Two-Stage Beamspace MUSIC-Based Near-Field Channel Estimation for Hybrid XL-MIMO

Kaiqian Qu, Shuaishuai Guo, *Senior Member, IEEE*, Jia Ye, *Member, IEEE*, and Hui Zhao

***Abstract*—In extremely large-scale multiple-input multiple-output (XL-MIMO) systems, channel estimation poses a key challenge due to the introduction of the unknown distance parameter in near-field scenarios. We propose a beamforming codebook that includes pre-compensated distances, which allows the application of the traditional beamspace multiple signal classification (BMUSIC) to near-field channel estimation. To determine the optimal pre-compensation distance, we introduce three strategies: Maximizing the correlation integral (MCI), maximizing the minimum correlation (MMC), and exceeding the minimum correlation threshold (EMCT). In addition, we develop a two-stage BMUSIC algorithm and a switch transformation design to further reduce the time-intensive 2-dimensional (2D) search processes and avoid the overlaps of multiple coherent paths. Simulation results confirm that the proposed method not only diminishes computational complexity but also notably outperforms existing methods in terms of estimation accuracy.**

***Index Terms*—Extremely large-scale MIMO (XL-MIMO), near-field, channel estimation, two-stage bemspace MUSIC.**

I. Introduction

EXTREMELY large-scale multiple-input multiple-output (XL-MIMO) is a promising technology to achieve the exceptional performance expected from 6G [1]. Equipped with a substantial number of antennas, XL-MIMO can achieve unprecedented array gain via beamforming [2]. The extensive array size introduces characteristics of near-field spherical waves, which alter the channel structure. To fully exploit XL-MIMO, acquiring accurate near-field channel state information (CSI) is essential, representing a challenge for research.

Recent studies have focused on near-field channel estimation [3]–[5] and beam training [6]–[8]. In [3], Cui et al. introduced a near-field polar-domain codebook along with a polar-domain orthogonal matching pursuit (P-OMP) algorithm for estimating the near-field channel. Similarly, Lu et al. [4] depicted the XL-MIMO channel employing a mixed line-of-sight (LoS) and non-LoS (NLoS) path components model, utilizing the P-OMP algorithm for NLoS component estimation. With the polar codebook, a look-ahead P-OMP is applied to XL- reconfigurable intelligent surface (RIS) aided MIMO channel estimation [5]. However, the inherent column coherence of the codebook limits the estimation accuracy of P-OMP. Besides, near-field beam training is also a conventional way to obtain CSI. To reduce the training overhead based on an exhaustive search with the two-dimensional (2D) polar-domain codebook, Zhang et al. developed a two-phase beam-training strategy that decouples the angular and distance dimensions [6]. Nonetheless, the scheme in [6] suffers from a scarcity of distance sampling points, leading to inadequate resolution and lower estimation accuracy. Additionally, increasing the number of distance sampling points significantly raises the training overhead. Subsequent research [7], [8] further proposed hierarchical codebook designs to reduce the training overhead, albeit at the expense of performance.

Unlike the above approaches, 2D multiple signal classification (MUSIC) algorithm stands out as a robust method for near-field channel estimation [9]. It remains unaffected by the coherence of distance domain codewords and does not require extensive pilot overhead. When integrated with the XL-array, 2D MUSIC provides ultra-high resolution, precisely estimating both the distance and angle of near-field paths. Nevertheless, its main drawbacks include the time-intensive 2D search and the complexity of Eigenvalue Decomposition (EVD). Although Zhang et.al. proposed a reduced-dimension near-field MUSIC approach, it is only applicable to full-digital systems where the antenna spacing is less than $\frac{1}{4}$ wavelength [10]. This approach is ill-suited for the hybrid precoding structures in practical XL-MIMO systems, which are designed to reduce radio frequency (RF) chains and thereby reducing power consumption and hardware complexity. To the best of our knowledge, there is no near-field MUSIC algorithm that avoids 2D search and also suit hybrid precoding structures.

Considering the above, this letter considers an XL-MIMO system with hybrid precoding and focuses on the near-field channel estimation. We aim to use the beamspace MUSIC (BMUSIC) technique [11] for tackling channel estimation. We introduce the pre-compensation distance to modify the DFT codebook to enhance the applicability of BMUSIC in the near field. The optimal pre-compensation distance is determined by three strategies: maximizing the correlation integral (MCI), maximizing the minimum correlation (MMC), and exceeding the minimum correlation threshold (EMCT). To address the high complexity of the 2D search, we introduce a two-stage BMUSIC strategy, effectively simplifying it into manageable 1D search problems. Furthermore, considering the signal coherence in multipath environments, we draw inspiration from spatial smoothing techniques to implement antenna subarray

The work is supported in part by the National Natural Science Foundation of China under Grant 62171262 and Grant 62301090; in part by Shandong Provincial Natural Science Foundation under Grant ZR2021YQ47, ZR2021LZHP003; in part by the Taishan Young Scholar under Grant tsqn201909043; in part by Major Scientific and Technological Innovation Project of Shandong Province under Grant 2020CXGC010109.

