









## Efficient DOA Estimation in Hybrid Analog–Digital Structures: Mitigating the Impact of Mutual Coupling

Ye Tian , Hongyun Zhao, Tuo Wu , *Member, IEEE*,  
Wei Liu , *Senior Member, IEEE*, Maged El Kashlan ,  
Henk Wymeersch , *Fellow, IEEE*,  
Fumiyuki Adachi , *Life Fellow, IEEE*,  
George K. Karagiannidis , *Fellow, IEEE*,  
and Chau Yuen , *Fellow, IEEE*

**Abstract**—Hybrid analog-digital structures (HADS) have emerged as an efficient solution for mitigating transmission loss and reducing power consumption in multiple-input multiple-output (MIMO) systems. However, the limited number of radio frequency (RF) chains and the presence of array mutual coupling (MC) pose significant challenges to achieving high-performance direction-of-arrival (DOA) estimation, thereby hindering effective downlink beamforming. To address these challenges, an efficient DOA estimation method specifically designed for HADS is proposed, effectively mitigating the impact of MC by employing a two-stage framework. The first stage reconstructs the spatial covariance matrix (SCM) by adjusting switch states and leveraging a middle subarray, combined with the real-valued subspace technique for initial DOA estimation. Using these initial estimates, MC is modeled and compensated by adjusting amplifiers and phase shifters. In the second stage, enhanced DOA estimation is achieved by fully exploiting the data from the entire array. Simulation results validate the effectiveness of the proposed method, demonstrating its capability to mitigate the MC effect and deliver accurate DOA estimation under practical conditions.

**Index Terms**—Direction-of-arrival (DOA) estimation, hybrid analog-digital structure (HADS), mutual coupling (MC), real-valued subspace technique.

Received 17 February 2025; revised 1 June 2025; accepted 24 July 2025. Date of publication 31 July 2025; date of current version 19 January 2026. The work of Ye Tian was supported in part by the Natural Science Foundation of Ningbo Municipality under Grant 2024J232 and in part by the Zhejiang Provincial Natural Science Foundation of China under Grant LY23F010004. The work of Chau Yuen was supported in part by the Ministry of Education, Singapore, through MOE Tier 2 under Award T2EP50124-0032. The review of this article was coordinated by Prof. Zhiguo Shi. (*Corresponding author: Tuo Wu.*)

Ye Tian and Hongyun Zhao are with the Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo 315211, China (e-mail: tianfield@126.com; 2311100199@nbu.edu.cn).

Tuo Wu and Chau Yuen are with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798 (e-mail: tuo.wu@ntu.edu.sg; chau.yuen@ntu.edu.sg).

Wei Liu is with the Department of Electrical and Electronic Engineering, Hong Kong Polytechnic University, Kowloon, China (e-mail: wliu.eee@gmail.com).

Maged El Kashlan is with the School of Electronic Engineering and Computer Science, Queen Mary University of London, E1 4NS London, U.K. (e-mail: maged.elkashlan@qmul.ac.uk).

Henk Wymeersch is with the Department of Electrical Engineering, Chalmers University of Technology, 41296 Gothenburg, Sweden (e-mail: henkw@chalmers.se).

Fumiyuki Adachi is with the International Research Institute of Disaster Science (IRIDeS), Tohoku University, Sendai 980-0845, Japan (e-mail: adachi@ecei.tohoku.ac.jp).

George K. Karagiannidis is with the Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece, and also with the Artificial Intelligence and Cyber Systems Research Center, Lebanese American University, Beirut 0379775, Lebanon (e-mail: geokarag@auth.gr).

Digital Object Identifier 10.1109/TVT.2025.3594548

## I. INTRODUCTION

Direction-of-arrival (DOA) estimation plays a critical role for various applications, including base station (BS) downlink beamforming, integrated sensing and communication (ISAC), and interference suppression [1], [2], [3]. Numerous DOA estimation methods have been proposed, such as those based on subspace processing [4], [5], sparse reconstruction [6], [7], deep learning [8], [9], [10], as well as efficient DOA methods based on these methods [11]. However, these methods rely on fully digital structures (FDS), where each antenna is connected to a dedicated radio frequency (RF) chain. With the advent of millimeter-wave and Terahertz communications, the high power consumption of large-scale arrays makes FDS impractical. As an alternative, hybrid analog-digital structures (HADS) with fewer RF chains have gained significant attention as an efficient solution to balance low power consumption and reliable communication performance [12].

Due to the analog combination of array outputs and the significant reduction of RF chains, traditional FDS-based methods are not directly applicable, posing challenges for high-performance DOA estimation. To address this, several HADS-specific methods have been proposed. An analog phase alignment (APA) method in [13] maximized array output power with low complexity but suffered from Rayleigh limitations, failing to fully utilize the array aperture. Alternatively, the spatial covariance matrix (SCM) is constructed via beam sweeping, and the DOA estimation is improved using the MUSIC algorithm [14]. Building on this, modified SCM reconstruction methods employing root-MUSIC and unitary root-MUSIC algorithms were introduced in [15], [16], [17], supporting single and multiple RF chain configurations. Furthermore, Liu et al. [18] proposed a low-computational-cost alternating switches algorithm for real-valued SCM reconstruction in HADS. Additionally, beam-space DOA methods [19], which are applicable to both analog and hybrid arrays, have also attracted significant attention for their potential in enhancing DOA estimation performance.

