From Partial Calibration to Full Potential: A Two-Stage Sparse DOA Estimation for Incoherently-Distributed Sources with Partly-Calibrated Arrays

He Xu, Tuo Wu, Wei Liu, Maged Elkashlan, Naofal Al-Dhahir, *Fellow, IEEE*, Mérouane Debbah, *Fellow, IEEE*, Chau Yuen, *Fellow, IEEE*, and Hing Cheung So, *Fellow, IEEE*

***Abstract*—Direction-of-arrival (DOA) estimation for incoherently distributed (ID) sources is crucial for Industrial Internet of Things (IIoT) applications operating in complex multipath environments, yet it remains challenging due to the combined effects of angular spread and gain-phase uncertainties in cost-sensitive antenna arrays. This paper presents a two-stage sparse DOA estimation framework, transitioning from *partial calibration to full potential*, under the generalized array manifold (GAM) framework. In the first stage, coarse DOA estimates are obtained by exploiting the output from a subset of partly-calibrated arrays (PCAs). In the second stage, these estimates are utilized to determine and compensate for gain-phase uncertainties across all array elements. Then a sparse total least-squares optimization problem is formulated and solved via alternating descent to refine the DOA estimates. Simulation results demonstrate that the proposed method achieves superior estimation accuracy compared to existing approaches, while maintaining robustness against both noise and angular spread effects in practical industrial environments.**

***Index Terms*—Direction-of-arrival (DOA) estimation, Industrial Internet of Things (IIoT).**

I. Introduction

DIRECTION-of-arrival (DOA) estimation plays a crucial role in emerging Industrial Internet of Things (IIoT) and Internet of Vehicles (IoV) applications [1]–[3], and massive multiple-input multiple-output (MIMO) communications [4], [5]. In these scenarios, accurate DOA estimation faces unprecedented challenges due to complex urban environments, dense network deployments, and dynamic mobility patterns. For instance, in autonomous driving, precise DOA estimation is essential for obstacle detection and vehicle-to-vehicle communication, while in smart factories, it enables efficient asset tracking and robotic navigation. These applications typically operate in environments with rich scattering and multiple reflections, making the multi-path propagation (MP) model indispensable for accurate system modeling [6]. Among the two primary MP models, coherently-distributed (CD) and incoherently-distributed (ID) sources [7], ID sources better characterize the practical wireless channels in such complex scenarios, where signals often arrive through multiple uncorrelated paths due to diverse scattering effects.

Over the years, numerous ID source estimation methods have been proposed, which is able to address IoT positioning challenges, including classical techniques such as the covariance matching estimator (COMET) [8], ESPRIT-based method [9], and beamspace transformation-based estimator (BTBET) [10]. For IIoT applications with mixed far-field and near-field ID sources, such as in smart warehouses where both stationary and mobile devices coexist, rank-reduction (RARE) estimator [11] and its two-stage variant [12] have shown promise. However, these subspace-based approaches typically require prior knowledge of the number of sources, which is impractical in dynamic IoT networks where devices frequently join and leave the network. Moreover, subspace methods often suffer performance degradation when the signal subspace dimension is uncertain or when noise levels are high, highlighting the need for more resilient alternatives. Another significant issue not adequately addressed by these methods is the gain-phase uncertainty inherent in practical radio frequency (RF) chains [13], which is particularly problematic in cost-sensitive IoT deployments where low-cost hardware is prevalent. Both theoretical analysis and algorithms investigations confirm that these uncertainties can severely degrade estimation accuracy [14]–[16], potentially compromising the reliability of IoT services. Although recent studies tackled this challenge through partly-calibrated arrays (PCAs) [17] and augmented generalized array manifold (GAM) approaches [18], most solutions focus on isolated aspects, such as merely compensating gain-phase errors, while leaving other critical factors, especially the angular spread in ID sources caused by multipath propagation, unaddressed in a unified framework. This is especially concerning for the angular spread in ID sources caused by multipath propagation in complex IoT environments, where devices are often deployed in cluttered indoor spaces or urban areas.

