_2 = -10°, \theta_3 = 5°$), $T = 100$.

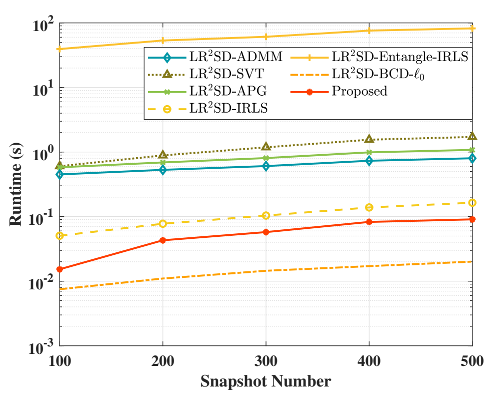

Fig. 4: Runtimes of different algorithms versus snapshot number at SNR = 0 dB, $M = 20$, $M_{\text{distort}} = 3$, $Q = 3$ ($\theta_1 = -25°, \theta_2 = -10°, \theta_3 = 5°$).

other hand, ϵ_γ is related to the update of $\boldsymbol{\gamma}$. Since $\mathbf{Z}$ and $\boldsymbol{\gamma}$ are iteratively optimized, the accuracies of the solutions $\boldsymbol{\gamma}$ and $\mathbf{Z}$ are related. Thus, ϵ_γ has an indirect impact on DOA estimation performance. Nevertheless, compared to $\epsilon_{\mathbf{Z}}$, the influence of ϵ_γ is minor. When the RMSE versus $\epsilon_{\mathbf{Z}}$ is relatively stable, it can be observed that ϵ_γ does not cause significant fluctuations in RMSE. For example, at SNR = 9 dB, the RMSE change induced by ϵ_γ is within $0.005°$.

As for distorted sensor detection, the probability of distorted sensor detection versus $\epsilon_{\mathbf{Z}}$ and ϵ_γ is plotted in Fig. 2. It is observed the change of ϵ_γ has greater impact on distorted sensor detection. When it comes to source enumeration, for 100 Monte Carlo trials, the number of correct enumeration trials for different $\epsilon_{\mathbf{Z}}$ and ϵ_γ values are shown in Fig. 3. Since the signal subspace is discriminated using the estimated source number, no fail enumeration is permitted, otherwise it will lead to performance degradation. Considering all these aspects, we select $\epsilon_{\mathbf{Z}}$ and ϵ_γ in intervals $[1.5, 3]$ and $[5.5, 7]$ respectively in the simulations.

B. Comparative Study

In this subsection, the proposed algorithm is compared with Capon [51], MUSIC [6], LR^{2}SD-SVT, LR^{2}SD-APG, LR^{2}SD-ADMM, LR^{2}SD-IRLS [32], LR^{2}SD-BCD-ℓ_0 [34],

This article has been accepted for publication in IEEE Transactions on Aerospace and Electronic Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAES.2025.3553467

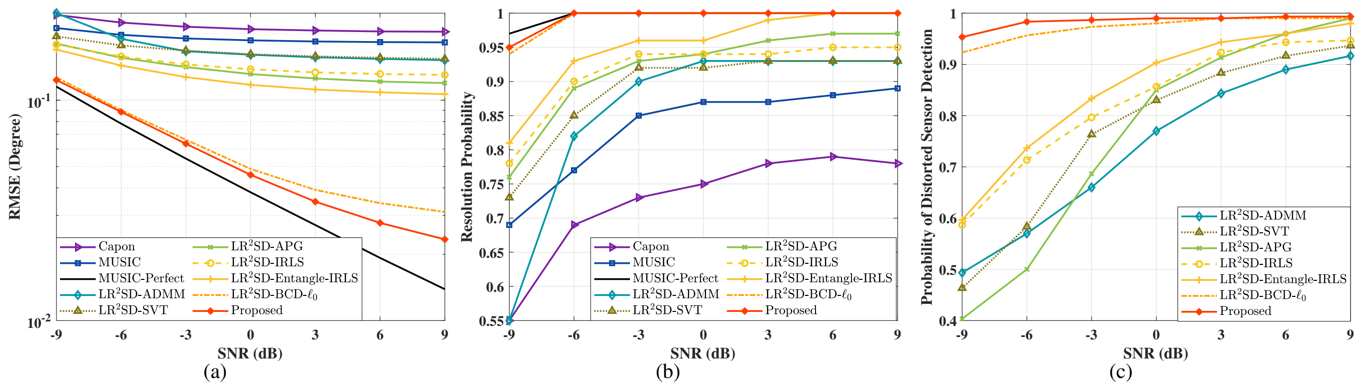

Fig. 5: (a) RMSE versus SNR. (b) Resolution probability versus SNR. (c) Probability of distorted sensor detection versus SNR. Here, $M = 20$ (three of which are distorted), $T = 100$, $Q = 3$ with $\theta_1 = -25°$, $\theta_2 = -10°$, and $\theta_3 = 5°$.

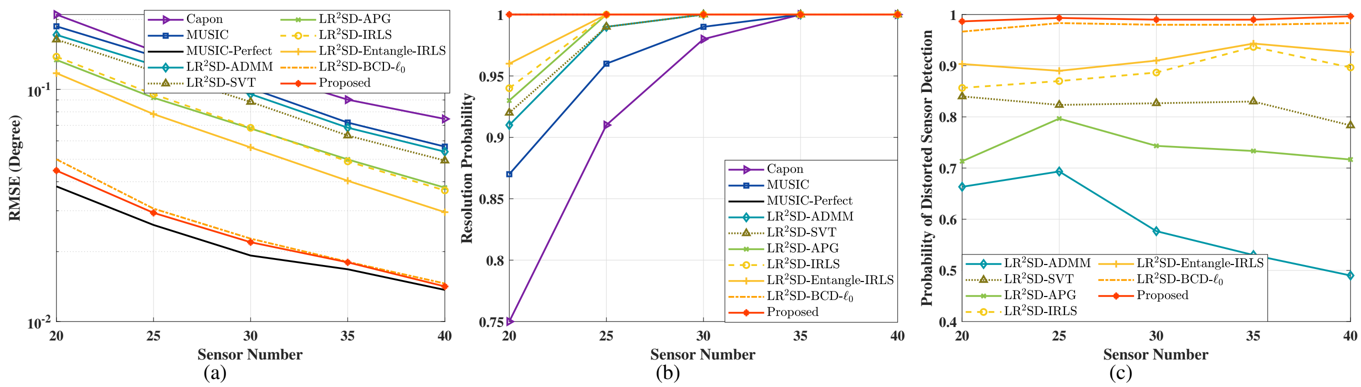

Fig. 6: (a) RMSE versus sensor number. (b) Resolution probability versus sensor number. (c) Probability of distorted sensor detection versus sensor number. We set SNR = 0 dB, $M_{\text{distort}} = 3$, $T = 100$, $Q = 3$ with $\theta_1 = -25°$, $\theta_2 = -10°$, and $\theta_3 = 5°$.

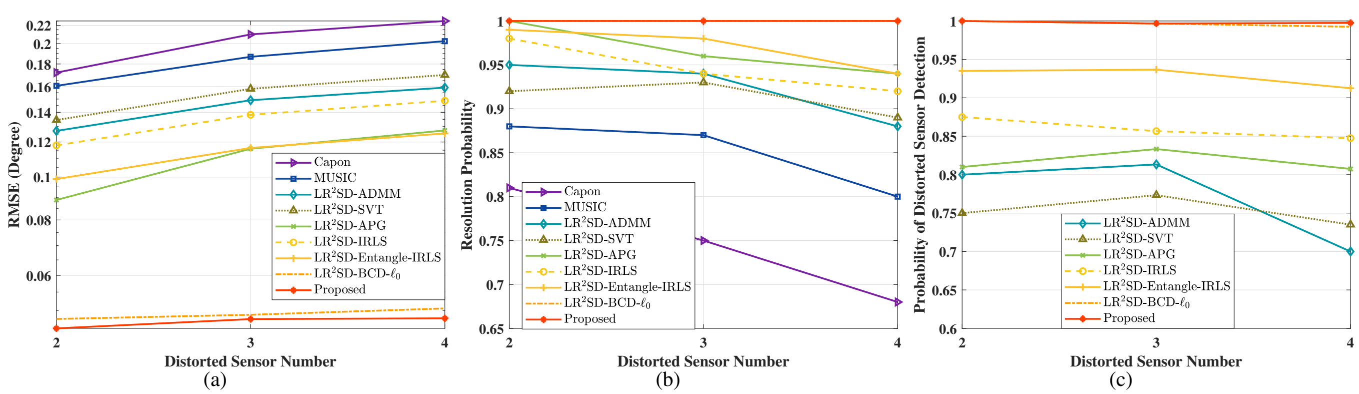

Fig. 7: (a) RMSE versus distorted sensor number. (b) Resolution probability versus distorted sensor number. (c) Probability of distorted sensor detection versus distorted sensor number. For other settings, SNR = 0 dB, $M = 20$, $T = 100$, $Q = 3$ with $\theta_1 = -25°$, $\theta_2 = -10°$, and $\theta_3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

and LR^{2}SD-Entangle-IRLS [33]. The MUSIC using perfect array without distorted sensors is considered as the benchmark, where the results are labeled as MUSIC-Perfect. The computational complexities of these algorithms are tabulated in Table I. Their runtimes versus snapshot number are plotted in Fig. 4. All codes are executed using MATLAB R2024b on a PC equipped with Intel(R) Core(TM) i9-14900K CPU and 64GB of RAM. LR^{2}SD-Entangle-IRLS is shown to have the maximum runtime, as its $\boldsymbol{\gamma}$ subproblem is a constrained Lasso problem and solved by quadratic programming, which is time-consuming. We see that LR^{2}SD-BCD-ℓ_0 requires the least runtime and the proposed algorithm comes the second. It is because our algorithm needs to calculate the proximal parameters $\mu_{\mathbf{Z}}^k$ and $\mu_{\boldsymbol{\gamma}}^k$ per iteration, and thus is a little bit slower than LR^{2}SD-BCD-ℓ_0. However, it is still faster than the remaining methods.

