Electromagnetic-Consistent Codebook Design for Emerging 3-D Arrays

Yuqi Hu, *Graduate Student Member, IEEE*, Chongwen Huang, *Senior Member, IEEE*, Li Wei, *Member, IEEE*, Xue Wang, *Senior Member, IEEE*, Wei E. I. Sha, *Fellow, IEEE*, Jun Yang, *Senior Member, IEEE*, Zhaoyang Zhang, *Senior Member, IEEE*, Jennifer Simonjan, *Senior Member, IEEE*, Osama M. Bushnaq, *Member, IEEE*, Sami Muhaidat, *Senior Member, IEEE*, and Mérouane Debbah, *Fellow, IEEE*

***Abstract*—The communication performance of traditional two-dimensional (2D) antenna arrays is approaching its theoretical limit under constraints of physical size and hardware costs, thus failing to meet the escalating demands of wireless communications. While double-layer three-dimensional (3D) antenna arrays presents a breakthrough for overcoming this bottleneck by exploiting the additional degrees of freedom, its implementation is hindered by several challenges, notably the issues of codebook design. In this paper, we propose a novel codebook scheme tailored for 3D antenna array structures. Specifically, an angle-distance-aware codebook for 3D antenna arrays is designed to cater to both near-field and far-field scenarios by minimizing inter-beam interference, with proven asymptotic orthogonality. Furthermore, evanescent codewords for both regions are effectively eliminated to improve codebook construction efficiency. Simulation results illustrate the superior performance of the proposed codebook over 2D baselines, with a 29% and 12% narrower angular and distance beamwidth of h=λ, and a 27% gain in spectral efficiency of $h = 0.5\lambda$, owing to the vertical dimension. Moreover, practical mutual coupling that manifests as beam deviations and broadening is analyzed to establish a basis for future work.**

***Index Terms*—3D antenna array, codebook, evanescent codewords, mutual coupling.**

Received 23 November 2025; revised 15 March 2026; accepted 13 May 2026. Date of publication 4 June 2026; date of current version 15 June 2026. The work was supported by National Natural Science Foundation of China under Grant 62331023 and 62394292, Zhejiang Provincial Natural Science Foundation for Distinguished Young Scholars under Grant LRG26F010001, Major Science and Technology Project of Changchun under Grant 2024WX04, China National Key R&D Program under Grant 2025ZD1301900, 2021YFA1000500 and 2023YFB2904804, Fundamental Research Funds for the Central Universities and Zhejiang University Global Partnership Fund. The associate editor coordinating the review of this article and approving it for publication was W. Mei. *(Corresponding authors: Chongwen Huang; Li Wei.)*

Yuqi Hu, Chongwen Huang, Wei E. I. Sha, and Zhaoyang Zhang are with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China, and also with Zhejiang Provincial Key Laboratory of Multi-Modal Communication Networks and Intelligent Information Processing and the National Key Laboratory of Millimeter-Wave and Terahertz Remote Sensing, Hangzhou 310027, China (e-mail: huyuqi@zju.edu.cn; chongwenhuang@zju.edu.cn; weisha@zju.edu.cn; zhzy@zju.edu.cn).

Li Wei is with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China, and also with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798 (e-mail: weili_xd@163.com).

Xue Wang is with the College of Communication Engineering, Jilin University, Changchun 130012, China (e-mail: txwangxue@jlu.edu.cn).

Jun Yang is with the State Key Laboratory of Mobile Network and Mobile Multimedia Technology, Shenzhen 518055, China, and also with the Wireless Product R&D Institute, ZTE Corporation, Shenzhen 518055, China (e-mail: yang.jun10@zte.com.cn).

Jennifer Simonjan and Osama M. Bushnaq are with the Technology Innovation Institute, Abu Dhabi, United Arab Emirates (e-mail: Jennifer.Simonjan@tii.ae; Osama.Bushnaq@tii.ae).

Sami Muhaidat is with the College of Computing and Mathematical Sciences, Khalifa University, Abu Dhabi, United Arab Emirates (e-mail: sami.muhaidat@ku.ac.ae).

Mérouane Debbah is with the Research Institute for Digital Future, Khalifa University, Abu Dhabi, United Arab Emirates (e-mail: merouane.debbah@ku.ac.ae).

Digital Object Identifier 10.1109/TCOMM.2026.3700533

I. Introduction

With emerging technologies advancing, application fields expanding, and the global market growing, the demand for wireless communications has exhibited an upward trend. In this case, multiple-input multiple-output (MIMO) technology, which deploys multiple antennas at the transmitting and receiving ends to make full utilization of spatial resources, has drawn considerable attention due to its outstanding spectral efficiency and capacity enhancement capabilities [1], [2], [3], [4]. The traditional MIMO systems are typically implemented through two-dimensional (2D) planar arrays, e.g., uniform linear array (ULA) and uniform planar array (UPA) [5]. Normally, to enhance the performance of them, enlarging aperture area is a primary and well-adopted approach, which is theoretically substantiated by the Hannan limit [6], [7].

However, in practical communication scenarios, the aperture extension is usually constrained by the physical size, construction costs, and other factors. Consequently, the gain of traditional 2D antenna arrays cannot constantly increase in practical scenarios, making the explosive data traffic and robust network performance inaccessible. In such a case, it is crucial to break through the physical constraints of 2D antenna arrays to achieve the desired capabilities, e.g., improved signal coverage, reduced interference, and enhanced energy efficiency. Resorting to the antenna analysis in [8], [9], and [10], the deployment of volumetric antenna arrays, i.e., three-dimensional (3D) antenna arrays, brings new opportunities to further exploit all available spatial resources, surpassing the limitations of conventional 2D antenna arrays.

Specifically, unlike the planar or linear antenna arrays arranged in horizontal directions, 3D arrays are built within

a volumetric infrastructure by introducing an additional layer, leading to a lower spatial correlation between adjacent antennas and more degrees of freedom (DoFs). According to [8], the performance of 3D-array-assisted MIMO across various multi-path communication scenarios is investigated, showing the significant spatial multiplexing performance and channel capacity enhancement of MIMO systems over traditional 2D-array-MIMO. The work [10] theoretically derives critical parameters (e.g., feasible region and array gain) for 3D antenna arrays, underscoring their substantial advantages over 2D arrays given the same aperture area. Numerical results in [10] reveal that the directivity of a 3D array with a layer spacing of 0.75λ is increased by more than 60% compared to that of a 2D array at a large scanning angle of 90°. The related works all demonstrate the potential of 3D arrays in satisfying the expanding requirements of future communication systems. However, research on 3D arrays is still in its early stages, with numerous challenges yet to be overcome.