To overcome these shortcomings, sparse reconstruction provides a promising solution for practical IoT deployments. Unlike subspace-based algorithms, sparsity-driven approaches do not hinge on prior knowledge of the number of sources, thereby achieving better performance in dynamic IoT networks with varying numbers of active devices. Besides, sparse methods inherently accommodate practical imperfections, such as gain-phase uncertainties common in low-cost IoT hardware, making them well-suited for the combined challenges posed by ID sources and multipath propagation. To this end, we propose a two-stage DOA estimator for ID sources that fully

This work was partially supported by the Zhejiang Provincial Natural Science Foundation of China under Grant LQN26F010011. (*Corresponding author: Tuo Wu.*)

H. Xu is with the School of Cyber Science and Engineering, Ningbo University of Technology, Ningbo 315211, China (E-mail: xuhebest@sina.com). T. Wu and C. Yuen are with the School of Electrical and Electronic Engineering, Nanyang Technological University 639798, Singapore (E-mail: {tuo.wu, chau.yuen}@ntu.edu.sg). W. Liu is with the Department of Electrical and Electronic Engineering, Hong Kong Polytechnic University, Kowloon, Hong Kong (E-mail: wliu.eee@gmail.com). M. Elkashlan is with the School of Electronic Engineering and Computer Science at Queen Mary University of London, London E1 4NS, U.K. (E-mail: maged.elkashlan@qmul.ac.uk). Naofal Al-Dhahir is with the Department of Electrical and Computer Engineering, The University of Texas at Dallas, Richardson, TX 75080 USA (E-mail: aldhahir@utdallas.edu). M. Debbah is with KU 6G Research Center, Department of Computer and Information Engineering, Khalifa University, Abu Dhabi 127788, UAE (E-mail: merouane.debbah@ku.ac.ae). H. C. So is with the Department of Electrical Engineering City University of Hong Kong, Hong Kong, China. (E-mail: hcso@ee.cityu.edu.hk).

capitalizes on the sparsity framework and leverages PCAs, while adopting GAM to accurately capture the small angular deviations characteristic of ID sources in IoT environments. Our approach proceeds in two stages: First, a set of consecutive calibrated sensors is used to extract the first column of the sample covariance matrix, and we form an augmented covariance vector to simplify the initial estimation while mitigating noise effects and partial angular spread interference. Subsequently, the coarse DOA estimates enable us to identify and compensate for gain-phase uncertainties in the uncalibrated sensors, allowing a refined covariance structure to be constructed using all array elements. This refined structure is then processed via a sparse total least-squares minimization, further enhancing DOA accuracy for reliable IoT positioning services. Simulation results demonstrate that this two-stage strategy outperforms competing subspace estimators across a broad range of SNR levels, making it particularly suitable for challenging IoT deployment scenarios.

Notations: **A**/**a** represent matrices/vectors, $(\cdot)^*$, the superscripts $(\cdot)^T$, $(\cdot)^{-1}$ and $(\cdot)^H$ denote the conjugate, transpose, inverse and conjugate transpose, respectively. $\|\cdot\|_F$, $\|\cdot\|_2$ and $\|\cdot\|_1$ denote the Frobenius norm, ℓ_2 norm and ℓ_1 norm, respectively. $\mathbb{E}\{\cdot\}$ is the statistical expectation, $\mathrm{diag}(\cdot)$ the diagonalization operation, and $\odot$ the Hadamard-Schur product. $\mathbf{\Pi}_M$ denotes an $M \times M$ exchange matrix, $\mathbf{I}_M$ an $M \times M$ identity matrix and $\mathbf{1}_{1\times M}$ a $1 \times M$ all-one row vector.

II. Problem Formulation

Consider K ID sources impinging on an M-element uniform linear array (ULA) with half wavelength inter-element spacing, and each ID source signal propagates along L distinct paths. Suppose that there are M_c consecutive calibrated sensors, then the array output after considering the gain-phase uncertainties at time instant t can be expressed as [11], [12]

$$\mathbf{z}(t) = \mathbf{\Xi}(\mathbf{g}) \sum_{k=1}^{K} s_k(t) \sum_{l=1}^{L} \gamma_{k,l}(t)\mathbf{a}(\theta_{k,l}(t)) + \mathbf{e}(t), \quad (1)$$