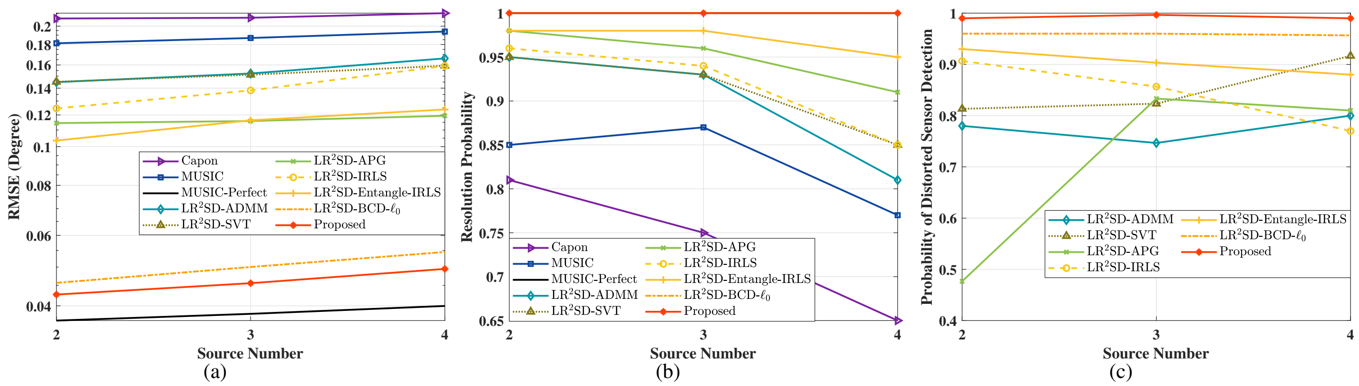

Fig. 8: (a) RMSE versus source number. (b) Resolution probability versus source number. (c) Probability of distorted sensor detection versus source number. Here, SNR = 0 dB, $M = 20$, $M_{\text{distort}} = 3$, $T = 100$. For $Q = 2$, $\theta_1 = -10°$ and $\theta_2 = 5°$. For $Q = 3$, $\theta_1 = -25°$,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

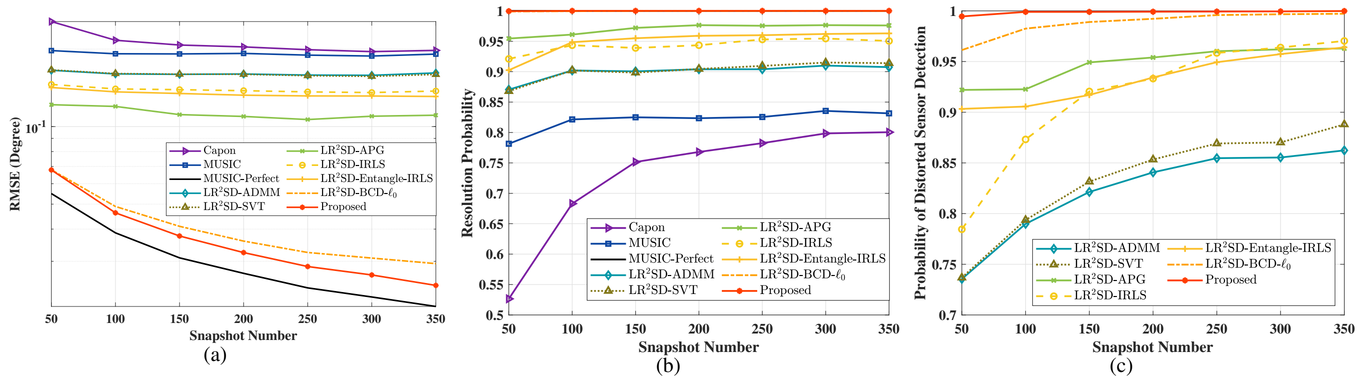

Fig. 9: (a) RMSE versus snapshot number. (b) Resolution probability versus snapshot number. (c) Probability of distorted sensor detection versus snapshot number. We set SNR = 0 dB, $M = 20$, $M_{\text{distort}} = 3$, $Q = 3$ with $\theta_1 = -25°$, $\theta_2 = -10°$, and $\theta_3 = 5°$.

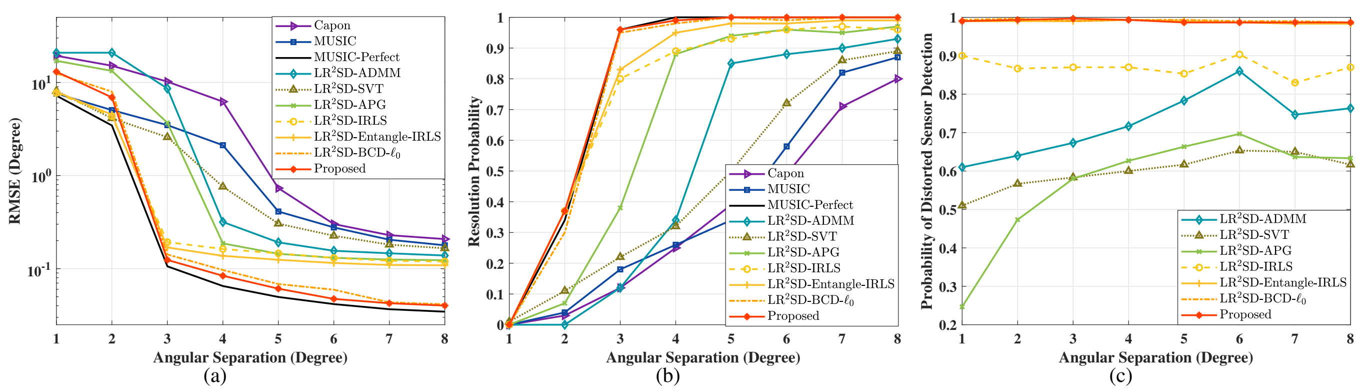

Fig. 10: (a) RMSE versus source separation angle. (b) Resolution probability versus source separation angle. (c) Probability of distorted sensor detection versus source separation angle. We set SNR = 0 dB, $M = 20$, $M_{\text{distort}} = 3$, $T = 100$, $Q = 2$ with $\theta_1 = 0°$, and θ_2 ranges from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

Under various settings, we conduct DOA estimation and distorted sensor detection using array observations with gain-phase errors. The source number is correctly identified for all simulations. In the following, various settings are tested, with only one variable being varied at one time while the others remain unchanged. Unless otherwise specified, the default settings are: SNR = 0 dB, 20 sensors (3 of which are distorted), 100 snapshots, and three sources from directions $-25°$, $-10°$, and $5°$.

We first evaluate the performance with SNR ranging from -9 dB to 9 dB. The results are plotted in Fig. 5. Comparing MUSIC-perfect with MUSIC, it is obvious that the existence of distorted sensors causes severe performance degradation. LR^{2}SD-BCD-ℓ_0 addresses the low-rank component and row-sparse part separately, while the proposed method considers the relationship between these two terms. It is found that the latter outperforms LR^{2}SD-BCD-ℓ_0, verifying the effectiveness of our idea. It is interesting that the RMSE differences between results by MUSIC-Perfect and proposed algorithm enlarges as SNR increases. This is because for low SNRs, the Gaussian noise is comparable with the gain-phase errors and has strong impacts on the results. Whereas a high SNR means a small Gaussian noise level, and it is mainly the gain-phase errors influence the performance, which is not much sensitive to the change of SNR.

The performance versus sensor number is investigated, and results are plotted in Fig. 6. We see that our algorithm outperforms its competitors in all metrics. The resolution probability remains 1 under all sensor numbers, consistent with the results by MUSIC-Perfect. Comparing the proposed method and MUSIC-perfect, it is found that their RMSE gaps narrow when sensor number increases. This is reasonable because the distorted sensor number is fixed, increasing the number of perfect sensors reduces the impact of distorted ones.