The primary challenge arises from efficient codebook design. Being an effective tool in addressing various technical issues such as mitigating signal transmission losses, ensuring comprehensive spatial coverage, and improving the efficiency of beam sweeping, codebook design has undoubtedly become an indispensable component in wireless communications [11], [12], [13]. Specifically, a codebook consists of a set of codewords, each representing analog phase shift or magnitude and phase shift combinations, to generate the desired beam [14], and extensive research has been conducted in this area. The IEEE 802.15.3c standard employs a codebook with only four phase shifts, offering a simpler design but yielding a lower beamforming gain [15]. In contrast, the discrete Fourier transform (DFT) codebook involves more phase shifts to achieve superior gain [16], [17], and beam steering codebook allows for finer quantization of angles [18], [19]. Additionally, the Kronecker-product codebook [20], [21], capturing both elevation and azimuth angles through the Kronecker product, is embraced by most work for its capability to approximate optimal beamforming vectors of spatially correlated channels [22]. There are various works contributing to codebook design from diverse perspectives, for further details, one may refer to [23], [24], [25], [26], [27], [28], [29], and [30].

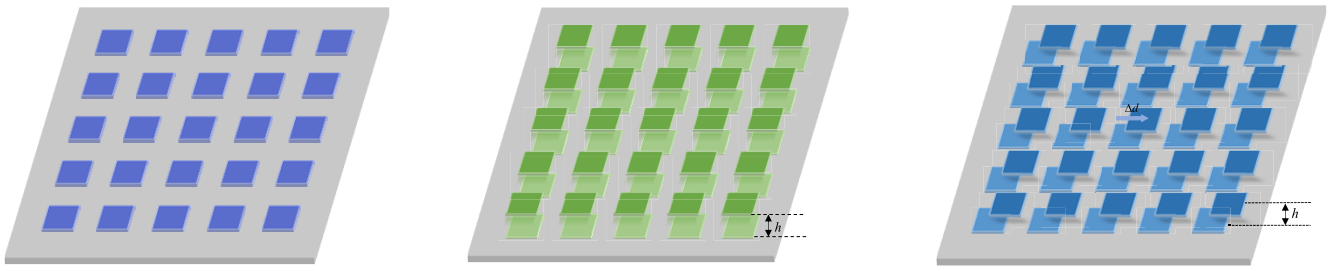
Although existing works have laid a solid foundation for codebook design, extending these designs to 3D arrays still confronts several difficulties. First, from a structural perspective, the emergence of 3D antenna arrays introduces a volumetric infrastructure that fundamentally changes the spatial degrees of freedom (DoFs). Existing codebook designs, predominantly centered on traditional 2D planar arrays, lack the architectural flexibility to exploit this additional vertical dimension. Second, from a signal modeling perspective, the employment of advanced technologies has shifted research core from conventional far-field regions to the burgeoning near-field regions, where characteristics are jointly determined by angles and distances [31]. Crucially, while some recent near-field codebooks [32], [33], [34] have introduced joint angle-distance sampling, they are mathematically derived based on planar apertures. This model neglects the non-linear phase variations introduced by the vertical layer spacing h and horizontal offsets Δd inherent in 3D structures. Consequently, directly applying these 2D-based joint sampling schemes leads to significant beam defocusing and misalignment, as the underlying signal models cannot align the spherical wavefronts across different volumetric layers. In parallel, research on holographic MIMO has also emphasized the necessity of angular-distance based channel modeling and estimation to capture the unique spherical wavefront characteristics [35], [36]. Notably, the integration of massive antenna elements and abundant range information in near-field scenarios exacerbate the codebook design complexity, which is neglected in the current near-field codebook design [37], [38], [39], thus a more efficient solution is required.

Given these challenges, the question arises: How to devise an efficient near-field codebook tailored for 3D antenna array structures?

In this paper, we devote our attention towards devising a universally applicable codebook design scheme aligning with the complex 3D array structures, which brings substantial enhancement of beam performance by exploiting the layer spacing characteristic. Specifically, our contributions are summarized as follows:

- We first propose a novel layer-spacing-aware codebook design scheme for the constructed 3D array structure, which effectively compensates for volumetric phase deviations through a joint angle-distance sampling mechanism applicable to both near-field and far-field regions. The design criterion is based on coherence minimization of codewords, establishing uniform angular sampling and nonuniform distance sampling schemes that explicitly incorporate the 3D geometric parameters. Simulations numerically prove the performance improvement of the beam stemming from the layer spacing of 3D arrays.
- To further relieve computational burden, we derive a more efficient codebook design through the analysis of channel characteristics, eliminating evanescent codewords that merely contribute to the beam generation. Based on this, a fundamental analysis of the wavenumber in 3D arrays is presented, revealing that the performance gain of 3D arrays stems from the unique capability in vertical dimension rather than from utilizing evanescent modes.
- Numerical results demonstrate the enhancement of system performance in terms of the beamwidth and the spectral efficiency attributed to the layer spacing. Additionally, we explore the potential impacts of implementing the proposed codebook design scheme with mutual coupling (MC), providing a theoretical foundation for future aplication and optimization.

The remainder of this paper is organized as follows. In Section II, we present the system model of 3D arrays, detailing the geometry and the corresponding channel models. Section III introduces the near-field codebook design for 3D arrays, with an analysis of its asymptotic orthogonality and an extension to far-field scenarios. In Section IV, the evanescent codewords are derived and the wavenumber properties is conducted. Numerical results are provided in Section V and conclusions are drawn in Section IV.



(a) Type 1: 2D planar array. (b) Type 2: 3D stacked array. (c) Type 3: 3D staggered array.

Fig. 1. Comparison of different array topologies.

Notations: The lowercase boldface **a** represents a vector and the uppercase boldface **A** represents a matrix. Operators $(\cdot)^T, (\cdot)^{-1}$ and $(\cdot)^H$ stand for transpose, inverse, and Hermitian, respectively. The notation $|\cdot|$ denotes the modulus.

II. System Model

This section firstly introduces the adopted 3D staggered array geometry, followed by the communication system and channel models for both near-field and far-field[1] regions.

A. Proposed 3D Array Geometry

The topology of 3D antenna arrays constitutes the physical foundation for its performance, with different arrangements leading to significant variations in MC, effective aperture, and engineering complexity. To establish the array model for this study, we compare a traditional 2D array and two typical 3D array configurations, as illustrated in Fig. 1.

Conventional 2D arrays with antenna elements in the same plane, offer structural simplicity and ease of integration. However, their performance is limited by the finite physical aperture area, which imposes a theoretical upper bound on the system DoF. Moreover, as element spacing approaching half-wavelength, an increase in spatial correlation further limits the capacity and multiplexing capability of MIMO systems.

In contrast, the stacked 3D array (i.e., Type 2) employs vertical stacking to exploit the extra DoF without expanding the aperture. In this configuration, the two layers are vertically separated by a height h, which is a key design parameter affecting the effective projected aperture and beam directivity. The additional dimension enlarges the effective projected aperture under oblique incidence, enhancing spatial diversity and strengthening robustness under complex channel conditions, especially during wide-angle scanning [8]. However, it suffers from significant electromagnetic blockage, where upper elements shield lower ones, drastically degrading their radiation efficiency.