K. Qu and S. Guo are with School of Control Science and Engineering, Shandong University, Jinan 250061, China, and also with Shandong Key Laboratory of Wireless Communication Technologies, Shandong University, China (e-mail: qukaiqian@mail.sdu.edu.cn,shuaishuai_guo@sdu.edu.cn).

Jia Ye is with the School of Electrical Engineering, Chongqing University, Chongqing, 400044, China (jia.ye@cqu.edu.cn).

Hui Zhao is with the Communication Systems Department, EURECOM, 06410 Sophia Antipolis, France (e-mail: hui.zhao@eurecom.fr).

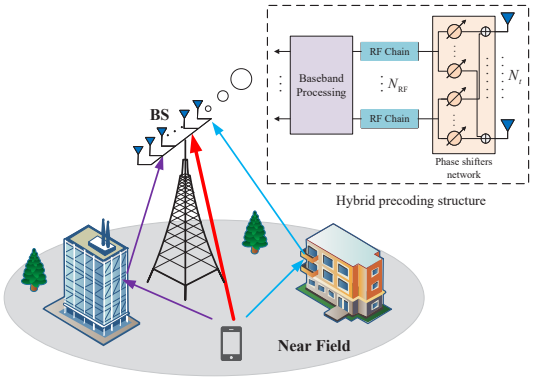


Fig. 1. Illustration of an XL-MIMO communication system, where a user is located in the near-field region.

selection via a switch transformation matrix. Simulation results demonstrate that our proposed method can improve estimation performance and reduce computational complexity.

II. System Model and BMUSIC Method

Consider a near-field XL-MIMO communication system shown in Fig. 1. The base station (BS) is equipped with N_t-antenna uniform linear array (ULA) to communicate with a single-antenna user. The BS employs the hybrid precoding structure, where N_t antennas are connected to N_{RF} RF chains with $N_{\mathrm{RF}} \ll N_t$. The channel estimation is performed by BS using the uplink pilot transmission from the user, with T time slots allocated for the uplink training process. Denoting the pilot signal at t-th time slot as $x(t)$, such that $|x(t)| = 1$, the signal output of the RF chains can be expressed as

$$\mathbf{y}(t) = \mathbf{W}^H \mathbf{h} x(t) + \mathbf{W}^H \mathbf{n}(t), \quad t = 1, 2, \cdots, T, \tag{1}$$

where $\mathbf{W} \in \mathbb{C}^{N_t \times N_{\mathrm{RF}}}$ represents the adopted analog combining matrix at BS; each element $W_{i,j}$ satisfies the constant modulus constraint $|W_{i,j}| = \frac{1}{\sqrt{N_t}}$; $\mathbf{n}(t)$ denotes the complex Gaussian white noise with zero mean and covariance $\sigma^2\mathbf{I}$; and $\mathbf{h}$ is the considered near-field channel modeled as [3]

$$\mathbf{h} = \sqrt{\frac{1}{L}} \sum_{l=1}^{L} \alpha_l \mathbf{a}(\theta_l, r_l), \tag{2}$$

where L denotes the number of paths; α_l, θ_l and r_l represent the gain, the direction and the distance of the l-th path to the BS, respectively; $\mathbf{a}(\theta, r) \in \mathbb{C}^{N_t}$ is the near-field steering vector at the direction θ and distance r defined as

$$\mathbf{a}(\theta, r) = \left[e^{-j2\pi \frac{d_1 - r}{\lambda}}, e^{-j2\pi \frac{d_2 - r}{\lambda}}, \ldots, e^{-j2\pi \frac{d_{N_t} - r}{\lambda}} \right]^T, \tag{3}$$

where $d_n = \sqrt{r^2 + (\delta_n \Delta)^2 - 2r\delta_n \Delta \sin\theta}$ and $\delta_n = \frac{2n - N_t + 1}{2}$ for $n = 0, 1, \ldots, N_t - 1$. Here, $\Delta = \frac{\lambda}{2}$ denotes the antenna spacing, and λ is the carrier wavelength.

In the following, we recall the conventional BMUSIC method for near-field communication, which delivers useful insights for the beamforming and channel estimation design. Following the MUSIC approach, the received pilot signals are collected to evaluate the covariance matrix, given as [11],

$$\mathbf{R}_y \triangleq \mathbb{E}\left\{\mathbf{y}(t)\mathbf{y}(t)^H\right\} \approx \frac{1}{T} \sum_{t=1}^{T} \mathbf{y}(t)\mathbf{y}(t)^H. \tag{4}$$