Next, the metrics versus distorted sensor number are plotted in Fig. 7. The results by MUSIC-Perfect are not involved because its performance is irrelevant with distorted sensors. It is seen that the performance of all algorithms degrades as distorted sensor number grows, where our algorithm is still superior to the competitors.

Besides, different numbers of source signals are examined. We set the source number as 2 (from -10° and 5°), 3 (from -25°, -10°, and 5°), and 4 (from -25°, -10°, 5°, and 20°). The performance versus source number is plotted in Fig. 8. Our method achieves a lower RMSE, improved resolution probability, and higher distorted sensor detection rate than the competing algorithms. It is worth mentioning that the source numbers are all correctly determined by the proposed algorithm.

We further compare the algorithm performance under snapshot number varying from 50 to 350 with a step of 50. The results are plotted in Fig. 9, and we observe that our algorithm still outperforms all competing schemes.

The performance w.r.t. angular separation is studied. We set the source number as two. The direction of one source is fixed as 0°, while that of the other changes from 1° to 8° in a step of 1°, resulting in the angular separation ranging from 1° to 8°. The results are displayed in Fig. 10. Regarding RMSE, the proposed algorithm does not perform the best when the angular differences are 1° and 2°, whereas its superiority is shown for larger separations. While for resolution probability and probability of distorted sensor detection, our algorithm always performs well.

IV. Conclusion

In this paper, we present an algorithm to tackle DOA estimation in the presence of sensor gain and phase uncertainties. The array observations are decomposed into a low-rank matrix and a row-sparse component, which are solved in a joint manner. Instead of applying convex surrogates for low-rankness and sparsity, we adopt the rank function and ℓ_0-norm for regularization. The subsequent problems are solved by hard-thresholding based on a scaled third quartile strategy. As a result, the proposed algorithm simultaneously attains DOA estimation, source enumeration, and distorted sensor detection. Its excellent performance is demonstrated under various experimental settings.

The gain-phase uncertainty studied in this work is introduced by distorted sensors, such that the gain-phase uncertainty matrix is diagonal with sparsely distributed diagonal elements. Apart from sensor distortion, mutual coupling and sensor location error will also cause gain-phase errors, where the uncertainty matrix has more complex structure and warrants further investigation.

Appendix A
Proof of Theorem II.1

According to Algorithm 1, the updates of $\mu_{\mathbf{Z}}^k$, μ_{γ}^k, $\omega_{\mathbf{Z}}^k$, and ω_{γ}^k satisfy:

$$\begin{aligned}
&\mu_{\mathbf{Z}}^k = \rho L_{\mathbf{Z}}^k, l \leq L_{\mathbf{Z}}^k \leq L, \\
&\mu_{\gamma}^k = \rho L_{\gamma}^k, l \leq L_{\gamma}^k \leq L, 1 < \rho < \infty, \\
&\omega_{\mathbf{Z}}^k \leq \bar{\tau}\sqrt{L_{\mathbf{Z}}^{k-1}/L_{\mathbf{Z}}^k}, \omega_{\gamma}^k \leq \bar{\tau}\sqrt{L_{\gamma}^{k-1}/L_{\gamma}^k}, 0 < \bar{\tau} < 1,
\end{aligned} \tag{37}$$

where l and L are the lower and upper bounds of the Lipschitz constants.

Since $f(\mathbf{Z}, \boldsymbol{\gamma})$ is gradient Lipschitz continuous w.r.t. each variable, it holds [52]

$$\begin{aligned}
f\left(\mathbf{Z}^k, \boldsymbol{\gamma}^{k-1}\right) \leq & f\left(\mathbf{Z}^{k-1}, \boldsymbol{\gamma}^{k-1}\right) \\
& + \left\langle \nabla_{\mathbf{Z}} f\left(\mathbf{Z}^{k-1}, \boldsymbol{\gamma}^{k-1}\right), \mathbf{Z}^k - \mathbf{Z}^{k-1}\right\rangle \\
& + \frac{L_{\mathbf{Z}}^k}{2}\left\|\mathbf{Z}^k - \mathbf{Z}^{k-1}\right\|_F^2,
\end{aligned} \tag{38}$$

$$\begin{aligned}
f\left(\mathbf{Z}^k, \boldsymbol{\gamma}^k\right) \leq & f\left(\mathbf{Z}^k, \boldsymbol{\gamma}^{k-1}\right) \\
& + \left\langle \nabla_{\boldsymbol{\gamma}} f\left(\mathbf{Z}^k, \boldsymbol{\gamma}^{k-1}\right), \boldsymbol{\gamma}^k - \boldsymbol{\gamma}^{k-1}\right\rangle \\
& + \frac{L_{\gamma}^k}{2}\left\|\boldsymbol{\gamma}^k - \boldsymbol{\gamma}^{k-1}\right\|_F^2.
\end{aligned} \tag{39}$$

Because $\mathbf{Z}^k$ and $\boldsymbol{\gamma}^k$ are the minimizers of (21) and (31), respectively, we have

$$\begin{aligned}
&\frac{\mu_{\mathbf{Z}}^k}{2}\left\|\mathbf{Z}^k - \hat{\mathbf{Z}}^{k-1}\right\|_F^2 + \left\langle \nabla_{\mathbf{Z}} f\left(\hat{\mathbf{Z}}^{k-1}, \boldsymbol{\gamma}^{k-1}\right), \mathbf{Z}^k\right\rangle \\
&+ \lambda_1 \operatorname{rank}\left(\mathbf{Z}^k\right) \leq \\
&\frac{\mu_{\mathbf{Z}}^k}{2}\left\|\mathbf{Z}^{k-1} - \hat{\mathbf{Z}}^{k-1}\right\|_F^2 + \left\langle \nabla_{\mathbf{Z}} f\left(\hat{\mathbf{Z}}^{k-1}, \boldsymbol{\gamma}^{k-1}\right), \mathbf{Z}^{k-1}\right\rangle \\
&+ \lambda_1 \operatorname{rank}\left(\mathbf{Z}^{k-1}\right),
\end{aligned} \tag{40}$$

and

$$\begin{aligned}
&\frac{\mu_{\gamma}^k}{2}\left\|\boldsymbol{\gamma}^k - \hat{\boldsymbol{\gamma}}^{k-1}\right\|_2^2 + \left\langle \nabla_{\boldsymbol{\gamma}} g\left(\mathbf{Z}^k, \hat{\boldsymbol{\gamma}}^{k-1}\right), \boldsymbol{\gamma}^k\right\rangle \\
&\quad + \lambda_2\left\|\boldsymbol{\gamma}^k\right\|_0 \leq \\
&\frac{\mu_{\gamma}^k}{2}\left\|\boldsymbol{\gamma}^{k-1} - \hat{\boldsymbol{\gamma}}^{k-1}\right\|_2^2 + \left\langle \nabla_{\boldsymbol{\gamma}} g\left(\mathbf{Z}^k, \hat{\boldsymbol{\gamma}}^{k-1}\right), \boldsymbol{\gamma}^{k-1}\right\rangle \\
&\quad + \lambda_2\left\|\boldsymbol{\gamma}^{k-1}\right\|_0.
\end{aligned} \tag{41}$$

This article has been accepted for publication in IEEE Transactions on Aerospace and Electronic Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAES.2025.3553467

After one iteration, the objective satisfies

$$F\left(\mathbf{L}^{k-1}\right)-F\left(\mathbf{L}^{k}\right)=F\left(\mathbf{Z}^{k-1},\boldsymbol{\gamma}^{k-1}\right)-F\left(\mathbf{Z}^{k},\boldsymbol{\gamma}^{k-1}\right)+F\left(\mathbf{Z}^{k},\boldsymbol{\gamma}^{k-1}\right)-F\left(\mathbf{Z}^{k},\boldsymbol{\gamma}^{k}\right). \tag{42}$$

For the update of $\mathbf{Z}$,

$$F\left(\mathbf{Z}^{k-1},\boldsymbol{\gamma}^{k-1}\right)-F\left(\mathbf{Z}^{k},\boldsymbol{\gamma}^{k-1}\right)=f\left(\mathbf{Z}^{k-1},\boldsymbol{\gamma}^{k-1}\right)-f\left(\mathbf{Z}^{k},\boldsymbol{\gamma}^{k-1}\right)+\lambda_1\text{rank}\left(\mathbf{Z}^{k-1}\right)-\lambda_1\text{rank}\left(\mathbf{Z}^{k}\right). \tag{43}$$