Furthermore, the Type 3 array improves upon Type 2 structure by horizontally offsetting the upper antenna layers to mitigate electromagnetic blockage and substantially improve the radiation performance of lower-layer elements. The horizontal offsets along the x-axis and y-axis, denoted as Δd_x and Δd_y respectively, are collectively represented as Δd in Fig. 1(c) for simplicity. While retaining the advantage of enlarged effective projected aperture, this design achieves lower spatial correlation and reduced MC due to the horizontal offsets compared to Type 2 configuration. From an implementation perspective, the Type 3 structure achieves a more effective balance between performance and engineering, thus offering greater practical value in real-world deployments.

Based on the foregoing analysis, the Type 3 structure is adopted in this study for subsequent codebook design and system performance evaluation, with the objective of fully leveraging its potential to enhance the DoF and spectral efficiency of MIMO systems.

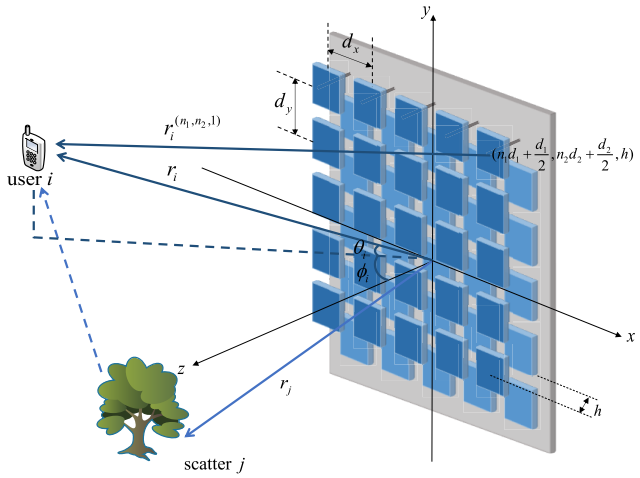


Fig. 2. Model of double-layer UPA systems.

B. Communication System Model for 3D Arrays

This work considers a downlink multi-user communication scenario as depicted in Fig. 2. The base station (BS) employs the Type 3 array with $N_x \times N_y \times 2$ antennas. Besides the conceptually theoretical 3D array, the practical structure comprising two stacked planar arrays is designed in [8], adopting an electromagnetic band-gap reflecting surface, as embraced by the application at BS side. Assuming single-antenna users,

[1]The "near-field" mentioned in this paper refers to radiative near-field, where reactive near-field is neglected. The boundary between the two regions is given by $d_r = 0.62\sqrt{\frac{D^3}{\lambda}}$, with D denoting the aperture area [40].

the channel between the BS and the k-th user is denoted as

$$\mathbf{h}_k = \sum_{l=0}^{L} g_l \mathbf{a}_l, \tag{1}$$

where $l = 0$ denotes the line-of-sight path, $l = 1, 2, \cdots, L$ represents the non-line-of-sight paths, and g_l is the complex gain of the l-th path. Notably, the corresponding array steering vector $\mathbf{a}_l$ depends on the specific electromagnetic field region, which significantly influences the channel model.

Rayleigh distance, defined as $d_F = \frac{2D^2}{\lambda}$, is employed to distinguish far-field from near-field [41]. In the far-field, where the electromagnetic wave approximates a plane wave, the steering matrix of the i-th path is defined as

$$\mathbf{A}_i^{\text{far}} = \frac{1}{\sqrt{2N_xN_y}} \begin{bmatrix} e^{j\Delta\varphi^{\text{far}}_{(i,1,1,0)}} \cdots & e^{j\Delta\varphi^{\text{far}}_{(i,N_x,1,0)}} \\ \vdots & \ddots & \vdots \\ e^{j\Delta\varphi^{\text{far}}_{(i,1,N_y,0)}} \cdots & e^{j\Delta\varphi^{\text{far}}_{(i,N_x,N_y,0)}} \end{bmatrix}, \tag{2}$$

where $\Delta\varphi^{\text{far}}_{(i,n_x,n_y,p)}$ denotes the phase difference of element (n_x, n_y) in the $(p+1)$-th layer relative to the reference element. With the center element of the first layer as the reference, it can be expressed as

$$\begin{aligned} &\Delta\varphi^{\text{far}}_{(i,n_x,n_y,p)} \\ &= k_c \left[(\tilde{n}_x d_x + \Delta d_x p)\cos\theta_i \sin\phi_i \right. \\ &\quad \left. + (\tilde{n}_y d_y + \Delta d_y p)\sin\theta_i + hp\cos\theta_i\cos\phi_i\right] \\ &= k_c[m_{x,p}\cos\theta_i\sin\phi_i + m_{y,p}\sin\theta_i + hp\cos\theta_i\cos\phi_i], \end{aligned} \tag{3}$$

where $k_c = \frac{2\pi}{\lambda}$ is wavenumber, with λ being the wavelength. The parameters d_x and d_y represent the spacing of adjacent elements of the same layer, while Δd_x and Δd_y denote the planar offsets between the two layers, which are vertically separated by h. The index $\tilde{n}_{\mathbf{i}} = \frac{2n_{\mathbf{i}} - N_{\mathbf{i}} - 1}{2}$, $\mathbf{i} = \{x, y\}$, $n_{\mathbf{i}} \in \{1, \cdots, N_{\mathbf{i}}\}$ and $N_{\mathbf{i}}$ is assumed to be odd.

Within the near-field region, the spherical wave propagation characteristics dominate, and the steering matrix is given by

$$\mathbf{A}_i^{\text{near}} = \frac{1}{\sqrt{2N_xN_y}} \begin{bmatrix} e^{-jk_c\left(r_i^{(1,1,p)} - r_i\right)} \cdots & e^{-jk_c\left(r_i^{(N_x,1,p)} - r_i\right)} \\ \vdots & \ddots & \vdots \\ e^{-jk_c\left(r_i^{(1,N_x,p)} - r_i\right)} \cdots & e^{-jk_c\left(r_i^{(N_x,N_y,p)} - r_i\right)} \end{bmatrix}, \tag{4}$$

where $r_i^{(n_x,n_y,p)} - r_i$ is the precise range difference between the element (n_x, n_y) in layer $p+1$ and the reference element for this path, determined by the array geometry, with $r_i^{(n_x,n_y,p)}$ defined as

$$\begin{aligned} r_i^{(n_x,n_y,p)} = &\left\{ (r_i\cos\theta_i\sin\phi_i - m_{x,p})^2 \right. \\ &\left. + (r_i\sin\theta_i - m_{y,p})^2 + (r_i\cos\theta_i\cos\phi_i - hp)^2 \right\}^{\frac{1}{2}}, \end{aligned} \tag{5}$$

and further simplification can be achieved via the Fresnel approximation [42], denoted by (6), as shown at the bottom of the page.