where $\mathbf{\Xi}(\mathbf{g}) = \mathrm{diag}(\mathbf{g})$ stands for the gain-phase uncertainty matrix with $\mathbf{g} = [\mathbf{1}_{1\times Mc}, g_{M_c+1}, \ldots, g_M]^T$ with $g_m = \rho_m e^{j\phi_m}$, $m \in [M_c+1, M]$, while ρ_m and ϕ_m represent the gain and phase uncertainties, respectively. $\{s_k(t)\}_{k=1}^K$ are K far-field narrowband uncorrelated ID signals and $\mathbf{e}(t)$ is the additive Gaussian white noise. $\gamma_{k,l}(t)$ is the complex path gain of the l-th path associated with the k-th ID source, which is zero-mean and temporally independent and identically distributed (i.i.d.) random variable with unknown variance $\sigma^2_{\gamma_k}/L$. $\mathbf{a}(\theta_{k,l}(t))$ represents the array steering vector of $\theta_{k,l}(t)$, with $\theta_{k,l}(t) = \theta_k + \tilde{\theta}_{k,l}(t)$, θ_k is the DOA of the k-th signal, and $\tilde{\theta}_{k,l}(t)$ is the real-valued angular deviation, which normally follows Gaussian or uniform distribution with unknown variance $\sigma^2_{\theta_k}$ [11], [19].

To simplify the mathematical modeling of multipath propagation in array signal processing, the GAM model with a small $\sigma^2_{\theta_k}$ is adopted, where $\mathbf{a}(\theta_{k,l}(t))$ is approximated as [6], [9]

$$\mathbf{a}(\theta_{k,l}(t)) \approx \mathbf{a}(\theta_k) + \mathbf{a}'(\theta_k)\tilde{\theta}_{k,l}(t), \quad (2)$$

where $\mathbf{a}'(\theta_k) = \partial\mathbf{a}(\theta_k)/\theta_k$ and

$$\mathbf{a}(\theta_k) = [1, e^{-j\pi\sin\theta_k}, \ldots, e^{-j(M-1)\pi\sin\theta_k}]^T. \quad (3)$$

Consequently, $\mathbf{z}(t)$ is expressed as

$$\mathbf{z}(t) \approx \mathbf{\Xi}(\mathbf{g})(\mathbf{A}\mathbf{s}(t) + \mathbf{A}'\tilde{\mathbf{s}}(t)) + \mathbf{e}(t), \quad (4)$$

where $\mathbf{A} = [\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_K)]$, $\mathbf{A}' = [\mathbf{a}'(\theta_1), \ldots, \mathbf{a}'(\theta_K)]$, $\mathbf{s}(t) = [s_1(t), \ldots, s_K(t)]^T$, $\tilde{\mathbf{s}}(t) = [\tilde{s}_1(t), \ldots, \tilde{s}_K(t)]^T$, $s_k(t) = s_k(t)\sum_{l=1}^{L}\gamma_{k,l}(t)$, and $\tilde{s}_k(t) = s_k(t)\sum_{l=1}^{L}\gamma_{k,l}(t)\tilde{\theta}_{k,l}(t)$.

III. Two-Stage DOA Estimation

A. First-Stage DOA Estimation

Collecting N snapshots, the matrix form of array output can be written as $\mathbf{Z}(t) = [\mathbf{z}(1), \mathbf{z}(2), \ldots, \mathbf{z}(N)]$. Based on the statistical expectation and sample average approximation, its covariance matrix can be calculated by

$$\begin{aligned}\mathbf{R} &= \mathbf{\Xi}(\mathbf{g})\mathbf{A}\mathbf{R}_s\mathbf{A}^H\mathbf{\Xi}^H(\mathbf{g}) + \mathbf{\Xi}(\mathbf{g})\mathbf{A}'\mathbf{R}_{\tilde{s}}\mathbf{A}'^H\mathbf{\Xi}^H(\mathbf{g}) + \sigma_n^2\mathbf{I}_M \\ &\approx 1/N\sum_{t=1}^{N}\mathbf{z}(t)\mathbf{z}^H(t). \end{aligned} \quad (5)$$