In the following, we omit $\boldsymbol{\gamma}^{k-1}$ for brevity. Based on (38) and (40), we further derive

$$\begin{aligned}
&F\left(\mathbf{Z}^{k-1}\right)-F\left(\mathbf{Z}^{k}\right)\\
&=f\left(\mathbf{Z}^{k-1}\right)-f\left(\mathbf{Z}^{k}\right)+\lambda_1\text{rank}\left(\mathbf{Z}^{k-1}\right)-\lambda_1\text{rank}\left(\mathbf{Z}^{k}\right)\\
&\geq\left\langle\nabla_{\mathbf{Z}}f\left(\hat{\mathbf{Z}}^{k-1}\right)-\nabla_{\mathbf{Z}}f\left(\mathbf{Z}^{k-1}\right),\mathbf{Z}^{k}-\mathbf{Z}^{k-1}\right\rangle\\
&\quad+\mu_{\mathbf{Z}}^{k}\left\langle\mathbf{Z}^{k}-\mathbf{Z}^{k-1},\mathbf{Z}^{k-1}-\hat{\mathbf{Z}}^{k-1}\right\rangle\\
&\quad+\left(\frac{\mu_{\mathbf{Z}}^{k}}{2}-\frac{L_{\mathbf{Z}}^{k}}{2}\right)\left\|\mathbf{Z}^{k}-\mathbf{Z}^{k-1}\right\|_F^2\\
&\geq-\left\|\mathbf{Z}^{k}-\mathbf{Z}^{k-1}\right\|_F\left(\left\|\nabla_{\mathbf{Z}}f\left(\hat{\mathbf{Z}}^{k-1}\right)-\nabla_{\mathbf{Z}}f\left(\mathbf{Z}^{k-1}\right)\right\|_F\right.\\
&\quad\left.+\mu_{\mathbf{Z}}^{k}\left\|\mathbf{Z}^{k-1}-\hat{\mathbf{Z}}^{k-1}\right\|_F\right)\\
&\quad+\left(\frac{\mu_{\mathbf{Z}}^{k}}{2}-\frac{L_{\mathbf{Z}}^{k}}{2}\right)\left\|\mathbf{Z}^{k}-\mathbf{Z}^{k-1}\right\|_F^2\\
&\geq-\left(\mu_{\mathbf{Z}}^{k}+L_{\mathbf{Z}}^{k}\right)\left\|\mathbf{Z}^{k}-\mathbf{Z}^{k-1}\right\|_F\left\|\mathbf{Z}^{k-1}-\hat{\mathbf{Z}}^{k-1}\right\|_F\\
&\quad+\left(\frac{\mu_{\mathbf{Z}}^{k}}{2}-\frac{L_{\mathbf{Z}}^{k}}{2}\right)\left\|\mathbf{Z}^{k}-\mathbf{Z}^{k-1}\right\|_F^2\\
&=-\left(\mu_{\mathbf{Z}}^{k}+L_{\mathbf{Z}}^{k}\right)\left\|\mathbf{Z}^{k}-\mathbf{Z}^{k-1}\right\|_F\omega_{\mathbf{Z}}^{k}\left\|\mathbf{Z}^{k-1}-\mathbf{Z}^{k-2}\right\|_F\\
&\quad+\left(\frac{\mu_{\mathbf{Z}}^{k}}{2}-\frac{L_{\mathbf{Z}}^{k}}{2}\right)\left\|\mathbf{Z}^{k}-\mathbf{Z}^{k-1}\right\|_F^2\\
&\geq\frac{1}{4}\left(\mu_{\mathbf{Z}}^{k}-L_{\mathbf{Z}}^{k}\right)\left\|\mathbf{Z}^{k}-\mathbf{Z}^{k-1}\right\|_F^2\\
&\quad-\frac{\left(\mu_{\mathbf{Z}}^{k}+L_{\mathbf{Z}}^{k}\right)^2}{\mu_{\mathbf{Z}}^{k}-L_{\mathbf{Z}}^{k}}{\omega_{\mathbf{Z}}^{k}}^2\left\|\mathbf{Z}^{k-1}-\mathbf{Z}^{k-2}\right\|_F^2\\
&=\frac{1}{4}\left(\rho-1\right)L_{\mathbf{Z}}^{k}\left\|\mathbf{Z}^{k}-\mathbf{Z}^{k-1}\right\|_F^2\\
&\quad-\frac{\left(\rho+1\right)^2}{\rho-1}L_{\mathbf{Z}}^{k}{\omega_{\mathbf{Z}}^{k}}^2\left\|\mathbf{Z}^{k-1}-\mathbf{Z}^{k-2}\right\|_F^2,
\end{aligned} \tag{44}$$

where the second inequality comes from the Cauchy-Schwarz inequality. The third inequality holds due to the Lipschitz gradient continuity of $f\left(\mathbf{Z},\boldsymbol{\gamma}\right)$. The fourth inequality is based on the Young's inequality.

Similar to (44), the following holds when updating $\boldsymbol{\gamma}$:

$$F\left(\boldsymbol{\gamma}^{k-1}\right)-F\left(\boldsymbol{\gamma}^{k}\right)\geq\frac{1}{4}\left(\rho-1\right)L_{\boldsymbol{\gamma}}^{k}\left\|\boldsymbol{\gamma}^{k}-\boldsymbol{\gamma}^{k-1}\right\|_F^2-\frac{\left(\rho+1\right)^2}{\rho-1}L_{\boldsymbol{\gamma}}^{k}{\omega_{\boldsymbol{\gamma}}^{k}}^2\left\|\boldsymbol{\gamma}^{k-1}-\boldsymbol{\gamma}^{k-2}\right\|_F^2, \tag{45}$$

where $\mathbf{Z}^k$ is omitted.

When the extrapolation weight is zero, merging (42), (44), and (45) yields:

$$\begin{aligned}
F\left(\mathbf{L}^{k-1}\right)-F\left(\mathbf{L}^{k}\right)&\geq\frac{1}{4}\left(\rho-1\right)L_{\mathbf{Z}}^{k}\left\|\mathbf{Z}^{k}-\mathbf{Z}^{k-1}\right\|_F^2\\
&\quad+\frac{1}{4}\left(\rho-1\right)L_{\boldsymbol{\gamma}}^{k}\left\|\boldsymbol{\gamma}^{k}-\boldsymbol{\gamma}^{k-1}\right\|_F^2\\
&\geq\frac{1}{4}\left(\rho-1\right)l\left\|\mathbf{L}^{k}-\mathbf{L}^{k-1}\right\|_F^2\geq 0.
\end{aligned} \tag{46}$$

According to Algorithm 1, in each iteration, if the solution generated by extrapolation strategy makes the objective increase, we set extrapolation weights to zero and redo this iteration. Then, the objective satisfies (46), implying its non-increasing property. By doing so, sequence $\left\{F\left(\mathbf{L}^{k}\right)\right\}_{k\in\mathbb{N}}$ is guaranteed to be non-increasing. Moreover, since the Frobenius norm, rank function, and ℓ_0-norm are all lower-bounded, we deduce that $\left\{F\left(\mathbf{L}^{k}\right)\right\}_{k\in\mathbb{N}}$ is also lower-bounded. Therefore, Theorem II.1 is proved.

Appendix B
Proof of Theorem II.2

We first prove the solution sequence is square summable, with which the sequence boundness is concluded. It is suitable to set

$$\omega_{\mathbf{Z}}^{k}\leq\frac{\tau\left(\rho-1\right)}{2\left(\rho+1\right)}\sqrt{L_{\mathbf{Z}}^{k-1}/L_{\mathbf{Z}}^{k}},\omega_{\boldsymbol{\gamma}}^{k}\leq\frac{\tau\left(\rho-1\right)}{2\left(\rho+1\right)}\sqrt{L_{\boldsymbol{\gamma}}^{k-1}/L_{\boldsymbol{\gamma}}^{k}}. \tag{47}$$

Combining (44), (45), and (47), we have

$$\begin{aligned}
F\left(\mathbf{Z}^{k-1}\right)-F\left(\mathbf{Z}^{k}\right)&\geq\frac{1}{4}\left(\rho-1\right)L_{\mathbf{Z}}^{k}\left\|\mathbf{Z}^{k}-\mathbf{Z}^{k-1}\right\|_F^2\\
&\quad-\frac{1}{4}\left(\rho-1\right)L_{\mathbf{Z}}^{k-1}\tau^2\left\|\mathbf{Z}^{k-1}-\mathbf{Z}^{k-2}\right\|_F^2,\\
F\left(\boldsymbol{\gamma}^{k-1}\right)-F\left(\boldsymbol{\gamma}^{k}\right)&\geq\frac{1}{4}\left(\rho-1\right)L_{\boldsymbol{\gamma}}^{k}\left\|\boldsymbol{\gamma}^{k}-\boldsymbol{\gamma}^{k-1}\right\|_F^2\\
&\quad-\frac{1}{4}\left(\rho-1\right)L_{\boldsymbol{\gamma}}^{k-1}\tau^2\left\|\boldsymbol{\gamma}^{k-1}-\boldsymbol{\gamma}^{k-2}\right\|_F^2,
\end{aligned} \tag{48}$$

which leads to

$$\begin{aligned}
F\left(\mathbf{L}^{k-1}\right)-F\left(\mathbf{L}^{k}\right)&\geq\frac{1}{4}\left(\rho-1\right)L_{\mathbf{Z}}^{k}\left\|\mathbf{Z}^{k}-\mathbf{Z}^{k-1}\right\|_F^2\\
&\quad-\frac{1}{4}\left(\rho-1\right)L_{\mathbf{Z}}^{k-1}\tau^2\left\|\mathbf{Z}^{k-1}-\mathbf{Z}^{k-2}\right\|_F^2\\
&\quad+\frac{1}{4}\left(\rho-1\right)L_{\boldsymbol{\gamma}}^{k}\left\|\boldsymbol{\gamma}^{k}-\boldsymbol{\gamma}^{k-1}\right\|_F^2\\
&\quad-\frac{1}{4}\left(\rho-1\right)L_{\boldsymbol{\gamma}}^{k-1}\tau^2\left\|\boldsymbol{\gamma}^{k-1}-\boldsymbol{\gamma}^{k-2}\right\|_F^2.
\end{aligned} \tag{49}$$