The aforementioned DOA estimation methods show promise for HADS, but they are typically derived ideal array conditions. In practical scenarios, however, the presence of array mutual coupling (MC) significantly complicates the DOA estimation process [20], [21]. Specifically, MC introduces uncertainties in the array manifold, which distort signal correlations and degrade array responses, making it difficult to accurately reconstruct the full-dimensional SCM required for super-resolution DOA estimation methods. These effects result in biased angle estimates and reduced resolution, rendering traditional methods ineffective for accurate DOA estimation. To address these challenges, a two-stage DOA estimation method that tackles both HADS and the unknown MC effect is proposed. Our approach integrates switch-state adjustment, middle subarray utilization, and SCM reconstruction techniques, effectively mitigating the impact of MC and improving DOA estimation accuracy. The main contributions are summarized as follows:

- A novel HADS configured with the switch network, amplifier and phase shifter is introduced. With such configurations, the unknown MC can be estimated and compensated in an efficient way. To the best of our knowledge, it is the first time to exploit such a structure to achieve DOA estimation for HADS with unknown MC.
- By jointly applying the middle subarray and SCM reconstruction techniques, a subspace based two-stage DOA estimation

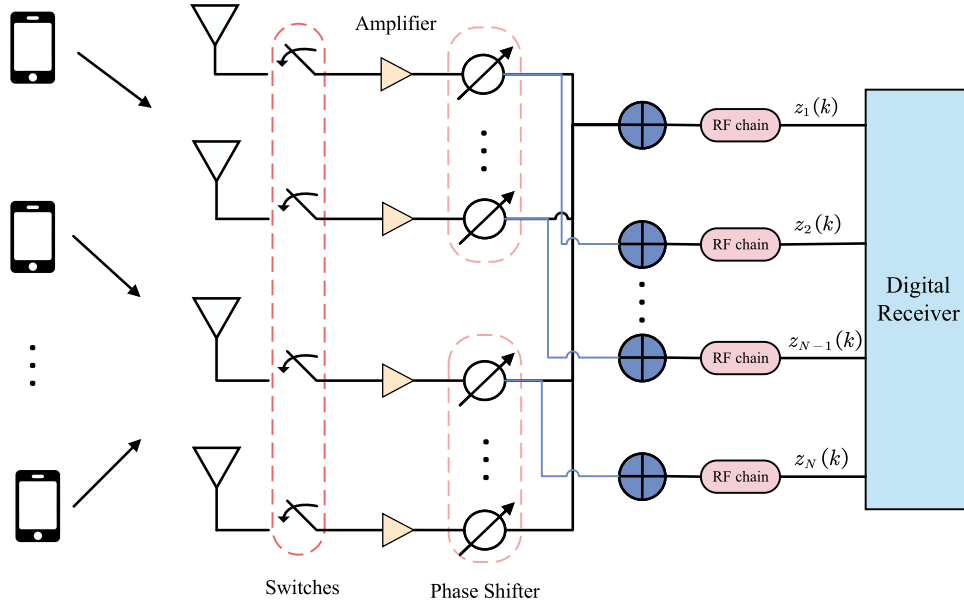


Fig. 1. The adopted HADS for DOA estimation.

method is proposed, which provides a satisfactory estimation performance.

- Extensive simulations under various conditions are performed, showing the superiority of the proposed method for migrating the MC effect and facilitating its practical engineering applications.

*Notations:* Boldface uppercase (lowercase) letters stand for matrices (vectors). The superscripts  $(\cdot)^{-1}$ ,  $(\cdot)^T$ ,  $(\cdot)^H$ , and  $(\cdot)^\dagger$  represent the inverse, transpose, conjugate transpose, and pseudo-inverse operations, respectively.  $\mathbf{0}_{M \times N}$  represents an  $M \times N$  all-zero matrix,  $\mathbf{I}_M$  the  $M \times M$  identity matrix,  $\mathbf{J}_M$  the  $M \times M$  exchange matrix with all ones on its anti-diagonal positions and zeros elsewhere.  $\mathbb{E}\{\cdot\}$ ,  $\text{diag}\{\cdot\}$ ,  $\text{blkdiag}\{\cdot\}$ , and  $\text{vec}\{\cdot\}$  stand for the statistical expectation, diagonalization, block diagonalization, and vectorization operations, respectively.  $\odot$  denotes the Hadamard product, and  $\otimes$  the Kronecker product.  $\text{Toeplitz}\{\cdot\}$  represents a symmetric Toeplitz matrix formed by the bracketed entries. Finally,  $\text{Re}(\cdot)$  returns the real part of the argument,  $\angle[\cdot]$  returns the phase of a complex number,  $|\cdot|$  denotes complex numbers or vectors modulo, and  $\text{unvec}(\mathbf{a})$  rearranges a column-arranged vector  $\mathbf{a}$  to form a matrix,  $(\mathbf{a})_i$  denotes the  $i$ th element in the vector  $\mathbf{a}$ .

## II. SYSTEM MODEL

Considering an HADS consists of  $M$  antennas and  $N$  ( $N < M$ ) RF chains shown in Fig. 1. The  $M$  antennas form a uniform linear array (ULA) with an inter-element spacing of half the carrier wavelength. The first RF chain is configured to connect to all antennas, while the other RF chains are sequentially connected to a single antenna. Each antenna is equipped with an auxiliary unit in the analog domain, which includes a switch, an amplifier, and a phase shifter. Consider  $L$  far-field incoherent narrowband signals  $s_1(t), s_2(t), \dots, s_L(t)$  impinging from  $L$  distinct DOAs  $\theta: \theta_1, \theta_2, \dots, \theta_L$ . To ensure the feasibility of subspace decomposition and the algorithm's validity, the constraint  $M - 2P > L$  must be strictly satisfied. Taking the unknown MC into consideration, the analog array output at time index  $t$  before combination can be expressed as

$$\mathbf{y}(t) = \mathbf{S}\mathbf{W}\mathbf{H}\mathbf{x}(t) = \mathbf{S}\mathbf{W}\mathbf{H}(\mathbf{C}\mathbf{A}(\theta)\mathbf{s}(t) + \mathbf{n}(t)), \quad (1)$$

where  $\mathbf{x}(t) = \mathbf{C}\mathbf{A}(\theta)\mathbf{s}(t) + \mathbf{n}(t)$  represents the output before amplifier,  $\mathbf{S}$  is a diagonal switch matrix, whose  $(m, m)$ -th element is one if the switch corresponding to the  $m$ -th antenna is on and zeros elsewhere,