where $\mathbf{R}s = \mathbb{E}\{s(t)s^H(t)\}$ is the source covariance matrix, and $\mathbf{R}\tilde{s} = \mathbb{E}\{\tilde{s}(t)\tilde{s}^H(t)\}$ is the covariance matrix of the angular spread components. Further, it should be noted that the first row of $\mathbf{A}'$ (the derivative matrix of the steering vector) is a $1 \times K$ all-zero vector. This property indicates that the first array element is immune to angular spread effects. Therefore, to eliminate the impact of angular spread, we extract the first $M_c \times 1$ elements from $\mathbf{R}$ to form the covariance vector $\mathbf{r}_c$ (see (6) below), which corresponds to the well-calibrated array elements. Assuming that the noise variance σ_n^2 is known *a priori* or pre-estimated via the maximum likelihood criterion, and subtracting the noise term from (5), $\mathbf{r}_c$ is given by

$$\mathbf{r}_c \approx \mathbf{A}_c\mathbf{R}_s\mathbf{1}_{K\times 1} = \mathbf{A}_c\mathbf{p}, \quad (6)$$

where $\mathbf{p} = \mathbf{R}_s\mathbf{1}_{K\times 1} = [p_1, \ldots, p_K]^T$ with p_k being the power of $s_k(t)$, and $\mathbf{A}_c$ represents the first M_c rows of $\mathbf{A}$. Since $\mathbf{p}$ is a positive vector, to further improve the resolution of DOA estimation, the conjugate symmetry strategy is adopted to construct the following augmented covariance vector as $\mathbf{r}_1 = \left[[\mathbf{\Pi}_{M_c}\mathbf{r}_c^*]^T, [\mathbf{r}_c(2:M_c]^T)\right]^T \approx \mathbf{B}\mathbf{p}$, where $\mathbf{B} = [\mathbf{b}(\theta_1), \ldots, \mathbf{b}(\theta_K)]$ with its k-th column given by $\mathbf{b}(\theta_k) = \left[e^{j(M_c-1)\pi\sin\theta_k}, \ldots, 1, \ldots, e^{-j(M_c-1)\pi\sin\theta_k}\right]^T$.

Let $\mathbf{\Phi}$ be the sparse representation matrix of $\mathbf{B}$, and $\mathbf{p}_G$ the K-sparsity vector with respect to $\mathbf{p}$. Then, initial DOA estimates can be obtained by solving the following ℓ_2-ℓ_1 norm minimization problem

$$\begin{aligned} &P_1 : \hat{\mathbf{p}}_G = \arg\min_{\mathbf{p}_G} \|\mathbf{r}_1 - \mathbf{B}\mathbf{p}_G\|_2^2 + \lambda\|\mathbf{p}_G\|_1, \\ &P_2 : \{\hat{\theta}_k\}_{k=1}^K = \arg\max_{\theta} f_d(\hat{\mathbf{p}}_G), \theta \in (-90^\circ, 90^\circ], \end{aligned} \quad (7)$$

where $f_d(\hat{\mathbf{p}}_G)$ returns the indices of nonzero elements in $\hat{\mathbf{p}}_G$, and λ is the penalty parameter and selected here by the L-curve method.

B. Second-Stage DOA Estimation

With the use of the DOA estimates $\{\hat{\theta}_k\}_{k=1}^K$, we can reconstruct the array manifold matrices $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ corresponding to these estimated angles. Here, $\hat{\mathbf{A}} = [\mathbf{a}(\hat{\theta}_1), \ldots, \mathbf{a}(\hat{\theta}_K)]$ contains the steering vectors for all M array elements, while

$\hat{\mathbf{B}} = [\mathbf{b}(\hat{\theta}_1), \ldots, \mathbf{b}(\hat{\theta}_K)]$ represents the augmented steering vectors using only the first M_c calibrated elements. This allows us to obtain a refined power estimate $\hat{\mathbf{p}} = (\hat{\mathbf{B}}^H \hat{\mathbf{B}})^{-1} \hat{\mathbf{B}}^H \mathbf{r}_1$. To estimate the gain-phase uncertainties, we extract the first column of $\mathbf{R}$ to form $\mathbf{r}_2$ which represents the correlation between all array elements and the first (calibrated) element:

$$\mathbf{r}_2 \approx \mathbf{\Xi}(\mathbf{g}) \mathbf{A} \mathbf{R}_s \mathbf{1}_{K\times 1} = \mathbf{\Xi}(\mathbf{g}) \mathbf{A}\mathbf{p} = \mathbf{g} \odot \mathbf{A}\mathbf{p}. \tag{8}$$