A hybrid precoding architecture is utilized to trade off system performance against hardware complexity [43]. This architecture separates the precoding into analog and digital beamforming, implemented by the codebook matrix $\mathbf{F}_{\text{RF}}$ and the baseband precoder $\mathbf{F}_{\text{BB}}$, respectively. The overall performance is evaluated by the spectral efficiency, expressed as

$$R = \sum_{k=1}^{K} \log_2 \left(1 + \frac{|\mathbf{h}_k \mathbf{P}_k|^2}{\sum_{j \neq k} |\mathbf{h}_k \mathbf{P}_j|^2 + \sigma_k^2} \right), \tag{7}$$

where K is the number of user, $\mathbf{P} = \mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}$, and σ_k^2 denotes the noise power at the k-th user. To simplify the analysis, the number of radio frequency (RF) chains $N_{\text{RF}} = K$ is assumed in this work.

III. Proposed Codebook Design Scheme

In this section, a codebook design scheme for the proposed 3D antenna array systems is presented, which aims to minimize inter-beam interference.[2] The high-dimensional joint problem is decoupled into two simpler subproblems, enabling independent sampling in angular and distance domains under a controlled coherence threshold to ensure effective beam orthogonality while maintaining a manageable codebook size.

A. Near-Field Codebook Design for 3D Arrays

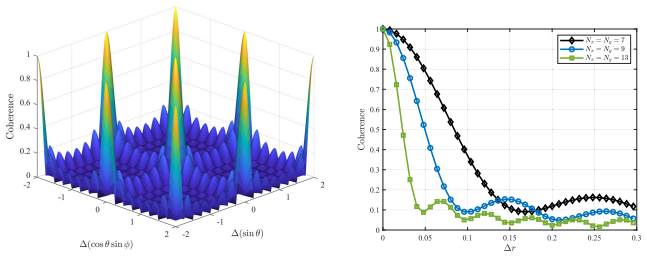
The above expressions (2) and (4) reveal a fundamental difference between near-field and far-field channels, manifested as a transition from purely angular to joint angular-range dependence. This distinction requires that efficient codebook designs be capable of adaptively generating the appropriate beam type based on the specific field region of users. Given the growing significance of near-field [44], [45], we first design its codebook to exploit the unique spatial DoF in this region.

The codebook design in this work is guided by the principle of minimizing mutual coherence among codewords, enabling precise beam focusing and effective user separation through nearly orthogonal beam patterns for enhanced spectral efficiency without additional overhead. Based on the vectorization of the steering matrix in (4), the coherence of the constructed 3D codebook is defined as

$$f(\mathbf{a}_a^{\text{near}}, \mathbf{a}_b^{\text{near}}) = \left| (\mathbf{a}_a^{\text{near}})^H \mathbf{a}_b^{\text{near}} \right|$$

[2]Note that we do not incorporate path-loss or shadowing explicitly here, as our primary emphasis is on the codebook design. Nevertheless, such effects can be superimposed in a real system.

$$\begin{aligned} r_i^{(n_x,n_y,p)} &\approx r_i - \left[m_{x,p}\cos\theta_i\sin\phi_i + m_{y,p}\sin\theta_i + hp\cos\theta_i\cos\phi_i\right] + \frac{1}{2r_i}\left[m_{x,p}^2\left(1-\cos^2\theta_i\sin^2\phi_i\right) + m_{y,p}^2\left(1-\sin^2\theta_i\right)\right. \\ &\left. + h^2p^2\left(1-\cos^2\theta_i\cos^2\phi_i\right) - 2m_{x,p}m_{y,p}\cos\theta_i\sin\theta_i\sin\phi_i - 2hpm_{x,p}\cos^2\theta_i\cos\phi_i\sin\phi_i - 2hpm_{y,p}\cos\theta_i\sin\theta_i\cos\phi_i\right] \\ &= r_i - g^{\text{angle}}_{\theta_i,\phi_i}(n_x,n_y,p) + \frac{1}{2}g^{\text{joint}}_{\theta_i,\phi_i,r_i}(n_x,n_y,p). \end{aligned} \tag{6}$$



(a) Angular coherence versus Δ_{angle}. (b) Range coherence versus Δr.

Fig. 3. Coherence with different sampling intervals.

$$= \left| \frac{1}{2N_xN_y} \sum_{p} \sum_{n_x} \sum_{n_y} e^{jk_c\left[\left(r_a^{(n_x,n_y,p)} - r_a\right) - \left(r_b^{(n_x,n_y,p)} - r_b\right)\right]} \right|$$

$$\approx \left| \frac{1}{2N_xN_y} \sum_{p} \sum_{n_x} \sum_{n_y} \underbrace{e^{jk_c\left[-g^{\text{angle}}_{\theta_a,\phi_a}(n_x,n_y,p) + g^{\text{angle}}_{\theta_b,\phi_b}(n_x,n_y,p)\right]}}_{f^{\text{angle}}_{\theta,\phi}(a,b)} \right.$$

$$\left. \cdot \underbrace{e^{j\frac{k_c}{2}\left[g^{\text{joint}}_{\theta_a,\phi_a,r_a}(n_x,n_y,p) - g^{\text{joint}}_{\theta_b,\phi_b,r_b}(n_x,n_y,p)\right]}}_{f^{\text{joint}}_{\theta,\phi,r}(a,b)} \right|. \tag{8}$$

This expression reveals that under the Fresnel approximation, the near-field coherence function can be represented as the product of two parts: 1) a linear component $f^{\text{angle}}_{\theta,\phi}(a,b)$, which captures the far-field-like phase behavior, 2) a quadratic component $f^{\text{joint}}_{\theta,\phi,r}(a,b)$, which embodies the distinctive near-field curvature effect, offering a clear physical interpretation.

Drawing on the idea of [37], this paper similarly decouples the joint optimization problem into two independent subproblems. Specifically, this framework achieves the sampling design for the angular domain and the distance domain through the following steps.

- **Angular-Domain Sampling**: For sampling points on the same range ring (where $g^{\text{joint}}_{\theta,\phi,r}(n_x,n_y,p) = \text{const}$), the quadratic term reduces to 1. The coherence function then collapses to $f^{\text{angle}}_{\theta,\phi}(a,b)$, transforming the minimization into a classical far-field problem and naturally yielding the angular sampling scheme.
- **Distance-Domain Sampling**: With a fixed sampling angle, the linear phase term $f^{\text{angle}}_{\theta,\phi}(a,b)$ vanishes. Consequently, coherence is governed solely by the quadratic term $f^{\text{joint}}_{\theta,\phi,r}(a,b)$, and analyzing it against a coherence threshold yields the required range sampling scheme.