and  $\mathbf{W} = \text{diag}(\rho_1, \dots, \rho_M)$  and  $\mathbf{H} = \text{diag}(e^{j\varphi_1}, \dots, e^{j\varphi_M})$  are the amplitude adjustment and phase adjustment matrices, respectively, with  $\rho_m$  and  $\varphi_m$  standing for amplifier gain and phase shift at time index  $t$ . Besides,  $\mathbf{s}(t) = [s_1(t), \dots, s_L(t)]^T$  and  $\mathbf{n}(t) = [n_1(t), \dots, n_M(t)]^T$  denote the signal vector and Gaussian white noise vector, respectively. Further,  $\mathbf{A}(\theta) = [\mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_L)]$  denotes the  $M \times L$  analog array manifold matrix with  $\mathbf{a}(\theta_l) = [1, e^{jw_l}, \dots, e^{j(M-1)w_l}]^T$ ,  $w_l = -\pi \sin \theta_l$ , and  $\mathbf{C}$  is an  $M \times M$  MC matrix with a symmetric Toeplitz structure, which is expressed as

$$\mathbf{C} = \text{Toeplitz} \{ [c_0, c_1, \dots, c_P, \mathbf{0}_{1 \times (M-P-1)}] \}, \quad (2)$$

where  $c_i$  stands for the MC coefficient between two antennas,  $0 < |c_P| \dots < |c_1| < c_0 = 1$ , and  $P$  is the length of unknown MC coefficients.

In the case of FDS, the SCM of  $\mathbf{y}(t)$  can be directly utilized for DOA estimation. However, the presence of MC and the limited number of RF chains in the HADS pose challenges in achieving high-resolution and high-accuracy DOA estimation. To tackle this challenge, the following section presents a method to enhance the DOA estimation performance in the considered scenario.

## III. TWO-STAGE DOA ESTIMATION

In this section, we present a novel two-stage approach for DOA estimation. The proposed method first addresses the challenge of reconstructing the SCM by adjusting the states of the auxiliary units and applying the diagonal loading technique to mitigate mutual coupling effects. In the second stage, we leverage a super-resolution subspace method to accurately estimate the DOA by exploiting the reconstructed SCM. This two-step process enhances DOA accuracy by effectively utilizing the limited RF chains in the hybrid analog-digital structure.

### A. First Stage: Initial DOA Estimation

To obtain initial DOA estimation while suppressing the impact of MC, we carefully configure the matrices  $\mathbf{S}$ ,  $\mathbf{W}$ , and  $\mathbf{H}$ .  $\mathbf{S}$  is set as a block diagonal matrix  $\mathbf{S} = \text{blkdiag}\{\mathbf{0}_{P \times P}, \mathbf{I}_{M-2P}, \mathbf{0}_{P \times P}\}$  to select only the middle  $M - 2P$  antennas. The selection of the middle subarray is not arbitrary but mathematically optimal. When mutual coupling is

present, the central  $M - 2P$  elements of the array exhibit a unique property: their MC-affected steering vector elements have identical amplitude gain profiles, differing only in phase offsets. This property allows the resulting spatial covariance matrix to maintain the same structure as would be obtained from an array without mutual coupling (differing only in the signal covariance matrix), which forms the theoretical foundation for subsequent MC suppression and reliable DOA estimation. In contrast, selecting boundary elements would introduce asymmetric coupling effects, resulting in a more complex mathematical structure that would degrade the accuracy of subspace-based DOA estimation.  $\mathbf{W}$  is set as an identity matrix, e.g.,  $\mathbf{W} = \mathbf{I}_M$ , to maintain the focus on phase distortions.  $\mathbf{H}$  is configured as a diagonal matrix  $\mathbf{H} = \text{diag}(1, e^{-j\pi \sin \theta}, \dots, e^{-j(M-1)\pi \sin \theta})$ , where  $\theta$  belongs to the observation space of phase shift, to compensate for the phase distortions caused by MC at different angles. After analog combination and digital sampling of the first RF chain, the output of the digital receiver can be written as

$$z_1(k) = \mathbf{h}_q^H \bar{\mathbf{x}}(k), k = 1, \dots, K, q = 1, \dots, Q, \quad (3)$$

where  $\mathbf{h}_q = [e^{-j\pi \sin \tilde{\theta}_q}, \dots, e^{-j(M-1)\pi \sin \tilde{\theta}_q}]^T$ ,  $\bar{\mathbf{x}}(k)$  represents the middle  $\tilde{M} = M - 2P$  elements of  $\mathbf{x}(k)$ ,  $K$  stands for the number of samples with respect to candidate phase shift  $\tilde{\theta}_q$ , and  $Q$  is the dimension of phase shift observation space at this stage.

Let us further assume that  $\tilde{\theta}_q$  remains unchanged within  $K$  samples. Subsequently, the average power of  $z_1(k)$  can be calculated as

$$\begin{aligned} P_q &= \frac{1}{K} \sum_{k=1}^K |z_1(k)|^2 = \frac{1}{K} \sum_{k=1}^K z_1(k) z_1^H(k) \\ &= \mathbf{h}_q^H \left( \frac{1}{K} \sum_{k=1}^K \bar{\mathbf{x}}(k) \bar{\mathbf{x}}^H(k) \right) \mathbf{h}_q \approx \mathbf{h}_q^H \mathbf{R}_{\bar{\mathbf{x}}} \mathbf{h}_q, \end{aligned} \quad (4)$$

where

$$\begin{aligned} \mathbf{R}_{\bar{\mathbf{x}}} &= \mathbb{E} \{ \bar{\mathbf{x}}(t) \bar{\mathbf{x}}^H(t) \} \\ &= \mathbf{F} \mathbf{C} \mathbf{A}(\theta) \mathbf{R}_s \mathbf{A}^H(\theta) \mathbf{C}^H \mathbf{F}^H + \sigma_n^2 \mathbf{I}_{M-2P} \\ &= \bar{\mathbf{A}}(\theta) \mathbf{\Pi} \mathbf{R}_s \mathbf{\Pi}^H \bar{\mathbf{A}}^H(\theta) + \sigma_n^2 \mathbf{I}_{M-2P} \\ &= \bar{\mathbf{A}}(\theta) \bar{\mathbf{R}}_s \bar{\mathbf{A}}^H(\theta) + \sigma_n^2 \mathbf{I}_{M-2P}. \end{aligned} \quad (5)$$