By stacking all the direction and distance information pairs as $\mathbf{A} = [\mathbf{a}(\theta_1, r_1), \cdots, \mathbf{a}(\theta_L, r_L)] \in \mathbb{C}^{N_t \times L}$ and stacking the associated complex path gain as $\mathbf{b} = \sqrt{\frac{1}{L}} \left[\alpha_1, \cdots, \alpha_L\right]^T \in \mathbb{C}^L$, the covariance matrix can be rewritten as

$$\mathbf{R}_y = \mathbf{W}^H \mathbf{A} \underbrace{\mathbb{E}\left\{\mathbf{s}(t)\mathbf{s}(t)^H\right\}}_{\mathbf{R}_s} \mathbf{A}^H \mathbf{W} + \sigma^2 \mathbf{I}, \tag{5}$$

where $\mathbf{s}(t) = \mathbf{b}x(t)$. Take the EVD of the covariance matrix given in (5), we can obtain

$$\mathbf{R}_y = [\mathbf{U}_S \mathbf{U}_N] \begin{bmatrix} \mathbf{\Sigma}_S & \\ & \mathbf{\Sigma}_N \end{bmatrix} [\mathbf{U}_S \mathbf{U}_N]^H, \tag{6}$$

where $\mathbf{U}_S$ and $\mathbf{U}_N$ represent the eigenvectors associated with the L largest eigenvalues, denoted by $\mathbf{\Sigma}_S$, and the remaining $N_{\mathrm{RF}} - L$ smallest eigenvalues, denoted by $\mathbf{\Sigma}_N$, respectively. Multiplying $\mathbf{U}_N$ in the right side of (5) and (6), we have $\mathbf{R}_y \mathbf{U}_N = \sigma^2 \mathbf{U}_N$ and $\mathbf{R}_y \mathbf{U}_N = \mathbf{W}^H \mathbf{A} \mathbf{R}_s \mathbf{A}^H \mathbf{W} \mathbf{U}_N + \sigma^2 \mathbf{U}_N$, which indicates that $\mathbf{W}\mathbf{U}_N^H \mathbf{W}^H \mathbf{A} \mathbf{R}_s \mathbf{A}^H \mathbf{W} \mathbf{U}_N = \mathbf{0}$. So, when $\mathbf{R}_s$ is non-singular, $\mathbf{A}^H \mathbf{W} \mathbf{U}_N = \mathbf{0}$. In other words, each path vector $\mathbf{a}^H(\theta, r)$ correlated with the beamforming matrix $\mathbf{W}$ contributes a zero to the null spectrum $\mathbf{a}^H(\theta, r)\mathbf{W}\mathbf{U}_N\mathbf{U}_N^H\mathbf{W}^H\mathbf{a}(\theta, r)$. Then the angle and distance pairs $\left[\left(\hat{\theta}_1, \hat{r}_1\right), \cdots, \left(\hat{\theta}_L, \hat{r}_L\right)\right]$ can be estimated by finding the L peak points in the BMUSIC spectrum[1]

$$P(\theta, r) = \frac{1}{\mathbf{a}^H(\theta, r)\mathbf{W}\mathbf{U}_N\mathbf{U}_N^H\mathbf{W}^H\mathbf{a}(\theta, r)}. \tag{7}$$

Furthermore, the path gain can be estimated as

$$\hat{\mathbf{b}} = \frac{1}{T} \sum_{t=1}^{T} \mathrm{pinv}\left(\mathbf{W}^{\mathrm{H}} \hat{\mathbf{A}}\right) \mathbf{y}(\mathrm{t}), \tag{8}$$

where $\mathrm{pinv}(\cdot)$ denotes pseudo-reverse operation. Then the estimated channel can be written as $\hat{\mathbf{h}} = \hat{\mathbf{A}}\hat{\mathbf{b}}$.

Obviously, the performance of the BMUSIC method is related to the beamforming matrix $\mathbf{W}$. In the conventional BMUSIC method for far-field channel estimation, $\mathbf{W}$ is generally generated from a DFT codebook [11] with N_t well-selected direction-related parameters $\{\phi_{N_t}\}_{i=1}^{N_t}$ [2]. However, the conventional DFT codebook mismatches with the near-filed channel structure, thereby resulting in significant gain loss for each signal path and consequently, substantial errors in channel estimation. To address this issue we propose introducing pre-compensation distance $r_{c,\theta}$ for each angle θ and re-formulated the beamforming codebook as

$$\mathbf{C} = [\mathbf{a}(\phi_1, r_{c,\phi_1}), \mathbf{a}(\phi_2, r_{c,\phi_2}), \cdots, \mathbf{a}(\phi_{N_t}, r_{c,\phi_{N_t}})], \tag{9}$$

[1] L is assumed to be known which can be obtained by the minim description length principle or the magnitude of the eigenvalues in practice [11].

[2] The angle range needed to construct $\mathbf{W}$ can be determined through a wide-beam search by activating a limited subset of antennas in the center of XL-array during the initialization phase[[12]. $\mathbf{W}$ can be obtained by selecting appropriate columns from $\mathbf{C}$ based on the angle range.

where ϕ_i is the uniform sampling angle satisfying that $\sin\phi_i = \frac{2i-N_t-1}{N_t}, i = 1, 2, \cdots, N_t$. In the following, we will propose three different strategies for optimizing $r_{c,\theta}$.