The detailed analysis regarding the sampling of angles and distances is provided below.

1) Angular-Domain Sampling: Since $g^{\text{joint}}_{\theta,\phi,r}(n_x,n_y,p) = \text{const}$ is assumed, the coherence $f(\mathbf{a}_a^{\text{near}}, \mathbf{a}_b^{\text{near}})$ converts to

$$f(\mathbf{a}_a^{\text{near}}, \mathbf{a}_b^{\text{near}}) \approx \left| \frac{1}{2N_xN_y} \sum_{n_x} \sum_{n_y} \sum_{p} f^{\text{angle}}_{\theta,\phi}(a,b) \right|, \tag{9}$$

which can be further solved as a geometric summation problem presented in (10), as shown at th bottom of the next page. It can be observed that the coherence function in (10) comprises two sinc functions, revealing a periodic zero-coherence property which can be seen in Fig. 3 (a). Specifically, zeros of it occur at

$$\cos\theta_b \sin\phi_b - \cos\theta_a \sin\phi_b = \frac{l_x\lambda}{N_x d_x},$$
$$\sin\theta_b - \sin\theta_a = \frac{l_y\lambda}{N_y d_y}, \tag{11}$$

where $l_x, l_y \in \mathbb{Z} \setminus \{0\}$. However, the design of practical codebooks is inherently governed by a fundamental trade-off between codebook size, mutual coherence, and angular coverage. Sampling at the first null (i.e., $l_x \pm 1, l_y \pm 1$) minimizes interference at the expense of a linearly growing codebook cardinality, which becomes prohibitively complex for massive arrays. Conversely, a sparser angular sampling reduces codebook size but creates coverage gaps, adversely affecting transmission reliability.

Therefore, a superior and more practical approach is to find a suboptimal sampling interval under a codebook size constraint, which can be formulated as a structured optimization problem. Firstly, we define a feasible set $\mathcal{S}$ of sampling intervals (l_x, l_y) that simultaneously satisfy the coverage condition $\Omega_{\text{cover}}(l_x,l_y) \supseteq \Omega_{\text{req}}$ and the mutual coherence constraint δ, where $(l_x,l_y) \in \mathbb{R}^+ \times \mathbb{R}^+$. The objective is to identify the optimal interval (l_x^*, l_y^*) within $\mathcal{S}$ that maximizes the number of beams under the codebook size constraint.

Remark 1: Sampling offsets from the zero point yield a path-difference phase in the mutual coherence, scaled by the parameter h. While a larger h enhances beam directivity, it can also exacerbate sidelobes or coherent interference at non-zero points. Consequently, designing the angular sampling interval requires a careful balancing of the joint impact of h on beam shape and inter-beam coherence.

2) Distance-Domain Sampling: A similar assumption is employed in this part. Given that $f^{\text{angle}}_{\theta,\phi}(a,b) = 1$, which can be realized by performing the sampling at a fixed and identical angle (θ,ϕ), the coherence (8) is simplified to (12), as shown at the bottom of the next page, where $\Delta_r = 1/r_a - 1/r_b, C_p = h^2p^2\left(1-\cos^2\theta\cos^2\phi\right), A_{xx} = 1-\cos^2\theta\sin^2\phi, A_{yy} = 1-\sin^2\theta, A_{xy} = -2\cos\theta\sin\theta\sin\phi, B_x = -2hp\cos^2\theta\cos\phi\sin\phi$, and $B_y = -2hp\cos\theta\sin\theta\cos\phi$.

Since (12) involves a summation with a quadratic phase term, the Fresnel integral offers an ideal analytical framework [46]. Firstly, an integral substitution transforms (12) into

$$f(\mathbf{a}_a^{\text{near}}, \mathbf{a}_b^{\text{near}}) \approx \left| \frac{1}{2} \sum_{p=0}^{1} \int_{-L_x}^{L_x} \int_{-L_y}^{L_y} e^{j\frac{k_c}{2}\Delta_r Q(m_x,m_y,p)} dm_x dm_y \right|, \tag{13}$$

where $Q(m_x,m_y,p)$ is the phase as derived from the exponent in (12). To eliminate the intractable cross term $A_{xy}m_{x,p}m_{y,p}$, a rotation matrix $\mathbf{R}(\alpha)$ is introduced

$$\begin{bmatrix} u \\ v \end{bmatrix} = \mathbf{R}(\alpha) \begin{bmatrix} m_x \\ m_y \end{bmatrix} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} m_x \\ m_y \end{bmatrix}, \tag{14}$$

where α satisfies $\tan(2\alpha) = \frac{A_{xy}}{A_{xx}-A_{yy}}$, yielding

$$\begin{aligned} &A_{xx}m_x^2 + A_{yy}m_y^2 + A_{xy}m_x m_y \\ &= \frac{1}{2}\left[(A_{xx}+A_{yy}) + \sqrt{(A_{xx}-A_{yy})^2 + A_{xy}^2}\right]u^2 \\ &\quad + \frac{1}{2}\left[(A_{xx}+A_{yy}) - \sqrt{(A_{xx}-A_{yy})^2 + A_{xy}^2}\right]v^2 \\ &= \lambda_u u^2 + \lambda_v v^2. \end{aligned} \tag{15}$$

Meanwhile, the linear terms are transformed into $\beta_u = B_x\cos\alpha + B_y\sin\alpha, \beta_v = -B_x\sin\alpha + B_y\cos\alpha$, and C_p remains the same. Assuming $\Delta_r > 0$, the coherence expression then becomes

$$f(\mathbf{a}_a^{\text{near}}, \mathbf{a}_b^{\text{near}}) \approx \left|\frac{1}{2}\sum_{p=0}^{1} e^{j\frac{k_c}{2}\Delta_r C_p} I_u(\Delta_r, p) I_v(\Delta_r, p)\right|, \tag{16}$$

where

$$\begin{aligned} I_u(\Delta_r, p) &= \int_{u_{\min}}^{u_{\max}} e^{j\frac{k_c}{2}\Delta_r(\lambda_u u^2 + \beta_u u)} du \\ &= e^{-j\frac{k_c\Delta_r\beta_u^2}{8\lambda_u}}\sqrt{\frac{\pi}{\mu_u}}\left[\mathcal{F}\left(\sqrt{\frac{\mu_u}{\pi}}u'_{\max}\right) - \mathcal{F}\left(\sqrt{\frac{\mu_u}{\pi}}u'_{\min}\right)\right], \\ I_v(\Delta_r, p) &= \int_{v_{\min}}^{v_{\max}} e^{j\frac{k_c}{2}\Delta_r(\lambda_v v^2 + \beta_v v)} dv \\ &= e^{-j\frac{k_c\Delta_r\beta_u^2}{8\lambda_u}}\sqrt{\frac{\pi}{\mu_v}}\left[\mathcal{F}\left(\sqrt{\frac{\mu_v}{\pi}}v'_{\max}\right) - \mathcal{F}\left(\sqrt{\frac{\mu_v}{\pi}}v'_{\min}\right)\right], \end{aligned} \tag{17}$$

where $\mathcal{F}(x) = C(x) + jS(x), u'_{\max} = u_{\max} + \frac{\beta_u}{2\lambda_u}, u'_{\min} = u_{\min} + \frac{\beta_u}{2\lambda_u}, v'_{\max} = v_{\max} + \frac{\beta_v}{2\lambda_v}, v'_{\min} = v_{\min} + \frac{\beta_v}{2\lambda_v}, \mu_u = k_c\Delta_r\lambda_u$ and $\mu_v = k_c\Delta_r\lambda_v$.