The gain-phase uncertainty of each uncalibrated element can then be estimated by comparing the actual observation $\mathbf{r}_2$ with the ideal response $\mathbf{v} = \hat{\mathbf{A}}\hat{\mathbf{p}}$: $\hat{g}_m = \mathbf{r}_2(m) / \mathbf{v}(m), m \in [M_c + 1, M]$, where $\mathbf{v}(m)$ stands for the m-th element of $\hat{\mathbf{A}}\hat{\mathbf{p}}$. Accordingly, we now utilize all M array elements for improved DOA estimation. First, we reconstruct the gain-phase uncertainty matrix $\hat{\mathbf{\Xi}}$ and compensate its effect in $\mathbf{r}_2$:

$$\mathbf{r}_3 = (\hat{\mathbf{\Xi}}^H(\hat{\mathbf{g}})\hat{\mathbf{\Xi}}(\hat{\mathbf{g}}))^{-1}\hat{\mathbf{\Xi}}^H(\hat{\mathbf{g}})\mathbf{r}_2 \approx \mathbf{A}\mathbf{p}. \tag{9}$$

With the compensated array response $\mathbf{r}_3$, we can now apply the same conjugate symmetry strategy as in the first stage, but this time using all M array elements instead of just M_c calibrated ones: $\mathbf{r}_4 = \left[[\mathbf{\Pi}_M \mathbf{r}_3^*]^T, [\mathbf{r}_3(2:M)]^T\right]^T \approx \mathbf{C}\mathbf{p}$, where $\mathbf{C} = [\mathbf{c}(\theta_1), \ldots, \mathbf{c}(\theta_K)]$ with its k-th column given by $\mathbf{c}(\theta_k) = \left[e^{j(M-1)\pi \sin\theta_k}, \ldots, 1, \ldots, e^{-j(M-1)\pi \sin\theta_k}\right]^T$.

On the other hand, it is emphasized that the first-stage DOA estimation does not consider the impact of the finite number of snapshots and the model bias between the GAM and true array manifold. Taking both adverse factors into consideration, we employ the perturbation matrix $\mathbf{\Gamma}$ and rewrite $\mathbf{r}_4$ accurately in a sparse representation manner as $\mathbf{r}_4 = (\mathbf{\Psi} + \mathbf{\Gamma})\mathbf{p}_{\bar{G}}$, where $\mathbf{\Psi}$ denotes the sparse representation of $\mathbf{C}$.

Next, the following sparse total least-squares minimization optimization problem is constructed to achieve robust and refined DOA estimation:

$$\begin{aligned} \{\hat{\mathbf{p}}_{\bar{G}}, \hat{\mathbf{\Gamma}}\} &= \arg\min_{\mathbf{p}_{\bar{G}}, \mathbf{\Gamma}} \|\mathbf{\Gamma}\|_F^2 + \lambda \|\mathbf{p}_{\bar{G}}\|_1 \\ &s.\,t.\ \ \mathbf{r}_4 = (\mathbf{\Psi} + \mathbf{\Gamma})\mathbf{p}_{\bar{G}} \end{aligned} \tag{10}$$

which can be relaxed as an unconstrained form as:

$$\{\hat{\mathbf{p}}_{\bar{G}}, \hat{\mathbf{\Gamma}}\} = \arg\min_{\mathbf{p}_{\bar{G}}, \mathbf{\Gamma}} \|\mathbf{r}_4 - (\mathbf{\Psi} + \mathbf{\Gamma})\mathbf{p}_{\bar{G}}\|_2^2 + \|\hat{\mathbf{\Gamma}}\|_F^2 + \lambda \|\mathbf{p}_{\bar{G}}\|_1. \tag{11}$$

By exploiting the alternating descent algorithm, (11) can be solved efficiently. Given $\hat{\mathbf{\Gamma}}^{(i)}$ at the $(i-1)$-th iteration with its initial value $\hat{\mathbf{\Gamma}}^{(0)} = \mathbf{0}$, $\mathbf{p}_{\bar{G}}$ is updated at i-th iteration as

$$\hat{\mathbf{p}}_{\bar{G}}^{(i)} = \arg\min_{\mathbf{p}_{\bar{G}}} \|\mathbf{r}_4 - (\mathbf{\Psi} + \hat{\mathbf{\Gamma}}^{(i-1)})\mathbf{p}_{\bar{G}}\|_2^2 + \lambda \|\mathbf{p}_{\bar{G}}\|_1 \tag{12}$$

and with available $\hat{\mathbf{p}}_{\bar{G}}^{(i)}$, $\hat{\mathbf{\Gamma}}^{(i)}$ is updated as

$$\hat{\mathbf{\Gamma}}^{(i)} = \arg\min_{\mathbf{\Gamma}} \|\mathbf{r}_4 - (\mathbf{\Psi} + \mathbf{\Gamma})\hat{\mathbf{p}}_{\bar{G}}^{(i)}\|_2^2 + \|\mathbf{\Gamma}\|_F^2. \tag{13}$$