III. Pre-compensation Distance Optimization

According to the previous analysis, the goal is to maximize the gain (correlation) of the fine-tuned codebook $\mathbf{C}$ in the near field. This involves maximizing the correlation coefficient between the steering vector corresponding to $r_{c,\theta}$ and the focusing vector corresponding to the true distance r. This ensures that the near-field signals are maximally strengthened through the combining matrix $\mathbf{W}$. Afterward, the signal subspace can be preserved as completely as possible and BMUSIC can be performed. Define $\rho(r_{c,\theta}, r)$ as the correlation coefficient between the steering vector corresponding to $(\theta, r_{c,\theta})$ and the steering vector corresponding to the true distance (θ, r),

$$\rho(r_{c,\theta}, r) = |\mathbf{a}(\theta, r_{c,\theta})^H \mathbf{a}(\theta, r)|. \tag{10}$$

We aim to maximize $\rho(r_{c,\theta}, r)$ where r is any true near-field distance. Similar pre-compensation distance methods are mentioned in [6], where far-field steering vectors are used, i.e., $r_{c,\theta} = \infty$. It is obvious that it is not the optimal choice and we should find a more suitable $r_{c,\theta}$. When the distance range of paths in the near field $[r_{\min}, r_{\max}]$ is known [3], we propose the following three strategies:

*1) **MCI**:* This strategy is to find $r_{c,\theta} \in [r_{\min}, r_{\max}]$ that maximizes the integral of $\rho(r_{c,\theta}, r)$ in $[r_{\min}, r_{\max}]$, i.e. maximizing the sum of the correlation between the pre-compensation steering vector and all potential true steering vectors. The problem can be formulated as $\max_{r_{c,\theta}} \int_{r_{\min}}^{r_{\max}} \rho(r_{c,\theta}, r)\mathrm{d}r$. Further, take the square of $\rho(r_{c,\theta}, r)$, we have $\rho^2(r_{c,\theta}, r) = \mathbf{a}(\theta, r_{c,\theta})^H \mathbf{a}(\theta, r)\mathbf{a}(\theta, r)^H \mathbf{a}(\theta, r_{c,\theta})$. The problem can be equivalently written as

$$\max_{r_{c,\theta}} \int_{r_{\min}}^{r_{\max}} \rho^2(r_{c,\theta}, r)\mathrm{d}r = \min_{r_{c,\theta}} \; -\mathbf{a}(\theta, r_{c,\theta})^H \mathbf{M}\mathbf{a}(\theta, r_{c,\theta}) \tag{11}$$

where $\mathbf{M} = \int_{r_{\min}}^{r_{\max}} \mathbf{a}(\theta, r)\mathbf{a}(\theta, r)^H \mathrm{d}r$ is a definite integral. This is a problem of finding the minimum of a single-variable function within a fixed interval and the function has a continuous derivative. Employing multiple initial points for the gradient-based local search to obtain the global optimal solution, which can be achieved by the Global Optimization Toolbox in MATLAB.

*2) **MMC**:* The MMC strategy aims to maximize the minimum value of $\rho(r_{c,\theta}, r)$ in $[r_{\min}, r_{\max}]$, i.e.

$$\max_{r_{c,\theta}} \min_{r \in [r_{\min}, r_{\max}]} \rho(r_{c,\theta}, r). \tag{12}$$

Furthermore, Lemma 1 in [3] has derived that $\rho(r_{c,\theta}, r) = |G(\beta)| = \frac{1}{\beta}|C(\beta) + jS(\beta)|$, where $C(\beta)$ and $S(\beta)$ are Fresnel functions, $\beta = \sqrt{\frac{N_t^2 d^2 \cos^2(\theta)}{2\lambda}\left|\frac{1}{r_{c,\theta}} - \frac{1}{r}\right|}$. And $|G(\beta)|$ shows an approximately monotonically decreasing trend with the increase of β. Clearly, the increase or decrease of β depends on $|\frac{1}{r_{c,\theta}} - \frac{1}{r}|$. So we have that $\rho(r_{c,\theta}, r)$ exhibits an approximately monotonically increasing trend when $r < r_{c,\theta}$ and a decreasing trend $r > r_{c,\theta}$, which is also shown in Fig. 2. For such a trend, the minimum value appears at the ends of $[r_{\min}, r_{\max}]$. Then the problem can be approximated as simultaneously maximizing both $\rho(r_{c,\theta}, r_{\min})$ and $\rho(r_{c,\theta}, r_{\max})$.

[3]The distance range can be obtained through several methods, including but not limited to.: the near-field operational range of BS, the predictability of user-dense areas (e.g. shopping malls), and BS traffic analysis prediction.