The rotation of the coordinate system complicates the coherence function by transforming its integration region from a rectangle to a parallelogram. Nevertheless, the envelope still exhibits a decaying trend during the initial oscillation cycles, as shown in Fig. 3 (b), where $h = 0.5\lambda$. This observation motivates a distance-domain sampling scheme governed by a coherence threshold $\Delta_{\text{threshold}}$, which determines the minimum distance interval Δ_r^* satisfying $f(\Delta_r) = \Delta_{\text{threshold}}$ for distinguishable samples. The final sampling interval $\Delta_r \geq \Delta_r^*$ is then jointly optimized with respect to the codebook size and the coverage distance, defining the non-uniform sampling strategy of r.

Remark 2: The parameter h exerts a direct influence on the distance-domain sampling design. It manifests in both the constant and linear phase terms of the mutual coherence expressions (12), directly controlling range focusing and coherence decay. While a larger h sharpens range resolution, it exacerbates oscillations of coherence curve, complicating the codebook design.

Notably, the effect of layer spacing on coherence is non-monotonic. While increasing h narrows the beamwidth and enhances DoF, coherence is dominated by inter-layer phase interference. As a metric defined by the inner product between different codewords, it exhibits oscillatory behavior and strong coupling with parameters such as angle and distance, with no direct correlation between the two.

This decoupled optimization strategy tackles the high-dimensional joint design by solving two independent low-dimensional subproblems. It leverages the approximate orthogonality between angle and distance to independently

$$\begin{aligned} f(\mathbf{a}_a^{\text{near}}, \mathbf{a}_b^{\text{near}}) &= \left|\frac{1}{2N_xN_y}\sum_{n_x=1}^{N_x}\sum_{n_y=1}^{N_y}\sum_{p=0}^{1} e^{jk_c[m_{x,p}(\cos\theta_b\sin\phi_b - \cos\theta_a\sin\phi_a) + m_{y,p}(\sin\theta_b - \sin\theta_a) + hp(\cos\theta_b\cos\phi_b - \cos\theta_a\cos\phi_a)]}\right| \\ &= \left|\frac{1}{2N_xN_y}\sum_{n_x=-\tilde{N}_x}^{\tilde{N}_x}\sum_{n_y=-\tilde{N}_y}^{\tilde{N}_y} \underbrace{\left(1 + e^{jk_c[\Delta d_x(\cos\theta_b\sin\phi_b - \cos\theta_a\sin\phi_a) + \Delta d_y(\sin\theta_b - \sin\theta_a) + h(\cos\theta_b\cos\phi_b - \cos\theta_a\cos\phi_a)]}\right)}_{w} \right. \\ &\qquad \left. \times e^{jk_c[n_x d_x(\cos\theta_b\sin\phi_b - \cos\theta_a\sin\phi_a) + n_y d_y(\sin\theta_b - \sin\theta_a)]}\right| \\ &= \left|\frac{1}{2N_xN_y}(1+w) \times \frac{\sin\left(\frac{1}{2}N_x k_c d_x(\cos\theta_b\sin\phi_b - \cos\theta_a\sin\phi_a)\right)}{\sin\left(\frac{1}{2}k_c d_x(\cos\theta_b\sin\phi_b - \cos\theta_a\sin\phi_a)\right)} \times \frac{\sin\left(\frac{1}{2}N_x k_c d_x(\sin\theta_b - \sin\theta_a)\right)}{\sin\left(\frac{1}{2}k_c d_x(\sin\theta_b - \sin\theta_a)\right)}\right|. \end{aligned} \tag{10}$$

$$\begin{aligned} f(\mathbf{a}_a^{\text{near}}, \mathbf{a}_b^{\text{near}}) &\approx \left|\frac{1}{2N_xN_y}\sum_{p=0}^{1}\sum_{n_x=1}^{N_x}\sum_{n_y=1}^{N_y} f_{\theta,\phi,r}^{\text{joint}}(a,b)\right| \\ &= \left|\frac{1}{2N_xN_y}\sum_{p=0}^{1}\sum_{n_x=1}^{N_x}\sum_{n_y=1}^{N_y} e^{j\frac{k_c}{2}\left[\left(\frac{1}{r_a}-\frac{1}{r_b}\right)\left(m_{x,p}^2\left(1-\cos^2\theta\sin^2\phi\right) + m_{y,p}^2\left(1-\sin^2\theta\right) + h^2p^2\left(1-\cos^2\theta\cos^2\phi\right)\right)\right]}\right. \\ &\qquad \left. \times e^{j\frac{k_c}{2}\left[\left(\frac{1}{r_a}-\frac{1}{r_b}\right)\left(-2m_x m_y\cos\theta\sin\theta\sin\phi - 2m_x hp\cos^2\theta\cos\phi\sin\phi - 2m_y hp\cos\theta\sin\theta\cos\phi\right)\right]}\right| \\ &= \left|\frac{1}{2N_xN_y}\sum_{p=0}^{1}\sum_{n_x=1}^{N_x}\sum_{n_y=1}^{N_y} e^{j\frac{k_c}{2}\Delta_r\left(A_{xx}m_{x,p}^2 + A_{yy}m_{y,p}^2 + B_x m_{x,p} + B_y m_{y,p} + A_{xy}m_{x,p}m_{y,p} + C_p\right)}\right|. \end{aligned} \tag{12}$$

tailor the sampling in each dimension, which achieves low global mutual coherence through the maintenance of low intra-dimensional coherence, thus providing a theoretical foundation for the codebook performance.

B. Asymptotic Orthogonality of Codewords in 3D Arrays

The asymptotic orthogonality of the near-field codebook is crucial for minimizing inter-beam interference, enabling precise signal focusing and effective suppression of signal leakage. Leveraging the UPA coherence analysis in (8), we establish the following fundamental property for the proposed 3D array system.