The optimization subproblems (10) and (12) are both convex, and therefore can be solved efficiently [1]. The iteration is terminated when $\|\hat{\mathbf{p}}_{\bar{G}}^{(i)} - \hat{\mathbf{p}}_{\bar{G}}^{(i-1)}\|_2 < \varepsilon$ or the user defined maximum number of iterations I_t is reached. Finally, the improved DOA estimates $\{\hat{\theta}_k^{(f)}\}_{k=1}^K$ are computed by identifying the positions of the K non-zero elements in $\hat{\mathbf{p}}_{\bar{G}}^{(I_t)}$.

The computational complexity of this method mainly stems from calculating array covariance vectors ($\mathcal{O}(MN)$) and solving two optimization problems: (7) with complexity $\mathcal{O}(G^3)$ and (11) with complexity $\mathcal{O}(I_t\bar{G}^3 + I_t(2M-1)\bar{G})$. Since $\bar{G} \ll G$, the overall complexity is approximately $\mathcal{O}(MN + G^3)$. Compared to subspace-based techniques like RARE, which requires $\mathcal{O}(M^2N + M^3 + G(8K^3 + 4(M-1)K^2 + 2(M-1)K))$, our method is computationally more efficient when $2\sqrt{M-1}K$ is comparable with G.

IV. Simulation Results

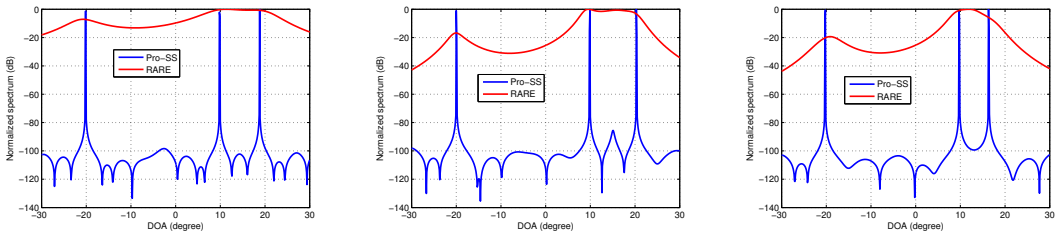
In this section, the performance of the two-stage DOA estimation method is assessed, and compared with that of the RARE in [12], ESPRIT-Like[2] in [17], conventional MUSIC using first M_c elements (designated as MUSIC-FS) and Cramér-Rao Bound (CRB) in [17], where the first-stage and second-stage estimations are respectively designated as Pro-FS and Pro-SS. It should be noted that the RARE based method is not suitable for dealing with array gain-phase uncertainties, and thus we utilize the gain-phase uncertainty estimation result obtained from our method to compensate it for comparison. In the simulations, a ULA with $M = 16$ sensors is employed and the root mean square error (RMSE) obtained by averaging 300 independent Monte-Carlo trials is selected to evaluate the performance of different methods. The regularization parameter λ selected by the L-curve method is $\lambda = 1.85$, the iteration associated with the sparse total least-squares minimization optimization problem is terminated when $I_t = 3$ or $\varepsilon = 10^{-3}$. Meanwhile, the number of paths for each ID source is $L = 100$. The gain and phase uncertainties are respectively modeled as $\rho_m = 1 + \sqrt{12}\sigma_\rho \eta_m$, $\phi_m = \sqrt{12}\sigma_\phi \mu_m$ [13], [17], [18], where η_m and μ_m are uniformly i.i.d. random variables within $[-0.5, 0.5]$, and σ_ρ and σ_ϕ represent the standard deviations of η_m and μ_m, respectively.