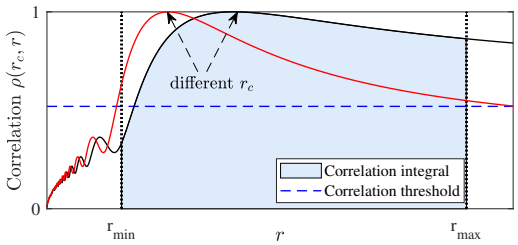


Fig. 2. The correlation schematic by different r_c over $[r_{\min}, r_{\max}]$. MCI aims to maximize the sum of correlations between r_c and all true distances, i.e. maximize the area under the blue region in Fig. 2; however, this can lead to poor correlation at some distances. MMC aims to maximize the minimum correlation between r_c and all distances, thus avoiding the drawbacks of MCI. EMCT is an extension of MMC, designed to ensure that the minimum correlation between r_c and all distances exceeds a certain threshold, i.e. the curve in Fig. 2 always remains above the blue line in the range $[r_{\min}, r_{\max}]$.

Similarly, since $r \in [r_{\min}, r_{\max}]$, when $r_{c,\theta}$ increases, $|\frac{1}{r_{c,\theta}} - \frac{1}{r_{\min}}|$ increases while $|\frac{1}{r_{c,\theta}} - \frac{1}{r_{\min}}|$ decreases, resulting that $\rho(r_{c,\theta}, r_{\min})$ decreases while $\rho(r_{c,\theta}, r_{\max})$ increases. So, $\rho(r_{c,\theta}, r_{\min})$ and $\rho(r_{c,\theta}, r_{\max})$ have opposite trends with r_c varies. Therefore, the solution to the problem is the balance state where $r_{c,\theta}$ satisfies $\rho(r_{c,\theta}, r_{\min}) = \rho(r_{c,\theta}, r_{\max})$, i.e.,

$$\left|\frac{1}{r_{c,\theta}} - \frac{1}{r_{\min}}\right| = \left|\frac{1}{r_{c,\theta}} - \frac{1}{r_{\max}}\right|. \tag{13}$$

Hence, we can obtain an approximate closed-form solution

$$r_{c,\theta} = \frac{2r_{\min}r_{\max}}{r_{\min} + r_{\max}}, \; \forall\theta. \tag{14}$$

*3) **EMCT**:* This strategy aims to find $r_{c,\theta}$ that ensures the minimum value of $\rho(r_{c,\theta}, r)$ in $[r_{\min}, r_{\max}]$ exceeds a threshold $\epsilon \in [0, 1]$. Based on the analysis in MMC, EMCT strategy is equal to make that $\rho(r_{c,\theta}, r_{\min}) > \epsilon$ and $\rho(r_{c,\theta}, r_{\max}) > \epsilon$ simultaneously. Utilizing numerical methods to solve $\rho(r_{c,\theta}, r_{\max}) = \epsilon$ and $\rho(r_{c,\theta}, r_{\min}) = \epsilon$ in $[r_{\min}, r_{\max}]$, we can obtain r_{c1}^{θ} and r_{c2}^{θ}, respectively. According the increasing and decreasing properties of $\rho(r_{c,\theta}, r_{\min})$ and $\rho(r_{c,\theta}, r_{\max})$, we have

$$\begin{cases} r_{c,\theta} > r_{c1}^{\theta}, & \text{when } \rho(r_{c,\theta}, r_{\max}) > \epsilon \\ r_{c,\theta} < r_{c2}^{\theta}, & \text{when } \rho(r_{c,\theta}, r_{\min}) > \epsilon \end{cases} \tag{15}$$

Therefore, the EMCT strategy obtains that $r_{c,\theta} \in [r_{c1}^{\theta}, r_{c2}^{\theta}]$.

IV. Two-Stage BMUSIC

With the optimized $r_{c,\theta}$, the channel can be estimated through the BMUSIC method introduced in Section II. Compared to the conventional 2D MUSIC [9], BMUSIC reduces the complexity of EVD from N_t^3 to N_{RF}^3. However, the time-consuming 2D search still remains. Additionally, a fundamental premise for effective MUSIC estimation is that $\mathbf{R}_s$

is non-singular, i.e. $\mathrm{rank}(\mathbf{R}_s) = L$. However, according to $\mathbf{s}(t) = \mathbf{b}x(t)$, $\mathrm{rank}(\mathbf{R}_\mathrm{s})$ is always 1. This implies that when multiple coherent paths are present $L > 1$, the aforementioned method becomes ineffective. To address the two drawbacks, we first propose a two-stage BMUSIC method to reduce the computational complexity of the 2D search and then deal with the issue of multipath coherence.

A. Two-stage BMUSIC

Specifically, we decouple the 2D search described in Eq. (7) into two distinct 1D searches. The initial stage involves searching for angles by leveraging the pre-compensation distance $r_{c,\theta}$ given in Section III to determine the path direction. Subsequently, the second stage focuses on searching for distances with the path angles obtained from the first stage. Thus, the two-stage BMUSIC can be summarized as: **Stage 1:** Find the peak point of the spectrum function $f(\theta)$ given as Eq. (16) to get the direction vector $\hat{\theta} = [\hat{\theta}_1, \cdots, \hat{\theta}_L]$:

$$f(\theta) = \frac{1}{\mathbf{a}^H(\theta, r_{c,\theta})\mathbf{W}\mathbf{U}_N\mathbf{U}_N^H\mathbf{W}^H\mathbf{a}(\theta, r_{c,\theta})}, \ \theta \in \Omega_\theta, \quad (16)$$

where Ω_θ is the angle search space[4], $\mathbf{U}_N$ is defined in Eq. (6).