Here,  $\mathbf{R}_{\bar{\mathbf{x}}}$  denotes the SCM of middle  $\tilde{M}$  analog antennas. Besides,  $\bar{\mathbf{A}}(\theta)$  is the middle  $\tilde{M}$  rows of  $\mathbf{A}(\theta)$ ,  $\mathbf{F} = [\mathbf{0}_{(M-2P) \times P} \quad \mathbf{I}_{M-2P} \quad \mathbf{0}_{(M-2P) \times P}]$ , and  $\bar{\mathbf{R}}_s = \mathbf{\Pi} \mathbf{R}_s \mathbf{\Pi}^H$ ,  $\mathbf{\Pi} = \text{diag}\{\sum_{p=-P}^P c_{|p|} e^{j p w_1}, \dots, \sum_{p=-P}^P c_{|p|} e^{j p w_L}\}$ ,  $\mathbf{R}_s = \mathbb{E}\{\mathbf{s}(t) \mathbf{s}^H(t)\} = \text{diag}(\sigma_{s1}^2, \dots, \sigma_{sL}^2)$  stands for the covariance matrix of source signals, and  $\sigma_n^2$  the noise variance [22].

With the aid of operation  $\text{vec}(\mathbf{EFG}) = (\mathbf{G}^T \otimes \mathbf{E}) \text{vec}(\mathbf{F})$  and the definition of  $\mathbf{r} = \text{vec}(\mathbf{R}_{\bar{\mathbf{x}}})$ ,  $P_q$  can be further derived as

$$P_q = \text{vec}(\mathbf{h}_q^H \mathbf{R}_{\bar{\mathbf{x}}} \mathbf{h}_q) = (\mathbf{h}_q \otimes \mathbf{h}_q^*)^T \mathbf{r}. \quad (6)$$

By continuously adjusting the phase shift with different  $\tilde{\theta}_q$ ,  $q = 1, \dots, Q$ , an overdetermined system is constructed as

$$\mathbf{p} = [P_1, \dots, P_Q]^T = \mathbf{B} \mathbf{r} = \begin{bmatrix} (\mathbf{h}_1 \otimes \mathbf{h}_1^*)^T \\ \vdots \\ (\mathbf{h}_Q \otimes \mathbf{h}_Q^*)^T \end{bmatrix} \mathbf{r}. \quad (7)$$

Given the fact that  $\mathbf{B}$  may be rank-deficient, directly solving (7) could lead to ill-posed solutions. To address this issue, we apply the diagonal

loading technique [23], which improves the numerical stability of the inverse. Thus, the solution is given by

$$\hat{\mathbf{r}} = (\mathbf{B}^H \mathbf{B} + \varepsilon \mathbf{I}_{(M-2P)^2})^{-1} \mathbf{B}^H \mathbf{p}, \quad (8)$$

where  $\varepsilon$  represents the diagonal loading coefficient, and  $\mathbf{I}_{(M-2P)^2}$  is the identity matrix of size  $(M - 2P)^2$ . This technique ensures a well-conditioned system, leading to a more stable estimation of the SCM. The reconstructed SCM corresponding to the middle  $\tilde{M}$  antennas is then obtained by

$$\hat{\mathbf{R}}_{\bar{\mathbf{x}}} = (\text{unvec}(\hat{\mathbf{r}}) + (\text{unvec}(\hat{\mathbf{r}}))^H) / 2, \quad (9)$$

where the operator  $\text{unvec}(\hat{\mathbf{r}})$  denotes the inverse vectorization process, transforming  $\hat{\mathbf{r}}$  back into the matrix form of the estimated covariance.

It is noted that complex multiplication costs four times as much as real multiplication. *Therefore, to reduce system overhead while maintaining estimation accuracy, we apply the real-valued subspace technique from [21] for initial DOA estimation.* First, we define the Hermitian persymmetric matrix  $\tilde{\Psi}$  and the unitary transformation matrix  $\mathbf{U}_{\tilde{M}}$  as follows:

$$\tilde{\Psi} = \frac{\bar{\mathbf{A}} (\bar{\mathbf{R}}_s + \Phi^{1-\tilde{M}} \bar{\mathbf{R}}_s^* \Phi^{\tilde{M}-1}) \bar{\mathbf{A}}^H}{2} + \sigma_n^2 \mathbf{I}_{\tilde{M}}, \quad (10)$$

$$\mathbf{U}_{\tilde{M}} = \begin{cases} \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbf{I}_{\frac{\tilde{M}}{2}} & \mathbf{J}_{\frac{\tilde{M}}{2}} \\ j\mathbf{J}_{\frac{\tilde{M}}{2}} & -j\mathbf{I}_{\frac{\tilde{M}}{2}} \end{pmatrix}, & \tilde{M} \text{ is even} \\ \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbf{I}_{\frac{\tilde{M}-1}{2}} & \mathbf{0}_{\frac{\tilde{M}-1}{2} \times 1} & \mathbf{J}_{\frac{\tilde{M}-1}{2}} \\ \mathbf{0}_{1 \times \frac{\tilde{M}-1}{2}} & \sqrt{2} & \mathbf{0}_{1 \times \frac{\tilde{M}-1}{2}} \\ j\mathbf{J}_{\frac{\tilde{M}-1}{2}} & \mathbf{0}_{\frac{\tilde{M}-1}{2} \times 1} & -j\mathbf{I}_{\frac{\tilde{M}-1}{2}} \end{pmatrix}, & \tilde{M} \text{ is odd} \end{cases} \quad (11)$$

where  $\Phi = \text{diag}\{e^{jw_1}, \dots, e^{jw_L}\}$ . Accordingly, we can construct the real-valued SCM as [20]

$$\begin{aligned} \tilde{\Psi} &= \mathbf{U}_{\tilde{M}} \tilde{\Psi} \mathbf{U}_{\tilde{M}}^H \\ &= \frac{\tilde{\mathbf{A}}(\theta) \cdot \text{Re}(\Phi^{\frac{\tilde{M}-1}{2}} \bar{\mathbf{R}}_s \Phi^{\frac{1-\tilde{M}}{2}}) \tilde{\mathbf{A}}^H(\theta)^T}{2} + \sigma_n^2 \mathbf{I}_{\tilde{M}}, \end{aligned} \quad (12)$$

where  $\tilde{\mathbf{A}}(\theta) = [\tilde{\mathbf{a}}(\theta_1), \dots, \tilde{\mathbf{a}}(\theta_L)]$ , with its  $l$ -th column being represented as

$$\tilde{\mathbf{a}}(\theta_l) = \begin{cases} \sqrt{2} \left[ \cos\left(\frac{\tilde{M}-1}{2} w_l\right), \dots, \cos\left(\frac{1}{2} w_l\right), \right. \\ \left. \sin\left(\frac{1}{2} w_l\right), \dots, \sin\left(\frac{\tilde{M}-1}{2} w_l\right) \right]^T, & \tilde{M} \text{ is even} \\ \sqrt{2} \left[ \cos\left(\frac{\tilde{M}-1}{2} w_l\right), \dots, \cos(w_l), 1, \right. \\ \left. \sin(w_l), \dots, \sin\left(\frac{\tilde{M}-1}{2} w_l\right) \right]^T, & \tilde{M} \text{ is odd.} \end{cases} \quad (13)$$