Lemma 1: (Asymptotic Orthogonality of Near-Field Codebook:) In a double-layer UPA system, the near-field codewords become asymptotically orthogonal as the array aperture increases, provided that their designated locations differ in at least one spatial dimension (distance or angle), i.e.,

$$\lim_{N_x,N_y\to\infty}\left|(\mathbf{a}_i^{\text{near}})^H\mathbf{a}_j^{\text{near}}\right|=0,\ \text{for}\ (r_i,\theta_i,\phi_i)\neq(r_j,\theta_j,\phi_j). \tag{18}$$

Proof:: The proof is provided in Appendix.

A key insight from the proof is that the conventional far-field scenario characterized by planar wavefronts is a special case of the near-field model. Consequently, the asymptotic orthogonality in the far-field region is inherently guaranteed by Lemma 1, i.e., orthogonality holds for far-field codewords when $(\theta_i,\phi_i)\neq(\theta_j,\phi_j)$.

The additional distance domain in the near-field provides a new dimension for spatial multiplexing and interference suppression beyond angle-only far-field beams. This leads to superior performance in multi-user scenarios, including more precise multi-path resolution, more efficient spatial multiplexing, and a substantial reduction in inter-user interference, thereby boosting the overall system performance.

Remark 3: The analysis confirms that the layer spacing h does not impact the asymptotic orthogonality. This offers considerable flexibility for practical system design, allowing the layer spacing to be optimized for hardware or cost constraints without compromising interference management performance.

C. Far-Field Scenarios

Given the fundamental correspondence between the two domains, a near-field codebook can be inherently compatible with far-field scenarios. A comparison of the steering vectors in different regions, i.e., (3) and (6), reveals that the far-field phase difference corresponds to the angular component under the Fresnel approximation of the near-field one. Specifically, expanding the maximum designed sampling distance far beyond the Rayleigh distance suppresses the joint exponent component in (8), which enables the proposed near-field codebook applicable for the entire hybrid field. However, this versatility comes at the expense of an increased codebook size.

Consequently, our near-field-optimized codebook is also far-field-compatible. Its behavior gracefully converges to a standard far-field solution as $r\to\infty$, ensuring robust operation across all distances.

IV. Evanescent Codewords

To address the challenge of escalating codebook complexity, this section focuses on the elimination of evanescent codewords. Drawing on classical electromagnetic theory that describes evanescent waves as exhibiting exponential decay [47], [48], we identify and eliminate these contribution-negligible wave components from the codebook. This approach substantially reduces computational complexity while maintaining strong codebook performance [49], [50].

For an antenna element located at $\mathbf{r}=(x,y,z)$, the channel response to a electromagnetic wave arriving from $\hat{\mathbf{k}}(\theta,\phi)$ can be modeled as a phase delay term $\mathbf{k}\cdot\mathbf{r}$, where the wave vector $\mathbf{k}$ has a fixed magnitude $k_c=\frac{2\pi}{\lambda}$ in free space and a direction given by the propagation direction [51]. Hence, the channel response can be expressed as

$$h(\mathbf{r})\propto e^{-j\mathbf{k}\cdot\mathbf{r}}=e^{-j(k_x x+k_y y+k_z z)}. \tag{19}$$

Compared with (1), the amplitude variations are omitted in this model, while the primary emphasis is placed on the phase information that determines the beam direction. Such a simplification is common and valid for the analysis of beamforming performance.

Under the far-field assumption, the electromagnetic wave is well modeled as a plane wave, characterized by a planar wavefront with a constant direction (θ,ϕ) across the antenna array. Thus, its spherical coordinate components are given by

$$k_x=k_c\cos\theta\sin\phi,\ \ k_y=k_c\sin\theta. \tag{20}$$

It follows from (11), as shown at th bottom of the next page, that the optimal sampling points form a uniform grid in the $(\cos\theta\sin\phi,\sin\theta)$ domain for the standard sampling case where the codebook size is not considered,

$$u_m=m\frac{\lambda}{N_x d_x},v_n=n\frac{\lambda}{N_y d_y}. \tag{21}$$

The relation $k_x=k_c\cdot u$ and $k_y=k_c\cdot v$ translates the direction cosine samples (u_m,v_n) to the projected wavenumber components (k_x,k_y) on the array plane, defining the spatial frequency components for each codeword

$$k_x=\frac{2\pi m}{N_x d_x},k_y=\frac{2\pi n}{N_y d_y}. \tag{22}$$

To avoid aliasing, the indices (m,n) are normalized as $m'=m/N_x,n'=n/N_y$ and confined to $(-0.5,0.5]$. According to electromagnetic wave theory, a codeword is classified as evanescent if its indices fall outside the propagating regime defined by $k_t^2=k_x^2+k_y^2\le k_c^2$ [49], i.e.,

$$\left(\frac{m'}{\alpha_x}\right)^2+\left(\frac{n'}{\alpha_y}\right)^2>1, \tag{23}$$

where $\alpha_x=d_x/\lambda,\alpha_y=d_y/\lambda$.

Furthermore, the channel redundancy extends into the radiative near-field, with the channel exhibiting spherical wavefronts and spatial frequencies confined to the propagating spectrum ($|\kappa|\le k$) [50].

In near-field, the phase delay $\psi(x,y,z)$ is proportional to the propagation path length $r(x,y,z)$ between the antenna and

the user which is located at (ρ, θ, ϕ), i.e., $\psi(x, y, z) = k_c \cdot r(x, y, z)$, where

$$r(x,y,z) = \Big[(x-\rho\cos\theta\sin\phi)^2 + (y-\rho\sin\theta)^2 + (z-\rho\cos\theta\cos\phi)^2\Big]^{\frac{1}{2}}. \tag{24}$$

The spatial frequencies k_x, k_y are defined as the spatial gradients of the phase distribution $k_x = \frac{\partial\psi(x,y,z)}{\partial x}, k_y = \frac{\partial\psi(x,y,z)}{\partial y}$, yielding

$$k_x = \frac{k_c(x-\rho\cos\theta\sin\phi)}{\sqrt{(x-\rho\cos\theta\sin\phi)^2+(y-\rho\sin\theta)^2+(z-\rho\cos\theta\cos\phi)^2}},$$
$$k_y = \frac{k_c(y-\rho\sin\theta)}{\sqrt{(x-\rho\cos\theta\sin\phi)^2+(y-\rho\sin\theta)^2+(z-\rho\cos\theta\cos\phi)^2}}. \tag{25}$$

Based on the design criteria for near-field codebook, optimal sampling points constitute a uniform grid in the transformed $\left(\cos\theta\sin\phi, \sin\theta, \frac{1}{r}\right)$ domain,

$$u_m = m\frac{\lambda}{N_x d_x}, v_n = n\frac{\lambda}{N_y d_y}, w_l = \frac{r_{\max}}{1+l\Delta_r r_{\max}}, \tag{26}$$

where $r_{\max}$ represents the maximum sampling distance.