Stage 2: Find the peak point of the spectrum function $g(r)$ given as Eq. (17) to get the distance vector $\hat{\mathbf{r}} = [\hat{r}_1, \cdots, \hat{r}_L]$:

$$g(r) = \frac{1}{\mathbf{a}^H(\hat{\theta}, r)\mathbf{W}\mathbf{U}_N\mathbf{U}_N^H\mathbf{W}^H\mathbf{a}(\hat{\theta}, r)}, r \in \Omega_r, \quad (17)$$

where Ω_r is the distance search space.

Remark 1. *As stated in [13], there exists a direct relationship between $\rho(r_{c,\theta}, r)$ and the angle estimation error in the first stage; specifically, a larger $\rho(r_{c,\theta}, r)$ results in a reduced angle estimation error. This observation aligns with the analysis in Section III, thereby justifying the direct application of $r_{c,\theta}$ from Section III in the two-stage BMUSIC.*

B. Switch Transformation for Multipath Coherence

To mitigate the issue of multipath coherence identified in the BMUSIC algorithm, we propose to leverage the concept of spatial smoothing, that is, employing a switching circuit at the BS. This enables dynamic selection of the active antenna subarray from the N_t-element antenna array. Specifically, we define Q transformation matrices $\mathbf{Z}_q, q = 1, 2, \cdots, Q$ as

$$\mathbf{Z}_q = \left[\mathbf{0}_{K\times(q-1)}|\mathbf{I}_K|\mathbf{0}_{K\times(Q-q)}\right]^T \in \mathbb{R}^{N_t\times K}, \quad (18)$$

where $K = N_t - Q + 1$, $\mathbf{0}$ denotes the zero matrix. Each transformation matrix $\mathbf{Z}_q$ represents the state of the switching circuit. In this regard, a new dimensional-reduced beamforming matrix $\tilde{\mathbf{W}} \in \mathbb{C}^{K\times N_\mathrm{RF}}$ is generated by taking the first K rows of $\mathbf{W}$. Therefore, the actual weight matrix configured on the antennas is $\mathbf{Z}_q\tilde{\mathbf{W}}$. For clarity, we denote the signal arriving at the antenna as $\hat{\mathbf{y}}(t) = \mathbf{h}x(t) + \mathbf{n}(t) \in \mathbb{C}^{N_t}$. Assuming that the switching circuit changes Q times within a coherence time duration, each of them gives us a signal output as

$$\tilde{\mathbf{y}}_q(t) = \tilde{\mathbf{W}}^H\mathbf{Z}_q^T\hat{\mathbf{y}}(t). \quad (19)$$

[4]Ω_θ can be the angle range corresponding to $\mathbf{W}$.

TABLE I: Complexity Comparison.

Method	Computational Complexity
Proposed method	$O\left(N_\mathrm{RF}^3 + N_\mathrm{RF}^2(G_1+G_2)\right)$
RD-MUSIC [10]	$O\left(N_t^3 + N_tG_1G_2\right)$
2D-MUSIC [9]	$O\left(N_t^3 + N_t^2G_1G_2\right)$
P-OMP [3]	$O\left(TN_\mathrm{RF}N_tG_1G_2\right)$

Accordingly, the average covariance matrix for the Q outputs can be written as

$$\begin{aligned}\tilde{\mathbf{R}}_y &= \frac{1}{Q}\sum_{q=1}^{Q}\frac{1}{T}\sum_{t=1}^{T}\tilde{\mathbf{y}}_q(t)\tilde{\mathbf{y}}_q(t)^H \\ &= \tilde{\mathbf{W}}^H\frac{1}{Q}\sum_{q=1}^{Q}\mathbf{Z}_q^T\hat{\mathbf{R}}_y\mathbf{Z}_q\tilde{\mathbf{W}} = \tilde{\mathbf{W}}^H\frac{1}{Q}\sum_{q=1}^{Q}\hat{\mathbf{R}}_q\tilde{\mathbf{W}},\end{aligned} \quad (20)$$

where $\hat{\mathbf{R}}_y = \frac{1}{T}\sum_{t=1}^{T}\hat{\mathbf{y}}(t)\hat{\mathbf{y}}(t)^H$, and $\hat{\mathbf{R}}_q = \mathbf{Z}_q^T\hat{\mathbf{R}}_y\mathbf{Z}_q$. When multipath coherence occurs, the estimated covariance matrix $\hat{\mathbf{R}}_y$ exhibits rank deficiency. By adopting the switch transformation, we obtain $\frac{1}{Q}\sum_{q=1}^{Q}\hat{\mathbf{R}}_q$, i.e., the classical forward-smoothed covariance matrix (Eq. (31), [14]). Forward smoothing can restore the rank of the matrix, thereby resolving the issue of multipath coherence [14]. Then, using $\tilde{\mathbf{R}}_y$ for EVD in (6), and employing $\tilde{\mathbf{W}}$ and $r_{c,\theta}$ for the 1D searches in (16) and (17), we can obtain the estimated parameters[5].