Employing (49) from 1 to K, we get

$$F\left(\mathbf{L}^{0}\right)-F\left(\mathbf{L}^{K}\right)\geq\sum_{k=1}^{K}\frac{(\rho-1)}{4}l\left(1-\tau^2\right)\left\|\mathbf{L}^{k-1}-\mathbf{L}^{k-2}\right\|_F^2. \tag{50}$$

Here, we use the lower bounds of $L_{\mathbf{Z}}^{k}$ and $L_{\boldsymbol{\gamma}}^{k}$. Besides, $\mathbf{L}^0=\mathbf{L}^{-1}$ is defined. In particular, (50) implies

$$\lim_{k\to\infty}\left\|\mathbf{L}^{k}-\mathbf{L}^{k-1}\right\|_F^2=0, \tag{51}$$

Hereafter, we denote $F\left(\mathbf{L}^{k}\right)$ as F^k for brevity. Since $F^k\to\infty$ if and only if $\left\|\mathbf{L}\right\|_F\to\infty$ and F^k is lower bounded,

$\{\mathbf{L}^k\}_{k\in\mathbb{N}}$ is bounded. Setting $\bar{\mathbf{L}}$ as a limit point of $\{\mathbf{L}^k\}_{k\in\mathbb{N}}$, then there exists a subsequence $\{\mathbf{L}_s^k\}_{k_s\in\mathbb{N}}$ which converges to $\bar{\mathbf{L}}$. That is, $\mathbf{L}^{k_s}\to\bar{\mathbf{L}}$ when $k_s\to\infty$.

Next, we show that $\bar{\mathbf{L}}$ is also a critical point of $F(\mathbf{L})$. The rank function and ℓ_0-norm are represented by $r_1(\cdot)$ and $r_2(\cdot)$, respectively. According to (20),

$$\begin{aligned}\mathbf{Z}^k \in \arg\min_{\mathbf{Z}}\ & \lambda_1 r_1(\mathbf{Z}) + \frac{\mu_{\mathbf{Z}}^k}{2}\left\|\mathbf{Z}-\hat{\mathbf{Z}}^{k-1}\right\|_F^2 \\ & + \left\langle \nabla_{\mathbf{Z}} f\left(\hat{\mathbf{Z}}^{k-1},\boldsymbol{\gamma}^{k-1}\right), \mathbf{Z}-\hat{\mathbf{Z}}^{k-1}\right\rangle \end{aligned} \tag{52}$$

Because $r_1(\cdot)$ is lower semicontinuous, it holds

$$\lim_{k_q\to\infty}\inf\ r_1\left(\mathbf{Z}^{k_q}\right) = r_1\left(\bar{\mathbf{Z}}\right). \tag{53}$$

Then, based on (52) and (53), we derive

$$\begin{aligned} r_1\left(\bar{\mathbf{Z}}\right) \le & \lim_{k_q\to\infty}\inf\ \lambda_1 r_1\left(\mathbf{Z}^{k_q}\right) + \frac{\mu_{\mathbf{Z}}^{k_q}}{2}\left\|\mathbf{Z}^{k_q}-\hat{\mathbf{Z}}^{k_q-1}\right\|_F^2 \\ & + \left\langle \nabla_{\mathbf{Z}} f\left(\hat{\mathbf{Z}}^{k_q-1},\boldsymbol{\gamma}^{k_q-1}\right), \mathbf{Z}-\hat{\mathbf{Z}}^{k_q-1}\right\rangle \\ \le & \lim_{k_q\to\infty}\inf\ \lambda_1 r_1(\mathbf{Z}) + \frac{\mu_{\mathbf{Z}}^{k_q}}{2}\left\|\mathbf{Z}-\hat{\mathbf{Z}}^{k_q-1}\right\|_F^2 \\ & + \left\langle \nabla_{\mathbf{Z}} f\left(\hat{\mathbf{Z}}^{k_q-1},\boldsymbol{\gamma}^{k_q-1}\right), \mathbf{Z}-\hat{\mathbf{Z}}^{k_q-1}\right\rangle \\ = & \lambda_1 r_1(\mathbf{Z}) + \frac{\bar{\mu}_{\mathbf{Z}}}{2}\left\|\mathbf{Z}-\bar{\mathbf{Z}}\right\|_F^2 + \left\langle \nabla_{\mathbf{Z}} f\left(\bar{\mathbf{Z}},\bar{\boldsymbol{\gamma}}\right), \mathbf{Z}-\bar{\mathbf{Z}}\right\rangle \end{aligned} \tag{54}$$

which holds for $\forall\mathbf{Z}\in\mathrm{dom}(F)$. Here, the last equality utilizes $\lim_{k_q\to\infty}\hat{\mathbf{Z}}^{k_q-1}=\bar{\mathbf{Z}}$, $\lim_{k_q\to\infty}\hat{\boldsymbol{\gamma}}^{k_q-1}=\bar{\boldsymbol{\gamma}}$, and $\lim_{k_q\to\infty}\mu_{\mathbf{Z}}^{k_q}=\bar{\mu}_{\mathbf{Z}}$.

Therefore,

$$\begin{aligned}\bar{\mathbf{Z}} \in \arg\min_{\mathbf{Z}}\ & \lambda_1 r_1(\mathbf{Z}) + \frac{\bar{\mu}_{\mathbf{Z}}}{2}\left\|\mathbf{Z}-\bar{\mathbf{Z}}\right\|_F^2 \\ & + \left\langle \nabla_{\mathbf{Z}} f\left(\bar{\mathbf{Z}},\bar{\boldsymbol{\gamma}}\right), \mathbf{Z}-\bar{\mathbf{Z}}\right\rangle. \end{aligned} \tag{55}$$

The first-order optimality condition yields

$$0 \in \nabla_{\mathbf{Z}} f\left(\bar{\mathbf{Z}},\bar{\boldsymbol{\gamma}}\right) + \lambda_1 \partial r_1\left(\bar{\mathbf{Z}}\right). \tag{56}$$

In the same way, we deduce

$$0 \in \nabla_{\boldsymbol{\gamma}} f\left(\bar{\mathbf{Z}},\bar{\boldsymbol{\gamma}}\right) + \lambda_2 \partial r_2\left(\bar{\boldsymbol{\gamma}}\right), \tag{57}$$

The above two results reveal that the limit point $\bar{\mathbf{L}}=\left(\bar{\mathbf{Z}},\bar{\boldsymbol{\gamma}}\right)$ is also a critical point of (17).

So far, we have proved that a subsequence of $\{\mathbf{L}^k\}_{k\in\mathbb{N}}$ converges to a critical point of (17). In the following, we will prove $\{\mathbf{L}^k\}_{k\in\mathbb{N}}$ is a Cauchy sequence and converges. The sequence convergence is attained using the Kurdyka-Łojasiewicz (KŁ) inequality [53]. The objective $F(\mathbf{L})$ bares o-minimal structure and satisfies KŁ property at any feasible point [54].

We first prove any $\mathbf{L}^k$ belongs to the neighborhood of $\bar{\mathbf{L}}$, which is a prerequisite for KŁ property. Let $B(\bar{v},h):=\{v:\|v-\bar{v}\|\le h\}$. One assumption is that $\mathbf{L}^0$ is sufficiently close to $\bar{\mathbf{L}}$, denoted as $\mathbf{L}^0\in B\left(\bar{\mathbf{L}},h\right)$. Besides, it is appropriate to set $\bar{F}:=F\left(\bar{\mathbf{L}}\right)$ as 0. If $\bar{F}\ne 0$, we can subtract $F(\mathbf{L})$ by $\bar{F}$.