Performing eigenvalue decomposition (EVD) on  $\tilde{\Psi}$  to obtain  $\tilde{M} \times (\tilde{M} - L)$  noise subspace matrix  $\mathbf{E}_n$  corresponding to the  $\tilde{M} - L$  largest eigenvalues, the initial DOA estimations can be obtained by searching  $L$  peaks of the following real-valued function

$$S(\theta) = \frac{1}{\tilde{\mathbf{a}}^H(\theta) \mathbf{E}_n \mathbf{E}_n^H \tilde{\mathbf{a}}(\theta)}. \quad (14)$$

## B. Second Stage: Enhanced DOA Estimation

It should be noted that the first-stage DOA estimation does not fully utilize the SCM of entire array. To tackle this issue, we reset  $\mathbf{S} = \mathbf{W} =$

$\mathbf{H} = \mathbf{I}_M$ . Since the first RF chain is occupied during the initial stage for analog combination and sampling, the outputs of the second to the  $N$ -th RF chains are leveraged in the second stage. Consequently, the outputs of the second to the  $N$ -th RF chains are expressed as

$$\mathbf{z}(k) = [z_2(k), \dots, z_N(k)]^T = \check{\mathbf{C}}\check{\mathbf{A}}(\theta)\mathbf{s}(t) + \check{\mathbf{n}}(t), \quad (15)$$

where  $\check{\mathbf{C}}, \check{\mathbf{A}}(\theta) = [\check{\mathbf{a}}(\theta_1), \dots, \check{\mathbf{a}}(\theta_L)]$  and  $\check{\mathbf{n}}(t)$  represent the submatrices that correspond to the first  $N - 1$  antennas, respectively.

For estimating and compensating MC, and subsequently making full use of the outputs of the entire array, we rewrite  $\check{\mathbf{C}}\check{\mathbf{A}}(\theta_l)$  as

$$\check{\mathbf{C}}\check{\mathbf{A}}(\theta_l) = \mathbf{T}_l \mathbf{c}, \quad (16)$$

where  $\mathbf{c} = [1, c_1, \dots, c_P]^T$  and

$$\begin{aligned} \mathbf{T}_l &= \mathbf{T}_{l,1} + \mathbf{T}_{l,2}, \\ \mathbf{T}_{l,1}(i, j) &= \begin{cases} [\check{\mathbf{a}}(\theta_l)]_{i+j-1}, & i + j \leq N \\ 0, & \text{otherwise,} \end{cases} \\ \mathbf{T}_{l,2}(i, j) &= \begin{cases} [\check{\mathbf{a}}(\theta_l)]_{i-j+1}, & i \geq j \geq 2 \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \quad (17)$$

Let  $\check{\mathbf{U}}_n$  denote the noise subspace matrix with respect to the SCM of  $\mathbf{z}(k)$ . Since the actual steering vector is orthogonal to the noise subspace  $\check{\mathbf{U}}_n^H \check{\mathbf{C}}\check{\mathbf{A}}(\theta_l) = \mathbf{0}_{(N-1-L) \times 1}$ , applying (16) yields  $\check{\mathbf{U}}_n^H \mathbf{T}[\check{\mathbf{a}}(\theta_l)]\mathbf{c} = \mathbf{0}_{(N-1-L) \times 1}$ . Then, by inserting initial DOA estimates  $\{\hat{\theta}_l\}_{l=1}^L$ , we can construct another overdetermined system as

$$\begin{bmatrix} \check{\mathbf{U}}_n^H \mathbf{T}[\check{\mathbf{a}}(\theta_1)] \\ \vdots \\ \check{\mathbf{U}}_n^H \mathbf{T}[\check{\mathbf{a}}(\theta_L)] \end{bmatrix} \mathbf{c} = \Xi_T \mathbf{c} = \mathbf{0}_{L(N-1-L) \times 1}, \quad (18)$$

which directly yields that

$$[\hat{c}_1, \dots, \hat{c}_P] = -\Xi_{T2}^\dagger \Xi_{T1}, \quad (19)$$

where  $\Xi_{T1}$  and  $\Xi_{T2}$  stand for the first row and remaining rows of  $\Xi_T$ , respectively.

With available  $\{\hat{\theta}_l\}_{l=1}^L$  and MC coefficients  $\{\hat{c}_i\}_{i=1}^P$ , we can reconstruct  $\mathbf{C}$  as

$$\hat{\mathbf{C}} = \text{Toeplitz} \{ [1, \hat{c}_1, \dots, \hat{c}_P, \mathbf{0}_{1 \times (M-P-1)}] \}. \quad (20)$$