Remark 2. *Switching the switching circuit in the coherence time maybe a challenge. A practical approach is conducting switch transitions across various pilot phases. Specifically, the pilots can be divided into Q segments, with each segment using a distinct switch state $\mathbf{Z}_q$ to obtain $\tilde{\mathbf{W}}^H\hat{\mathbf{R}}_q\tilde{\mathbf{W}}$. Furthermore, to ensure effective decorrelation, Q is generally recommended to be more than twice the number of paths L [14].*

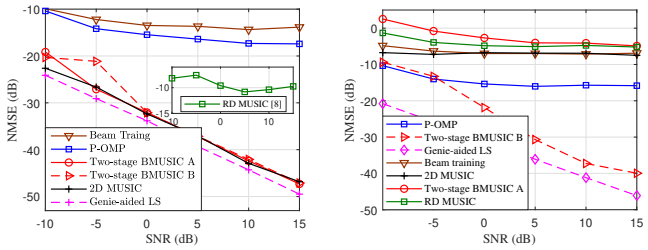
C. Complexity Analysis

The main computational complexity of the proposed algorithm comes from: 1) the complexity of EVD in (8), which is $O\left(N_\mathrm{RF}^3\right)$; 2) the complexity of the grid search in (16) and (17), which is $O\left(G_1N_\mathrm{RF}^2 + LG_2N_\mathrm{RF}^2\right)$, with G_1, G_2 standing for the grid number of angle and distance; 3) the complexity for the path gain estimation, which is $O\left(L^3\right)$. Due to the characteristics of sparse channels, L is usually small. Therefore, the total complexity is expressed as $O\left(N_\mathrm{RF}^3 + N_\mathrm{RF}^2(G_1+G_2)\right)$. Table I shows the computational complexity of different methods which demonstrates that the proposed method achieves a significant reduction in computational complexity.

V. Simulation Results

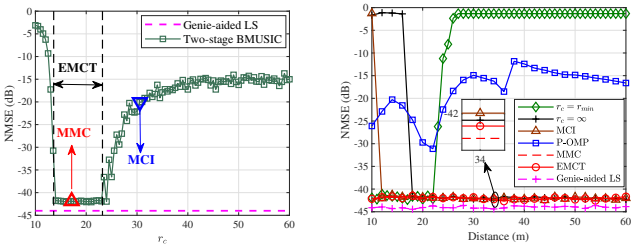
In this section, we provide simulation results for validating the performance of the proposed method. The simulation parameters are set as: $f_c = 60$ GHz, $N_t = 256$, $N_\mathrm{RF} = 10$. Furthermore, we set $Q = 5$, following the parameter settings from reference [14]. We assume that all path distances are

[5]The dimension of the steering vector used for 1D search should match the dimension of the column vectors of $\tilde{\mathbf{W}}$.



(a) $L = 1$. (b) $L = 2$.

Fig. 3. NMSE performance comparison versus SNR, $T = 10$, the MMC strategy is used in the proposed methods.



(a) Pre-compensation distance. (b) User distance.

Fig. 4. NMSE performance comparison versus pre-compensation distance (a) and user distance (b), $T = 10, L = 1$, SNR=10 dB.

generated within $[10, 60]$ meters and the path angles are generated within $[-\frac{\pi}{3}, \frac{\pi}{3}]$. We adopted the normally distributed channel path gain in this work. The normalized mean square error (NMSE) is chosen as the metric, i.e. $\mathbb{E}\left\{\frac{\|\mathbf{h}-\hat{\mathbf{h}}\|_2^2}{\|\mathbf{h}\|_2^2}\right\}$. For clarity, we denote the method without the switching circuit as Two-stage BMUSIC A, and the method employing the switching circuit as Two-stage BMUSIC B. The original 2D MUSIC [9], RD MUSIC [10], P-OMP [3], and beam training [6] algorithms are chosen for comparison. The Genie-aided-LS algorithm serves as the performance lower bound where the path distances and angles are perfectly known.

In Fig. 3(a), we show the NMSE performance versus the SNRs for all algorithms with $L = 1$. It can be observed that the proposed algorithm outperforms the comparison algorithms. Since the RD MUSIC algorithm experiences angle ambiguity in systems with half-wavelength spacing, its performance is significantly poor. Our proposed method with the MMC strategy achieves nearly the same performance as 2D MUSIC with a much lower complexity when the SNR is greater than -5dB. Additionally, both proposed method **A** and method **B** are applicable in this scenario. The results for a multipath scenario with $L = 2$ are shown in Fig. 3(b). Clearly, the coherence caused by multipath greatly diminishes the performance of all algorithms related to MUSIC. The proposed method **B**, which introduces a switch transformation matrix, can effectively counteract multipath effects and demonstrates performance closest to the lower bound.