Setting $k=1$ in (49), we derive

$$F^0 \ge F^0 - F^1 \ge \frac{1}{4}(\rho-1)\, l \left\|\mathbf{L}^1-\mathbf{L}^0\right\|_F^2, \tag{58}$$

and we get

$$\left\|\mathbf{L}^1-\mathbf{L}^0\right\|_F \le \sqrt{\frac{F^0}{C_1 l}}, \tag{59}$$

where $C_1=\frac{1}{4}(\rho-1)$. Then,

$$\begin{aligned}\left\|\mathbf{L}^1-\bar{\mathbf{L}}\right\|_F &\le \left\|\mathbf{L}^1-\mathbf{L}^0\right\|_F + \left\|\mathbf{L}^0-\bar{\mathbf{L}}\right\|_F \\ &\le \sqrt{\frac{F^0}{C_1 l}} + \left\|\mathbf{L}^0-\bar{\mathbf{L}}\right\|_F. \end{aligned} \tag{60}$$

We adopt a proper h which satisfies $\mathbf{L}^1\in B\left(\bar{\mathbf{L}},h\right)$.

Setting $k=2$ in (49),

$$F^0 \ge F^1 - F^2 \ge C_1 l \left\|\mathbf{L}^2-\mathbf{L}^1\right\|_F^2 - C_2 l \left\|\mathbf{L}^1-\mathbf{L}^0\right\|_F^2, \tag{61}$$

where $C_2=C_1\tau^2$. Based on (59) and (61), it holds

$$\left\|\mathbf{L}^2-\mathbf{L}^1\right\|_F \le \sqrt{\frac{C_1+C_2}{C_1^2 l}F^0}. \tag{62}$$

Then, we obtain

$$\begin{aligned}\left\|\mathbf{L}^2-\bar{\mathbf{L}}\right\|_F &\le \left\|\mathbf{L}^2-\mathbf{L}^1\right\|_F + \left\|\mathbf{L}^1-\bar{\mathbf{L}}\right\|_F \\ &\le \sqrt{\frac{C_1+C_2}{C_1^2 l}F^0} + \sqrt{\frac{F^0}{C_1 l}} + \left\|\mathbf{L}^0-\bar{\mathbf{L}}\right\|_F, \end{aligned} \tag{63}$$

which indicates $\mathbf{L}^2\in B\left(\bar{\mathbf{L}},h\right)$.

We assume $\mathbf{L}^k\in B\left(\bar{\mathbf{L}},h\right)$ for $0\le k\le K$ and further verify $\mathbf{L}^{K+1}\in B\left(\bar{\mathbf{L}},h\right)$. The first-order derivative of $F\left(\mathbf{L}^k\right)$ is

$$\partial F\left(\mathbf{L}^k\right) = \left(\lambda_1\partial r_1\left(\mathbf{Z}^k\right)+\nabla_{\mathbf{Z}} f\left(\mathbf{L}^k\right), \lambda_2\partial r_2\left(\boldsymbol{\gamma}^k\right)+\nabla_{\boldsymbol{\gamma}} f\left(\mathbf{L}^k\right)\right), \tag{64}$$

Then, utilizing the optimality condition of (20) and (30), we obtain

$$\begin{aligned} \lambda_1\partial r_1\left(\mathbf{Z}^k\right)+\nabla_{\mathbf{Z}} f\left(\mathbf{L}^k\right) &= \nabla_{\mathbf{Z}} f\left(\mathbf{L}^k\right) \\ & - \mu_{\mathbf{Z}}^k\left(\mathbf{Z}^k-\hat{\mathbf{Z}}^{k-1}\right) - \nabla_{\mathbf{Z}} f\left(\hat{\mathbf{Z}}^{k-1},\boldsymbol{\gamma}^{k-1}\right), \\ \lambda_2\partial r_2\left(\boldsymbol{\gamma}^k\right)+\nabla_{\boldsymbol{\gamma}} f\left(\mathbf{L}^k\right) &= \nabla_{\boldsymbol{\gamma}} f\left(\mathbf{L}^k\right) \\ & - \mu_{\boldsymbol{\gamma}}^k\left(\boldsymbol{\gamma}^k-\hat{\boldsymbol{\gamma}}^{k-1}\right) - \nabla_{\boldsymbol{\gamma}} f\left(\mathbf{Z}^k,\hat{\boldsymbol{\gamma}}^{k-1}\right). \end{aligned} \tag{65}$$

Based on (64) and (65), we derive

$$\begin{aligned}\mathrm{dist}\left(0,\partial F\left(\mathbf{L}^k\right)\right) \le & \left\|\left(\mu_{\mathbf{Z}}^k\left(\mathbf{Z}^k-\hat{\mathbf{Z}}^{k-1}\right), \mu_{\boldsymbol{\gamma}}^k\left(\boldsymbol{\gamma}^k-\hat{\boldsymbol{\gamma}}^{k-1}\right)\right)\right\|_F \\ & + \left\|\nabla_{\mathbf{Z}} f\left(\mathbf{L}^k\right), \nabla_{\mathbf{Z}} f\left(\hat{\mathbf{Z}}^{k-1},\boldsymbol{\gamma}^{k-1}\right)\right\|_F \\ & + \left\|\nabla_{\boldsymbol{\gamma}} f\left(\mathbf{L}^k\right) - \nabla_{\boldsymbol{\gamma}} f\left(\mathbf{Z}^k,\hat{\boldsymbol{\gamma}}^{k-1}\right)\right\|_F. \end{aligned} \tag{66}$$

For the first term on the right side, due to

$$\begin{aligned} & \left\|\mu_{\mathbf{Z}}^k\left(\mathbf{Z}^k-\hat{\mathbf{Z}}^{k-1}\right)\right\|_F \\ & \le \rho L\left\|\mathbf{Z}^k-\mathbf{Z}^{k-1}\right\|_F + \bar{\tau}\rho L\left\|\mathbf{Z}^{k-1}-\mathbf{Z}^{k-2}\right\|_F \\ & \le \rho L\left(\left\|\mathbf{Z}^k-\mathbf{Z}^{k-1}\right\|_F + \left\|\mathbf{Z}^{k-1}-\mathbf{Z}^{k-2}\right\|_F\right), \end{aligned} \tag{67}$$

This article has been accepted for publication in IEEE Transactions on Aerospace and Electronic Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAES.2025.3553467

we have

$$\begin{aligned}&\left\|\left(\mu_{\mathbf{Z}}^k\left(\mathbf{Z}^k-\hat{\mathbf{Z}}^{k-1}\right),\mu_{\boldsymbol{\gamma}}^k\left(\boldsymbol{\gamma}^k-\hat{\boldsymbol{\gamma}}^{k-1}\right)\right)\right\|_F\\&\quad\leq\rho L\left(\left\|\mathbf{L}^k-\mathbf{L}^{k-1}\right\|_F+\left\|\mathbf{L}^{k-1}-\mathbf{L}^{k-2}\right\|_F\right).\end{aligned}\tag{68}$$

We define the global gradient Lipschitz constant of $f(\mathbf{L})$ with each coordinate as L_G. Therefore,

$$\begin{aligned}&\left\|\nabla_{\mathbf{Z}}f\left(\mathbf{L}^k\right),\nabla_{\mathbf{Z}}f\left(\hat{\mathbf{Z}}^{k-1},\boldsymbol{\gamma}^{k-1}\right)\right\|_F\\&\quad+\left\|\nabla_{\boldsymbol{\gamma}}f\left(\mathbf{L}^k\right)-\nabla_{\boldsymbol{\gamma}}f\left(\mathbf{Z}^k,\hat{\boldsymbol{\gamma}}^{k-1}\right)\right\|_F\\&\leq L_G\left(\left\|\left(\hat{\mathbf{Z}}^{k-1},\boldsymbol{\gamma}^{k-1}\right)-\left(\mathbf{Z}^k,\boldsymbol{\gamma}^k\right)\right\|_F\right.\\&\quad\left.+\left\|\left(\mathbf{Z}^k,\hat{\boldsymbol{\gamma}}^{k-1}\right)-\left(\mathbf{Z}^k,\boldsymbol{\gamma}^k\right)\right\|_F\right)\\&\leq 2L_G\left(\left\|\mathbf{L}^k-\mathbf{L}^{k-1}\right\|_F+\left\|\mathbf{L}^{k-1}-\mathbf{L}^{k-2}\right\|_F\right).\end{aligned}\tag{69}$$

Combining (68), (69), and (66) results in

$$\begin{aligned}\operatorname{dist}\left(0,\partial F\left(\mathbf{L}^k\right)\right)\leq(\rho L+2L_G)&\left(\left\|\mathbf{L}^k-\mathbf{L}^{k-1}\right\|_F\right.\\&\left.+\left\|\mathbf{L}^{k-1}-\mathbf{L}^{k-2}\right\|_F\right).\end{aligned}\tag{70}$$

Since $F(\mathbf{L})$ satisfies KŁ inequality, it holds

$$\varphi^{'}\left(F^k\right)\operatorname{dist}\left(0,\partial F\left(\mathbf{L}^k\right)\right)\geq 1,\tag{71}$$

where $\varphi(\cdot)$ represents a concave and continuous function. Then, based on (70) and (71), we further obtain:

$$\begin{aligned}\varphi^{'}\left(F^k\right)\geq(\rho L+2L_G)^{-1}&\left(\left\|\mathbf{L}^k-\mathbf{L}^{k-1}\right\|_F\right.\\&\left.+\left\|\mathbf{L}^{k-1}-\mathbf{L}^{k-2}\right\|_F\right)^{-1}.\end{aligned}\tag{72}$$