In this refinement step of the second stage, we further configure  $\mathbf{S}$ ,  $\mathbf{W}$  and  $\mathbf{H}$  with  $\mathbf{S} = \mathbf{I}_M$ ,  $\mathbf{W} = \text{diag}\{ |(\hat{\mathbf{C}}\hat{\mathbf{a}}(\hat{\theta}))_1|, |(\hat{\mathbf{C}}\hat{\mathbf{a}}(\hat{\theta}))_2|, \dots, |(\hat{\mathbf{C}}\hat{\mathbf{a}}(\hat{\theta}))_M| \}$  and  $\mathbf{H} = \text{diag}\{\exp\{-\angle \hat{\mathbf{C}}\hat{\mathbf{a}}(\hat{\theta})\}\}$ . This configuration refines the parameters of the second stage to better utilize the reconstructed coupling matrix and achieve improved estimation accuracy. Hence, the average power of  $z_1(k)$  can be calculated by

$$P_q = \mathbf{h}_q^H \left( \frac{1}{K} \sum_{k=1}^K \mathbf{x}(k) \mathbf{x}^H(k) \right) \mathbf{h}_q \approx \mathbf{h}_q^H \mathbf{R}_x \mathbf{h}_q, \quad (21)$$

where  $\mathbf{R}_x$  is the  $M \times M$  SCM of the whole antenna array, and  $\mathbf{h}_q = \hat{\mathbf{C}}\hat{\mathbf{a}}(\hat{\theta}_q)$ . Taking the same procedure with (7) to reconstruct full-dimensional SCM  $\hat{\mathbf{R}}_x$ . The enhanced DOA estimation are obtained by searching  $L$  peaks of the following function

$$S'(\theta) = \frac{1}{\mathbf{a}^H(\theta) \hat{\mathbf{C}}^H \Sigma_n \Sigma_n^H \hat{\mathbf{C}} \mathbf{a}(\theta)}, \quad (22)$$

where  $\Sigma_n$  is the  $M - L$  noise subspace matrix of  $\hat{\mathbf{R}}_x$ .

#### IV. CRB ANALYSIS

The derivation process in the previous section shows that the proposed method sacrifices the number of samples in favor of enabling large-scale SCM reconstruction. To justify this approach, we present the following proposition, which demonstrates why SCM reconstruction is a more advantageous choice in this context.

*Proposition 1:* The SCM reconstruction based approach in second stage yields better estimation accuracy compared to directly using the SCM of the RF chain outputs, provided that the number of antennas  $M$  is far larger than that of the RF chains  $N$ .

*Proof:* We employ the Cramér-Rao Bound (CRB) to verify this proposition. The CRB itself is independent of the estimation method; rather, it depends on the Fisher information contained in the measurements. The proposed approach, by reconstructing the SCM, effectively increases the Fisher information, resulting in a lower CRB. For simplicity, one source signal with DOA  $\theta$  is considered, whose closed-form CRB expression with given number of antennas  $\check{M}$  and samples  $\check{K}$  is written as [22]

$$\text{CRB} = \frac{1}{\check{K}} \cdot \frac{\check{M}\kappa + 1}{\check{M}^2(\check{M}^2 - 1)} \cdot \frac{6}{\kappa^2 \pi^2 \cos^2 \theta}, \quad (23)$$

where  $\kappa = \sigma_s^2 / \sigma_n^2$  stands for the SNR. For a large-scale array and a relatively high SNR, the CRB in (23) reduces to

$$\text{CRB} \approx \frac{1}{\check{K}\check{M}^3} \cdot \frac{6}{\kappa \pi^2 \cos^2 \theta} \propto \frac{1}{\check{K}\check{M}^3}, \quad (24)$$

while for a large-scale array and sufficiently low SNRs (i.e.,  $\kappa \ll 1/\check{M}$ ), the CRB in (23) reduces to

$$\text{CRB} \approx \frac{1}{\check{K}\check{M}^4} \cdot \frac{6}{\kappa^2 \pi^2 \cos^2 \theta} \propto \frac{1}{\check{K}\check{M}^4}. \quad (25)$$

It is emphasized that although there are  $(Q + \bar{Q} + 1)K$  samples for DOA estimation of the proposed method, where  $\bar{Q}$  is the dimension of phase shift observation space at second stage, the final  $M \times M$  SCM has the same form with that obtained by  $M$ -element FDS array with  $K$  samples. For comparison, if one applies all samples for DOA estimation utilizing the output data of  $N$  RF chains directly, the SCM can be regarded as the one obtained by  $N$ -element FDS array with  $(Q + \bar{Q} + 1)K$  samples. Meanwhile, since the candidate phase shifts at the stage are chosen around a very small range of the initial DOA estimates, we have  $\bar{Q} \ll Q$ . Consequently, it can be obtained that the SCM reconstruction based approach can yield higher estimation accuracy in comparison with using the SCM of RF chains output directly, provided that

$$M > \begin{cases} \sqrt[3]{Q + \bar{Q} + 1}N \approx \sqrt[3]{Q}N, & \text{high } \kappa \\ \sqrt[4]{Q + \bar{Q} + 1}N \approx \sqrt[4]{Q}N, & \text{small } \kappa. \end{cases} \quad (26)$$

This ends the proof of proposition 1.  $\square$

It is worth noting that while CRB provides a useful local performance bound for our proposed method in high-SNR regimes, alternative global bounds such as the Ziv-Zakai bound (ZZB) [24] may offer more comprehensive characterization of estimator performance across all SNR regimes, particularly in capturing threshold effects at low SNR. However, for the high-SNR applications targeted in this work, the CRB provides sufficient insights into the performance limits of our algorithm with more tractable analytical expressions.

#### V. SIMULATION RESULTS

In this section, the performance of the proposed method is assessed. For convenience of comparison, we designate Proposed-FS



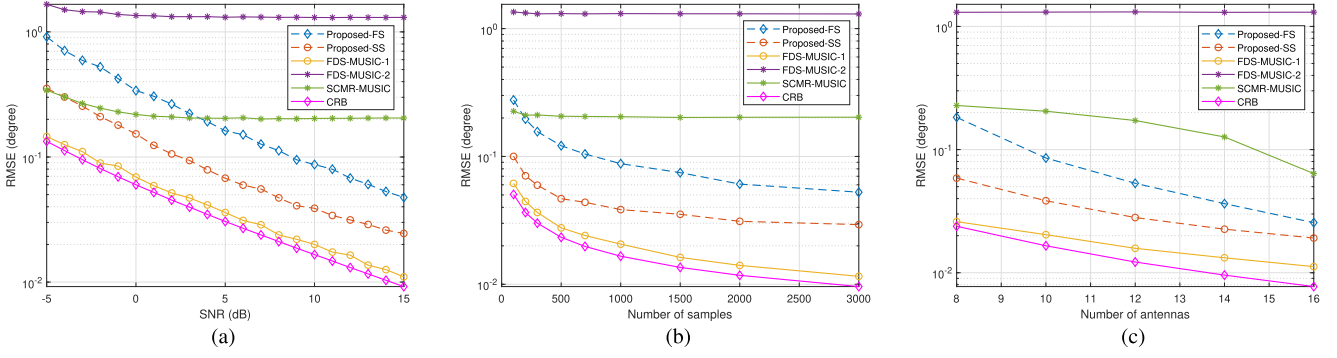


Fig. 2. RMSE of DOA estimations under various simulation configurations, with  $\theta_1 = -32.7^\circ$  and  $\theta_2 = 24.1^\circ$ : (a) RMSE versus SNR; (b) RMSE versus the number of samples; (c) RMSE versus the number of antennas.