In addition, the NMSE performance with respect to different pre-compensation distances when $\theta = 0$ is evaluated in Fig. 4(a). For the EMCT, we establish a threshold value of $\epsilon = 0.52$. Among different pre-compensation distances, the EMCT and MMC strategies can achieve the best estimation performance. This indicates that the minimum value of $\rho(r_{c,\theta}, r)$ plays a pivotal role in determining overall estimation accuracy. Fig. 4(b) provides a comparison of different schemes' performance as a function of user distance. It can be seen that MCI has poor estimation performance when the user distance is less than 12 m. This is due to the MCI strategy yielding low correlations $\rho(r_{c,\theta}, r)$, for certain user distances. For the same reason, when $r_c = r_{\min}$ and $r_c = \infty$, estimation performance degrades at certain user distances. Moreover, due to the sparseness of the distance sampling points, the performance of P-OMP fluctuates with the user distance. In contrast, the proposed two-stage BMUSIC method using MMC or EMCT has stable performance and outperforms the comparison methods.

VI. Conclusions

In this letter, we propose a BMUSIC-based near-field channel estimation method for hybrid XL-MIMO. We introduce a pre-compensation distance and detail three strategies for determining it. Additionally, we simplify the 2D search problem into 1D searches with a two-stage strategy. To tackle signal coherence in multipath environments, we develop a switch transformation matrix for effective antenna array selection. Simulation results show that the proposed methods improve estimation performance and reduce computational complexity.

References

[1] Y. Liu, Z. Wang, J. Xu, C. Ouyang, X. Mu, and R. Schober, "Near-field communications: A tutorial review," *IEEE Open J. Commun. Soc.*, vol. 4, pp. 1999–2049, Aug. 2023.

[2] S. Guo and K. Qu, "Beamspace modulation for near field capacity improvement in XL-MIMO communications," *IEEE Wireless Commun. Lett.*, vol. 12, no. 8, pp. 1434–1438, Aug. 2023.

[3] M. Cui and L. Dai, "Channel estimation for extremely large-scale MIMO: Far-field or near-field," *IEEE Trans. Commun.*, vol. 70, no. 4, pp. 2663–2677, Apr. 2022.

[4] Y. Lu and L. Dai, "Near-field channel estimation in mixed LoS/NLoS environments for extremely large-scale MIMO systems," *IEEE Trans. Commun.*, vol. 71, no. 6, pp. 3694–3707, June 2023.

[5] S. Yang, W. Lyu, Z. Hu, Z. Zhang, and C. Yuen, "Channel estimation for near-field XL-RIS-aided mmwave hybrid beamforming architectures," *IEEE Trans. Veh. Technol.*, vol. 72, no. 8, pp. 11 029–11 034, Aug. 2023.

[6] Y. Zhang, X. Wu, and C. You, "Fast near-field beam training for extremely large-scale array," *IEEE Wireless Commun. Lett.*, vol. 11, no. 12, pp. 2625–2629, Dec. 2022.

[7] J. Chen, F. Gao, M. Jian, and W. Yuan, "Hierarchical codebook design for near-field mmWave MIMO communications systems," *IEEE Wireless Commun. Lett.*, vol. 12, no. 11, pp. 1926–1930, Nov. 2023.

[8] C. Wu, C. You, Y. Liu, L. Chen, and S. Shi, "Two-stage hierarchical beam training for near-field communications," *IEEE Trans. Veh. Technol.*, vol. 73, no. 2, pp. 2032–2044, Feb. 2024.

[9] Y.-D. Huang and M. Barkat, "Near-field multiple source localization by passive sensor array," *IEEE Trans. Antennas Propag.*, vol. 39, no. 7, pp. 968–975, July 1991.

[10] X. Zhang, W. Chen, W. Zheng, Z. Xia, and Y. Wang, "Localization of near-field sources: A reduced-dimension MUSIC algorithm," *IEEE Commun. Lett.*, vol. 22, no. 7, pp. 1422–1425, July 2018.

[11] Z. Guo, X. Wang, and W. Heng, "Millimeter-wave channel estimation based on 2-D beamspace MUSIC method," *IEEE Trans. Wireless Commun.*, vol. 16, no. 8, pp. 5384–5394, Aug. 2017.

[12] Z. Xiao, T. He, P. Xia, and X.-G. Xia, "Hierarchical codebook design for beamforming training in millimeter-wave communication," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3380–3392, May 2016.

[13] J. He and T. Shu, "Effect of approximate planar wavefront on far-field direction finding," *IEEE Commun. Lett.*, vol. 26, pp. 657–661, Mar. 2022.

[14] Y. Zhong, S. Yuan, and L. Qiu, "Multiple damage detection on aircraft composite structures using near-field MUSIC algorithm," *Sens. Actuator A Phys*, vol. 214, pp. 234–244, Aug. 2014.