In the first simulation, the normalized spatial spectra of the proposed method and RARE for three ID sources are plotted in Fig. 1, with $\sigma_\rho = 0.1$ and $\sigma_\phi = 40°$. The parameters of three ID sources $(\theta_1, \theta_2, \theta_3) = (-20°, 10°, 20°)$, $\sigma_\theta = \{\sigma_{\theta_k}\}_{k=1}^3 = 1.5°$, with $N = 200$ and SNR=-6 dB in Fig. 1(a), from which we observe that the proposed method resolve three sources well, albeit with some estimation bias, while the compared RARE fails. In Fig. 1(b), SNR is increased to 6 dB, and we see that although RARE distinguish two closely-spaced sources, its spectral peaks are much less pronounced. In contrast, our solution distinguish them very clearly. In Fig. 1(c), we fix SNR at 6 dB, while setting the minimum angle separation to 6° and $\sigma_\theta = 2.5°$.

It is observed from Fig. 2 that the proposed method provide an improved performance compared with the remaining algorithms. Meanwhile, due to the increase of DOFs and robustness to perturbations of the sparse total least-squares approach adopted at the second stage, Pro-SS performs better

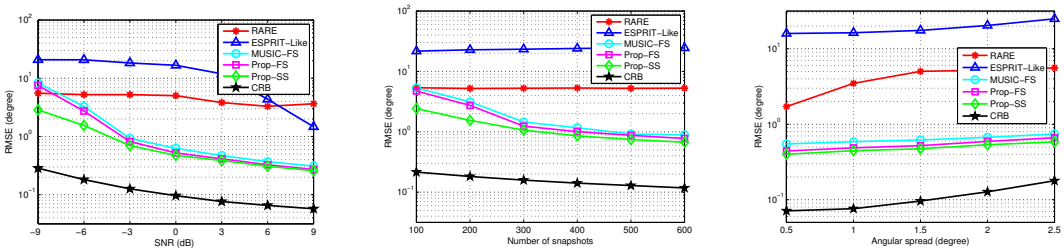
[1]With the available DOA estimates, angular spread estimation can be readily obtained by applying the approach in [11]. This additional parameter can provide valuable insights in scenarios requiring comprehensive spatial characterization of the signal environment, such as in multipath-rich IoT deployments or when characterizing scattering environments.

[2]Such an ESPRIT-Like refers to the application of the algorithm proposed in [17] under a ULA, and is also an extension of point-source model based ESPRIT-Like [20] for ID sources.



(a) (b) (c)

Fig. 1. Normalized spectrum for Pro-SS and RARE under different settings: (a) $\sigma_\theta = 1.5°$, SNR=-6 dB, minimum angle separation = $10°$; (b) $\sigma_\theta = 1.5°$, SNR=6 dB, minimum angle separation = $10°$; (c) $\sigma_\theta = 2.5°$, SNR=6 dB, minimum angle separation = $6°$.



(a) (b) (c)

Fig. 2. RMSE of DOA estimates for different methods: (a) RMSE versus SNR; (b) RMSE versus number of snapshots; (c) RMSE versus angular spread.

than Pro-FS and MUSIC-FS in the whole observed regions, especially under low SNRs and small number of snapshots, showing the effectiveness and superiority of the two-stage DOA estimator. In addition, although the performance of MUSIC-FS is close to that of Pro-FS, it should be emphasized that Pro-FS does not require prior knowledge of the number of sources, making it more competitive. On the other hand, it can be also seen that the performance improvement of Pro-SS gradually slows down, with the increase of SNR and snapshot number. This is because the gain-phase uncertainties, as well as the model bias between the GAM and true array manifold, cannot be perfectly suppressed. Under relatively high SNR and large snapshot conditions, their influence gradually becomes dominant, leading to a slight decrease in the level of performance improvement in the second stage. In Fig. 2(c), SNR and N are respectively set to 0 dB and 200, two DOAs remain unchanged, but the angular spread σ_θ for both ID sources varies from $0.5°$ to $2.5°$.

V. Conclusion

In this paper, we have presented a two-stage sparse DOA estimation framework for ID sources under gain-phase uncertainties. The first stage aims to obtain initial DOA estimates using CPAs, while the second stage has compensated for gain-phase uncertainties and has refined the estimation through sparse total least-squares optimization.