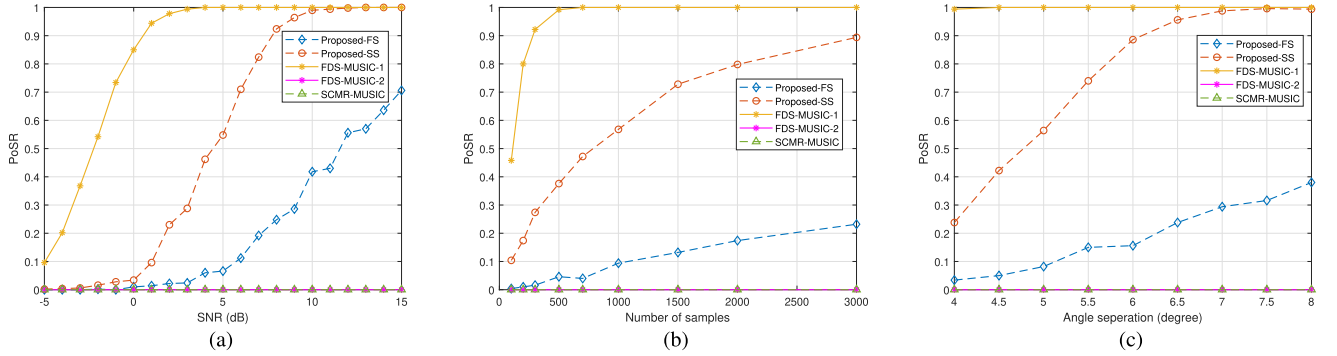


Fig. 3. PoSR of DOA estimations under various simulation configurations: (a) PoSR versus SNR; (b) PoSR versus the number of samples; (c) PoSR versus angle separation.

and Proposed-SS as the proposed techniques at the first and second stages, respectively. The FDS based MUSIC method with known MC utilizing  $M$  sensors (named as FDS-MUSIC-1) and  $N$  sensors (named as FDS-MUSIC2, where  $N$  equals the number of RF chains of the adopted HDAS), the SCM reconstruction based MUSIC without mitigating the impact of MC (named as SCMR-MUSIC), and CRB are selected for comparison. In the simulations, the length of unknown MC coefficients is set to  $P = 2$  with  $c_1 = 0.4045 + 0.2939j$  and  $c_2 = 0.2898 - 0.0776j$ , the number of phase shifts at the first stage and second stage are set to  $Q = 181$  and  $\bar{Q} = 30$ , respectively. The diagonal loading coefficient is set to  $\varepsilon = 1$ . The root mean square error (RMSE) and probability of successful resolution (PoSR) obtained by 500 independent Monte-Carlo trials are used to evaluate the performance of the proposed method.

The first simulation compares RMSEs of DOA estimation obtained by different methods under different SNRs, number of samples  $K$  and antennas  $M$ , where two sources located at  $\theta_1 = -32.7^\circ$  and  $\theta_2 = 24.1^\circ$  are considered. The simulation results are shown in Fig. 2, and the corresponding simulation configurations are set as follows: in Fig. 2(a),  $K = 1000$ ,  $M = 10$ ,  $N = 5$ , and SNR changes from  $-5$  dB to  $15$  dB; in Fig. 2(b),  $M = 10$ ,  $N = 5$ , SNR=10 dB and  $K$  varies from 100 to 3000; in Fig. 2(c),  $K = 1000$ ,  $N = 5$ , SNR=10 dB, and  $M$  varies from 8 to 20. The results demonstrate that the proposed methods can mitigate the impact of MC effectively, verified by comparison of the performance between our solutions and SCMR-MUSIC. Meanwhile, the performance of Proposed-SS exceeds that of Proposed-FS, thereby validating the effectiveness of the dual-stage estimation strategy. Furthermore, the proposed solution leverages the aperture advantage of the entire array, achieving significantly better performance compared to FDS-MUSIC-2. Notably, when  $M$  is relatively large, the proposed

method approaches the performance of FDS-MUSIC-1, highlighting its robustness and efficiency under practical conditions.

The second simulation evaluates the PoSRs of DOA estimation with respect to SNR,  $K$  and angle separation. Two DOAs are considered to be successfully resolved, provided that the absolute biases of two DOA estimates are all less than  $0.5^\circ$ . In Fig. 3(a) and (b), the simulation conditions are respectively the same with those in Fig. 3(a) and 3(b), respectively, except that two closely spaced sources are located at  $\theta_1 = -32.7^\circ$  and  $\theta_2 = -27.7^\circ$  with  $5^\circ$  separation, and SNR=5dB in Fig. 3(b). In Fig. 3(c),  $K$ ,  $M$ ,  $N$  and SNR are set to 1000, 10, 5 and 5 dB, respectively. The first DOA is fixed at  $\theta_1 = -32.7^\circ$ , while the second DOA  $\theta_2$  varies from  $-28.7^\circ$  to  $-24.7^\circ$ . From the simulation results in Fig. 3, it is evident that the PoSR of the proposed method improves with increasing SNR,  $K$ , and angle separation. Notably, Proposed-SS demonstrates clear super-resolution capability, which can be attributed to its ability to fully leverage the aperture of the entire array. In contrast, FDS-MUSIC-2 fails to resolve the two sources across the entire observed region due to its smaller effective array aperture.