Utilizing the concavity of $\varphi(\cdot)$, we have

$$\varphi\left(F^k\right)-\varphi\left(F^{k+1}\right)\geq\varphi^{'}\left(F^k\right)\left(F^k-F^{k+1}\right).\tag{73}$$

It follows from (72), (73), and (49) that

$$\begin{aligned}&\varphi\left(F^k\right)-\varphi\left(F^{k+1}\right)\geq\\&\quad\frac{C_1\left\|\mathbf{L}^k-\mathbf{L}^{k+1}\right\|_F^2-C_2\left\|\mathbf{L}^{k-1}-\mathbf{L}^k\right\|_F^2}{(\rho L+2L_G)\left(\left\|\mathbf{L}^k-\mathbf{L}^{k-1}\right\|_F+\left\|\mathbf{L}^{k-1}-\mathbf{L}^{k-2}\right\|_F\right)},\end{aligned}\tag{74}$$

which equals

$$\begin{aligned}&C_1\left\|\mathbf{L}^k-\mathbf{L}^{k+1}\right\|_F^2\leq C_2\left\|\mathbf{L}^{k-1}-\mathbf{L}^k\right\|_F^2\\&\quad+(\rho L+2L_G)\left(\left\|\mathbf{L}^k-\mathbf{L}^{k-1}\right\|_F+\left\|\mathbf{L}^{k-1}-\mathbf{L}^{k-2}\right\|_F\right)\\&\quad\left(\varphi\left(F^k\right)-\varphi\left(F^{k+1}\right)\right).\end{aligned}\tag{75}$$

Employing $a^2+b^2\leq(a+b)^2$ and $ab\leq ca^2+\frac{b^2}{4c}$ for $c>0$, based on (75), we derive

$$\begin{aligned}&\sqrt{C_1}\left\|\mathbf{L}^k-\mathbf{L}^{k+1}\right\|_F\leq\sqrt{C_2}\left\|\mathbf{L}^{k-1}-\mathbf{L}^k\right\|_F\\&\quad+\sqrt{(\rho L+2L_G)\left(\left\|\mathbf{L}^k-\mathbf{L}^{k-1}\right\|_F+\left\|\mathbf{L}^{k-1}-\mathbf{L}^{k-2}\right\|_F\right)}\\&\quad\sqrt{\left(\varphi\left(F^k\right)-\varphi\left(F^{k+1}\right)\right)}\\&\leq\sqrt{C_2}\left\|\mathbf{L}^{k-1}-\mathbf{L}^k\right\|_F\\&\quad+\frac{\sqrt{C_1}-\sqrt{C_2}}{3}\left(\left\|\mathbf{L}^k-\mathbf{L}^{k-1}\right\|_F+\left\|\mathbf{L}^{k-1}-\mathbf{L}^{k-2}\right\|_F\right)\\&\quad+\frac{3(\rho L+2L_G)}{4\left(\sqrt{C_1}-\sqrt{C_2}\right)}\left(\varphi\left(F^k\right)-\varphi\left(F^{k+1}\right)\right).\end{aligned}\tag{76}$$

Taking summation of (76) for k from 2 to K yields

$$\begin{aligned}&\sqrt{C_1}\left\|\mathbf{L}^K-\mathbf{L}^{K+1}\right\|_F\\&\quad+\sum_{k=2}^{K-1}\left(\sqrt{C_1}-\sqrt{C_2}\right)\left\|\mathbf{L}^k-\mathbf{L}^{k+1}\right\|_F\\&\leq\sqrt{C_2}\left\|\mathbf{L}^1-\mathbf{L}^2\right\|_F\\&\quad+\frac{\sqrt{C_1}-\sqrt{C_2}}{3}\sum_{k=2}^{K}\left(\left\|\mathbf{L}^k-\mathbf{L}^{k-1}\right\|_F+\left\|\mathbf{L}^{k-1}-\mathbf{L}^{k-2}\right\|_F\right)\\&\quad+\frac{3(\rho L+2L_G)}{4\left(\sqrt{C_1}-\sqrt{C_2}\right)}\left(\varphi\left(F^2\right)-\varphi\left(F^{K+1}\right)\right),\end{aligned}\tag{77}$$

We further obtain from the above inequality:

$$\begin{aligned}&\sum_{k=2}^{K}\left(\sqrt{C_1}-\sqrt{C_2}\right)\left\|\mathbf{L}^k-\mathbf{L}^{k+1}\right\|_F\\&\leq\sqrt{C_2}\left\|\mathbf{L}^1-\mathbf{L}^2\right\|_F\\&\quad+\frac{\sqrt{C_1}-\sqrt{C_2}}{3}\sum_{k=2}^{K}\left(\left\|\mathbf{L}^k-\mathbf{L}^{k-1}\right\|_F+\left\|\mathbf{L}^{k-1}-\mathbf{L}^{k-2}\right\|_F\right)\\&\quad+\frac{3(\rho L+2L_G)}{4\left(\sqrt{C_1}-\sqrt{C_2}\right)}\left(\varphi\left(F^2\right)-\varphi\left(F^{K+1}\right)\right),\end{aligned}\tag{78}$$

which implies

$$\begin{aligned}&\sum_{k=2}^{K}\left\|\mathbf{L}^k-\mathbf{L}^{k+1}\right\|_F\\&\leq\left(\frac{3\sqrt{C_2}}{\sqrt{C_1}-\sqrt{C_2}}+2\right)\left\|\mathbf{L}^1-\mathbf{L}^2\right\|_F+\left\|\mathbf{L}^0-\mathbf{L}^1\right\|_F\\&\quad+\frac{9(\rho L+2L_G)}{4\left(\sqrt{C_1}-\sqrt{C_2}\right)^2}\left(\varphi\left(F^2\right)-\varphi\left(F^{K+1}\right)\right).\end{aligned}\tag{79}$$

Recalling (59), (62), and (63), and from (79), we have

$$\begin{aligned}&\left\|\mathbf{L}^{K+1}-\bar{\mathbf{L}}\right\|_F\\&\leq\sum_{k=2}^{K}\left\|\mathbf{L}^k-\mathbf{L}^{k+1}\right\|_F+\left\|\mathbf{L}^2-\bar{\mathbf{L}}\right\|_F\\&\leq\left(\frac{3\sqrt{C_2}}{\sqrt{C_1}-\sqrt{C_2}}+3\right)\sqrt{\frac{C_1+C_2}{C_1^2l}F^0}+2\sqrt{\frac{F^0}{C_1l}}\\&\quad+\frac{9(\rho L+2L_G)}{4\left(\sqrt{C_1}-\sqrt{C_2}\right)^2}\varphi\left(F^0\right)+\left\|\mathbf{L}^0-\bar{\mathbf{L}}\right\|_F\end{aligned}\tag{80}$$

where $\varphi\left(F^0\right)\geq\varphi\left(F^k\right)$ is used. Therefore, we successfully prove $\mathbf{L}^{K+1}\in B\left(\bar{\mathbf{L}},h\right)$, where h should meet

$$\begin{aligned}h\geq&\left(\frac{3\sqrt{C_2}}{\sqrt{C_1}-\sqrt{C_2}}+3\right)\sqrt{\frac{C_1+C_2}{C_1^2l}F^0}+2\sqrt{\frac{F^0}{C_1l}}\\&+\frac{9(\rho L+2L_G)}{4\left(\sqrt{C_1}-\sqrt{C_2}\right)^2}\varphi\left(F^0\right)+\left\|\mathbf{L}^0-\bar{\mathbf{L}}\right\|_F.\end{aligned}\tag{81}$$

In other words, (81) defines the required closeness between $\mathbf{L}^0$ and $\bar{\mathbf{L}}$.

For (76), summing up from $\bar{k}$ to K, we obtain

$$\begin{aligned}
&\sum_{k=\bar{k}}^{K} \left\| \mathbf{L}^{\bar{k}} - \mathbf{L}^{\bar{k}+1} \right\|_F \\
&\leq \left(\frac{3\sqrt{C_2}}{\sqrt{C_1}-\sqrt{C_2}} + 2 \right) \left\| \mathbf{L}^{\bar{k}-1} - \mathbf{L}^{\bar{k}} \right\|_F + \left\| \mathbf{L}^{\bar{k}-2} - \mathbf{L}^{\bar{k}-1} \right\|_F \\
&\quad + \frac{9\left(\rho L + 2L_G\right)}{4\left(\sqrt{C_1}-\sqrt{C_2}\right)^2} \left(\varphi\left(F^{\bar{k}}\right) - \varphi\left(F^{K+1}\right) \right), \qquad (82)
\end{aligned}$$

which indicates that

$$\lim_{\bar{k}\to\infty} \sum_{k=\bar{k}}^{\infty} \left\| \mathbf{L}^{\bar{k}} - \mathbf{L}^{\bar{k}+1} \right\|_F = 0. \qquad (83)$$

Thus, $\left\{\mathbf{L}^k\right\}_{k\in\mathbf{N}}$ is a Cauchy sequence and converges to $\bar{\mathbf{L}}$.