References

[1] R. S. Zakariyya, *et al.* "Joint DoA and CFO estimation scheme with received beam scanned leaky wave antenna for industrial internet of things (IIoT) systems," *IEEE Internet of Things J.*, vol. 10, no. 15, pp. 13686-13696, Aug. 2023.

[2] Y. Cui, *et al.* "Gridless underdetermined DOA estimation of wideband LFM signals with unknown amplitude distortion based on fractional fourier transform," *IEEE Internet of Things J.*, vol. 7, no. 12, pp. 11612-11625, Dec. 2020.

[3] M. Lin, *et al.* "Single sensor to estimate DOA with programmable metasurface," *IEEE Internet of Things J.*, vol. 8, no. 12, pp. 10187-10197, Jun. 2021.

[4] R. Zhang, *et al.* "Channel-training-aided target sensing for terahertz integrated sensing and massive MIMO communications," *IEEE Internet of Things J.*, vol. 12, no. 4, pp. 3755-3770, 15 Feb.15, 2025.

[5] R. Zhang, *et al.* "Direction-of-arrival estimation for large antenna arrays with hybrid analog and digital architectures," *IEEE Trans. Signal Process.*, vol. 70, pp. 72-88, 2022.

[6] Y. Tian, *et al.*. "Vehicle positioning with deep-learning-based direction-of-arrival estimation of incoherently distributed sources," *IEEE Internet of Things J.*, vol. 9, no. 20, pp. 20083-20095, Oct. 2022.

[7] Y. Tian, *et al.*. "Localization of mixed coherently and incoherently distributed sources based on generalized array manifold," *Signal Process.*, vol. 209, pp. 109038, Aug. 2023.

[8] O. Besson, *et al.*. "Decoupled estimation of DOA and angular spread for a spatially distributed source," *IEEE Trans. Signal Process.*, vol. 48, no. 7, pp. 1872-1882, Jul. 2000.

[9] A. Hu, T. Lv, H. Gao, Z. Zhang and S. Yang, "An ESPRIT-based approach for 2-D localization of incoherently distributed sources in massive MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 996-1011, Oct. 2014.

[10] Z. Zheng, *et al.*. "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. 11776-11789, Dec. 2018.

[11] R. Cao, *et al.*. "An angular parameter estimation method for incoherently distributed sources via generalized shift invariance," *IEEE Trans. Signal Process.*,vol. 64, no. 7, pp. 4493-4503, Sept. 2016.

[12] Y. Tian, *et al.*. "Localization of mixed far-field and near-field incoherently distributed sources using two-stage RARE estimator," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 59, no. 2, pp. 1482-1494, Apr. 2023.

[13] H. Xu, *et al.*. "Direction of arrival estimation with gain-phase uncertainties in unknown nonuniform noise," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 59, no. 6, pp. 9686-9696, Dec. 2023.

[14] Z. Dai, *et al.*. "A gain and phase autocalibration approach for large-scale planar antenna arrays," *IEEE Commun. Lett.*, vol. 25, no. 5, pp. 1645-1649, May 2021.

[15] C. Lu, *et al.*. "DOA estimation based on coherent integration-sparse bayesian learning with time-variant gain-phase errors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 59, no. 6, pp. 7951-7962, Dec. 2023.

[16] Y. Liu, *et al.*. "Low-complexity DOA estimation for coherently distributed sources with gain-phase errors in massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 73, no. 6, pp. 7939-7948, Jun. 2024.

[17] Y. Tian, *et al.*. "2-D DOA estimation of incoherently distributed sources considering gain-phase perturbations in massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 1143-1155, Feb. 2022.

[18] Y. Liu, *et al.*. "Noncircularity-based DOA estimation for incoherently distributed sources with gain-phase uncertainties," *IEEE Sensors J.*, vol. 24, no. 16, pp. 26022-26033, Aug. 2024.

[19] H. Chen, *et al.*. "Two-dimensional angular parameter estimation for noncircular incoherently distributed sources based on an L-shaped array," *IEEE Sensors J.*, vol. 20, no. 22, pp. 13704-13715, Nov. 2020.

[20] B. Liao and S. C. Chan, "Direction finding with partly calibrated uniform linear arrays," *IEEE Trans. Antenna Propag.*, vol. 60, no. 2, pp. 922-929, Feb. 2012.