## VI. CONCLUSION

A novel DOA estimation method for HADS in the presence of unknown MC has been presented in this correspondence. By appropriately adjusting the state of switch network, amplifier and phase shifters, a two-stage estimation approach is described in detail, where the middle subarray, diagonal loading based covariance matrix reconstruction, high-resolution and/or real-valued subspace techniques are jointly exploited. As demonstrated by simulations, the proposed method can provide good estimation accuracy super-resolution performance.

## REFERENCES

- [1] Y. Tian, W. Liu, H. Xu, S. Liu, and Z. Dong, "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.
- [2] T. Wu et al., "Joint angle estimation error analysis and 3-D positioning algorithm design for mmWave positioning system," *IEEE Int. Things J.*, vol. 11, no. 2, pp. 2181–2197, Jan. 2024.
- [3] K. Xu, X. Xia, C. Li, G. Hu, Q. Su, and W. Xie, "Spatial structure matching-based DOA estimation and tracking for integrated sensing and communication massive MIMO OFDM system," *IEEE Trans. Cogn. Commun. Netw.*, vol. 10, no. 2, pp. 526–540, Apr. 2024.
- [4] H. Xu, W. Liu, M. Jin, and Y. Tian, "Positioning and contour extraction of autonomous vehicles based on enhanced DOA estimation by large-scale arrays," *IEEE Int. Things J.*, vol. 10, no. 13, pp. 11792–11803, Jul. 2023.
- [5] Y. Tian, X. Gao, W. Liu, H. Chen, G. Wang, and Y. Qin, "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.
- [6] Y. Mao, D. Gao, Q. Guo, and M. Jin, "Message passing based block sparse signal recovery for DOA estimation using large arrays," *IEEE Signal Process. Lett.*, vol. 30, pp. 1552–1556, 2023.
- [7] L. Han, X. Liu, N. Zhang, S. Wu, J. Zhu, and Z. Xu, "Two-dimensional multi-snapshot newtonized orthogonal matching pursuit for DOA estimation," *Digit. Signal Process.*, vol. 121, Mar. 2022, Art. no. 103313.
- [8] J. Yu and Y. Wang, "Deep learning-based multipath DoAs estimation method for mmWave massive MIMO systems in low SNR," *IEEE Trans. Veh. Technol.*, vol. 72, no. 6, pp. 7480–7490, Jun. 2023.
- [9] T. Wu et al., "Exploit high-dimensional RIS information to localization: What is the impact of faulty element?," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 10, pp. 2803–2819, Oct. 2024.
- [10] H. Huang, J. Yang, H. Huang, Y. Song, and G. Gui, "Deep learning for super-resolution channel estimation and DOA estimation based massive MIMO system," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8549–8560, Sep. 2018.
- [11] X. Meng, B. X. Cao, F. G. Yan, M. Greco, F. Gini, and Y. Zhang, "Real-valued MUSIC for efficient direction of arrival estimation with arbitrary arrays: Mirror suppression and resolution improvement," *Signal Process.*, vol. 202, 2023, Art. no. 108766.
- [12] Y. Chen, L. Yan, C. Han, and M. Tao, "Millidegree-level direction-of-arrival estimation and tracking for terahertz ultra-massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 869–883, Feb. 2022.
- [13] F. Shu et al., "Low-complexity and high-resolution DOA estimation for hybrid analog and digital massive MIMO receive array," *IEEE Trans. Commun.*, vol. 66, no. 6, pp. 2487–2501, Jun. 2018.
- [14] S. Li, Y. Liu, L. You, W. Wang, H. Duan, and X. Li, "Covariance matrix reconstruction for DOA estimation in hybrid massive MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 9, no. 8, pp. 1196–1200, Aug. 2020.
- [15] Y. Liu, Y. Yan, L. You, W. Wang, and H. Duan, "Spatial covariance matrix reconstruction for DOA estimation in hybrid massive MIMO systems with multiple radio frequency chains," *IEEE Trans. Veh. Technol.*, vol. 70, no. 11, pp. 12185–12190, Nov. 2021.
- [16] Y. Zhou, B. Dang, Y. Li, and G. Liu, "An efficient spatial covariance matrix reconstruction algorithm in the hybrid analog-digital structure," *IEEE Trans. Veh. Technol.*, vol. 71, no. 7, pp. 7930–7935, Jul. 2022.
- [17] Y. Zhou, G. Liu, J. Li, Y. Li, S. Ye, and L. Li, "A high-efficiency beam sweeping algorithm for DOA estimation in the hybrid analog-digital structure," *IEEE Wireless Commun. Lett.*, vol. 10, no. 10, pp. 2323–2327, Oct. 2021.
- [18] C. Liu, Y. Zhou, and B. Dang, "A low-computational-cost alternating switches algorithm for real-valued spatial covariance matrix reconstruction," *Electron. Lett.*, vol. 58, no. 14, pp. 557–559, 2022.
- [19] M. D. Zoltowski, G. M. Kautz, and S. D. Silverstein, "Beamspace Root-MUSIC," *IEEE Trans. Signal Process.*, vol. 41, no. 1, pp. 344–364, Jan. 1993.
- [20] W. Hu and Q. Wang, "DOA estimation for UCA in the presence of mutual coupling via error model equivalence," *IEEE Wireless Commun. Lett.*, vol. 9, no. 1, pp. 121–124, Jan. 2020.
- [21] Y. Tian, R. Wang, H. Chen, Y. Qin, and M. Jin, "Real-valued DOA estimation utilizing enhanced covariance matrix with unknown mutual coupling," *IEEE Commun. Lett.*, vol. 26, no. 4, pp. 912–916, Apr. 2022.
- [22] J. Li, P. Stoica, and Z. Wang, "On robust capon beamforming and diagonal loading," *IEEE Trans. Signal Process.*, vol. 51, no. 7, pp. 1702–1715, Jul. 2003.
- [23] B. Friedlander and A. Weiss, "Direction finding in the presence of mutual coupling," *IEEE Trans. Antennas Propag.*, vol. 39, no. 3, pp. 273–284, Mar. 1991.
- [24] Z. Zhang, Z. Shi, and Y. Gu, "Ziv-Zakai bound for DOAs estimation," *IEEE Trans. Signal Process.*, vol. 71, pp. 136–149, 2023.