REFERENCES

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

[2] Z. Zheng, W.-Q. Wang, H. Meng, H. C. So, and H. Zhang, "Efficient beamspace-based algorithm for two-dimensional DOA estimation of incoherently distributed sources in massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 12, pp. 11 776–11 789, Dec. 2018.

[3] M. Koivisto, M. Costa, J. Werner, K. Heiska, J. Talvitie, K. Leppänen, V. Koivunen, and M. Valkama, "Joint device positioning and clock synchronization in 5G ultra-dense networks," *IEEE Trans. Wirel. Commun.*, vol. 16, no. 5, pp. 2866–2881, May 2017.

[4] L. Li, R. Ying, Y. Li, L. He, and P. S. R. Diniz, "RIS array diagnosis for mmWave communication systems," *IEEE Signal Process. Lett.*, vol. 31, pp. 1980–1984, Jul. 2024.

[5] P.-J. Chung, M. Viberg, and J. Yu, "DOA estimation methods and algorithms," in *Academic Press Library in Signal Processing*. Elsevier, 2014, vol. 3, pp. 599–650.

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

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] E. A. Santiago and M. Saquib, "Noise subspace-based iterative technique for direction finding," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 49, no. 4, pp. 2281–2295, Oct. 2013.

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

[13] E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Math.*, vol. 346, no. 9-10, pp. 589–592, May 2008.

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

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

[16] H. Huang, H. C. So, and A. M. Zoubir, "Off-grid direction-of-arrival estimation using second-order Taylor approximation," *Signal Process.*, vol. 196, p. 108513, Jul. 2022.

[17] H. Zamani, H. Zayyani, and F. Marvasti, "An iterative dictionary learning-based algorithm for DOA estimation," *IEEE Commun. Lett.*, vol. 20, no. 9, pp. 1784–1787, Sep. 2016.

[18] W. Tan, X. Feng, W. Tan, G. Liu, X. Ye, and C. Li, "An iterative adaptive dictionary learning approach for multiple snapshot DOA estimation," in *Proc. IEEE Int. Conf. Signal Process.*, Beijing, China, Aug. 2018, pp. 214–219.

[19] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, "Compressed sensing off the grid," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7465–7490, Nov. 2013.

[20] Y. Chi, "Guaranteed blind sparse spikes deconvolution via lifting and convex optimization," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 4, pp. 782–794, Jun. 2016.

[21] Y. Chi and M. F. Da Costa, "Harnessing sparsity over the continuum: Atomic norm minimization for superresolution," *IEEE Signal Process. Mag.*, vol. 37, no. 2, pp. 39–57, Mar. 2020.

[22] B. N. Bhaskar, G. Tang, and B. Recht, "Atomic norm denoising with applications to line spectral estimation," *IEEE Trans. Signal Process.*, vol. 61, no. 23, pp. 5987–5999, Dec. 2013.

[23] A. Paulraj and T. Kailath, "Direction of arrival estimation by eigenstructure methods with unknown sensor gain and phase," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, vol. 10, Tampa, FL, USA, Apr. 1985, pp. 640–643.

[24] B. Friedlander and A. Weiss, "Eigenstructure methods for direction finding with sensor gain and phase uncertainties," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, New York, NY, USA, Apr. 1988, pp. 2681–2684.

[25] A. Liu, G. Liao, C. Zeng, Z. Yang, and Q. Xu, "An eigenstructure method for estimating DOA and sensor gain-phase errors," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5944–5956, Dec. 2011.

[26] M. Wylie, S. Roy, and H. Messer, "Joint DOA estimation and phase calibration of linear equispaced (LES) arrays," *IEEE Trans. Signal Process.*, vol. 42, no. 12, pp. 3449–3459, Dec. 1994.

[27] 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.

[28] L. Zhao, H. Liu, Y. Li, and Y. Zhou, "DOA estimation under sensor gain and phase uncertainties," in *Proceedings of International Conference on Estimation, Detection and Information Fusion*, Harbin, China, Jan. 2015, pp. 209–213.

[29] H. Liu, L. Zhao, Y. Li, X. Jing, and T.-K. Truong, "A sparse-based approach for DOA estimation and array calibration in uniform linear array," *IEEE Sens. J.*, vol. 16, no. 15, pp. 6018–6027, Aug. 2016.

[30] B. Hu, X. Wu, X. Zhang, Q. Yang, and W. Deng, "DOA estimation based on compressed sensing with gain/phase uncertainties," *IET Radar Sonar Navig.*, vol. 12, no. 11, pp. 1346–1352, Sep. 2018.

[31] P. Chen, Z. Chen, Z. Cao, and X. Wang, "A new atomic norm for DOA estimation with gain-phase errors," *IEEE Trans. Signal Process.*, vol. 68, pp. 4293–4306, Jul. 2020.

[32] 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.

[33] H. Huang, T. Zhang, F. Yin, B. Liao, and H. Wymeersch, "Joint DOA estimation and distorted sensor detection under entangled low-rank and row-sparse constraints," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Seoul, Korea, Apr. 2024, pp. 12 851–12 855.

[34] X.-Y. Wang, X.-P. Li, H. Huang, and H. C. So, "Robust DOA estimation with distorted sensors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 60, no. 5, pp. 5730–5741, Oct. 2024.

[35] 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, Feb. 2015.

[36] 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.

[37] 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.

[38] 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.

[39] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imaging Sci.*, vol. 2, no. 1, pp. 183–202, Mar. 2009.

[40] 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.

[41] Z. Lin, M. Chen, and Y. Ma, "The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices," *arXiv preprint arXiv:1009.5055*, Sep. 2010.

This article has been accepted for publication in IEEE Transactions on Aerospace and Electronic Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAES.2025.3553467

[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] Y. Xu and W. Yin, "A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion," *SIAM J. Imaging Sci.*, vol. 6, no. 3, pp. 1758–1789, Sep. 2013.

[44] P. Tseng and S. Yun, "A coordinate gradient descent method for nonsmooth separable minimization," *Math. Program.*, vol. 117, pp. 387–423, 2009.

[45] L. Mirsky, "A trace inequality of John von Neumann," *Monatshefte für math.*, vol. 79, no. 4, pp. 303–306, 1975.

[46] X.-Y. Wang, X. P. Li, and H. C. So, "Truncated quadratic norm minimization for bilinear factorization based matrix completion," *Signal Process.*, vol. 214, p. 109219, Jan. 2024.

[47] 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.

[48] 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.

[49] Y. Xu, "Alternating proximal gradient method for sparse nonnegative Tucker decomposition," *Math. Program. Comput.*, vol. 7, pp. 39–70, 2015.

[50] X. Li, S. Wang, and Y. Cai, "Tutorial: Complexity analysis of singular value decomposition and its variants," *arXiv preprint arXiv:1906.12085*, Jun. 2019.

[51] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, no. 8, pp. 1408–1418, Aug. 1969.

[52] D. P. Bertsekas, "Nonlinear programming," *J. Oper. Res. Soc.*, vol. 48, no. 3, pp. 334–334, Dec. 2017.

[53] 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.

[54] L. Van den Dries and C. Miller, "Geometric categories and o-minimal structures," *Duke Math. J.*, vol. 84, no. 2, pp. 497–540, Aug. 1996.

Xiang-Yu Wang received the BE degree from Northwestern Polytechnical University, Xi'an, China, in 2021. He is currently working toward the PhD degree in the Department of Electrical Engineering, City University of Hong Kong. His research interests include array signal processing and tensor analysis.

Xiao-Peng Li (M'23) 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. He was a Research Assistant with the Department of Information Engineering, Shenzhen University, Shenzhen, China from 2018 to 2019, and a Postdoctoral Fellow with the Department of Electrical Engineering, City University of Hong Kong from 2022 to 2023. He is currently an Research Associate 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 (Member, IEEE) is a Marie Curie postdoctoral researcher of the Communication Systems Group at Chalmers University of Technology, Gothenburg, Sweden. Prior to that, he received the Ph.D. degree from Darmstadt University of Technology, Darmstadt, Germany, in 2023. He received the B.Eng. and M.Sc. degrees from Shenzhen University, Shenzhen, China, in 2015 and 2018, respectively, all in Electrical and Electronic Engineering.

His research interests lie in signal processing and wireless communications, with main focuses 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 (S'90–M'95–SM'07-F'15) was born in Hong Kong. He received the B.Eng. degree from the City University of Hong Kong and the Ph.D. degree from The Chinese University of Hong Kong, both in electronic engineering, in 1990 and 1995, respectively. From 1990 to 1991, he was an Electronic Engineer with the Research and Development Division, Everex Systems Engineering Ltd., Hong Kong. During 1995–1996, 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.

He 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).