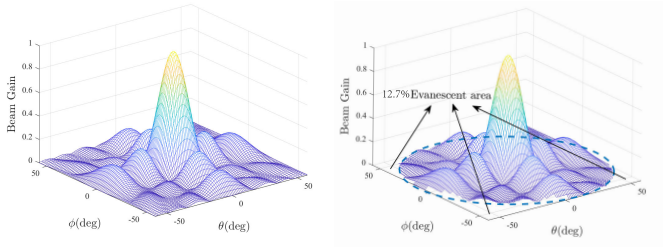
The mapping from (u, v, w) to (k_x, k_y) deviates from a simple linear relationship, specifically shown in (27), as shown at the bottom of the page. Analogous to operations in the far-field, the indices (m, n) are normalized to avoid spatial aliasing. Therefore, a codeword index in near-field is evanescent if it satisfies the inequality

$$\left(z-(l'+\alpha_z)^{-1}\sqrt{1-\frac{m'^2}{\alpha_x^2}-\frac{n'^2}{\alpha_y^2}}\right)^2 < 0, \tag{28}$$

where $x = m_{x,p}, y = m_{y,p}, z = hp, m' = m/N_x, n' = n/N_y, l' = l\Delta_r, \alpha_x = d_x/\lambda, \alpha_y = d_y/\lambda, \alpha_z = 1/r_{\max}$.

For conventional 2D arrays, inequality (28) holds if and only if (23) is satisfied, thus making the criteria for near-field and far-field evanescent codewords identical. In the case of the 3D arrays, the left-hand side of (28) can be complex when $z = h$, rendering a direct comparison meaningless. Nevertheless, this scenario requires the transverse wavenumbers k_x and k_y to be imaginary, which defines physically inadmissible codewords. Therefore, the applicable criterion is still equivalent to (23).

Remark 4: This analysis reveals that while 3D arrays provide enhanced DoF, the criterion for discriminating evanescent codewords remains identical to that for 2D arrays. This invariance stems from the fact that evanescent modes are solely determined by the transverse wavenumber exceeding the free-space wavenumber, which is independent of the vertical geometry of arrays. Thus the transverse wavenumber criterion extends directly to 3D configurations.



(a) Beam with evanescent codewords. (b) Beam w/o evanescent codewords.

Fig. 4. Comparison of original beam and beam without evanescent codewords.

Physically, the invariance of the evanescent criterion across 2D and 3D configurations stems from basic dispersion relation of free space, i.e., $k_x^2 + k_y^2 \le k^2$. The transition from a propagating wave to an evanescent one is strictly dictated by whether the transverse wavenumbers exceed the total wavenumber supported by the medium. While a 3D array introduces a vertical dimension $z = h$ that enables double-layer sampling of the field, it remains a passive observer of the wave's propagation state. If the phase gradient in the k_x-k_y plane defined by the beamforming direction is too steep to be supported as a traveling wave, the wave must decay along the z-axis to satisfy the Helmholtz equation, regardless of whether the antennas are arranged in a single layer or multiple layers. Thus, the 3D geometry enhances the DoFs in beam control but does not alter the intrinsic physical boundaries of the propagating regime.

The comparison of original beam and beam without evanescent codewords is shown in Fig. 4, with an antenna configuration of $N_x = N_y = 7$ and a pointing angle of $(0°, 0°)$. It can be seen that the evanescent codewords make a small contribution to wave formation, i.e., the codebook are generated within evanescent area as shown in Fig. 4 (b). Specifically, under this configuration, approximately 12.7% of the total codewords fall into this region. Thus, eliminating evanescent codewords in codebook construction reduces the complexity of it. Finally, the codebook designed will be stored in a lookup-table at the base station for key operations like beam scanning and channel estimation.

$$k_x = \frac{2\pi\left(x-\frac{r_{\max}}{1+l\Delta_r r_{\max}}\frac{m\lambda}{N_x d_x}\right)}{\lambda\sqrt{\left(x-\frac{r_{\max}}{1+l\Delta_r r_{\max}}\frac{m\lambda}{N_x d_x}\right)^2+\left(y-\frac{r_{\max}}{1+l\Delta_r r_{\max}}\frac{n\lambda}{N_y d_y}\right)^2+\left(z-\frac{r_{\max}}{1+l\Delta_r r_{\max}}\sqrt{1-\frac{m^2\lambda^2}{N_x^2 d_x^2}-\frac{n^2\lambda^2}{N_y^2 d_y^2}}\right)^2}},$$
$$k_y = \frac{2\pi\left(y-\frac{r_{\max}}{1+l\Delta_r r_{\max}}\frac{n\lambda}{N_y d_y}\right)}{\lambda\sqrt{\left(x-\frac{r_{\max}}{1+l\Delta_r r_{\max}}\frac{m\lambda}{N_x d_x}\right)^2+\left(y-\frac{r_{\max}}{1+l\Delta_r r_{\max}}\frac{n\lambda}{N_y d_y}\right)^2+\left(z-\frac{r_{\max}}{1+l\Delta_r r_{\max}}\sqrt{1-\frac{m^2\lambda^2}{N_x^2 d_x^2}-\frac{n^2\lambda^2}{N_y^2 d_y^2}}\right)^2}}, \tag{27}$$

While 2D and 3D arrays follow the same mathematical criterion for identifying evanescent wave codewords, a fundamental distinction exists in their physical interpretation and signal sensing capabilities. In the near-field, the volumetric arrangement of 3D arrays provides enhanced sampling of the wavenumber spectrum corresponding to spherical wavefronts. Specifically, for the same transverse sampling points (u_m, v_n) from $k_x^2 + k_y^2 \leq k^2$, antennas located at different layers perceive different wavenumbers, enabling more precise joint control over the angular and range dimensions of the beam. In the far field, active phase control along the z-direction enhances beamforming resolution and user separation in elevation, yielding narrower beamwidth and higher spatial multiplexing efficiency. This analysis reveals that the performance enhancement in 3D arrays originates from vertical control via layer spacing h, not evanescent mode utilization.

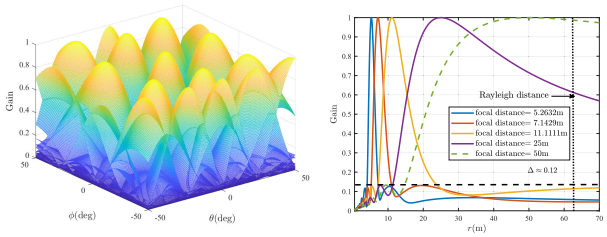
Notably, this work focuses on codebook design for communication systems, where evanescent codewords are deliberately omitted. However, such modes are essential for enhancing resolution and enabling subwavelength beam focusing [52], [53], [54], [55]. Observations from (27) reveal fundamental differences between near-field evanescent waves in 2D and 3D arrays. Particularly, the upper-layer antennas in a 3D array exhibit complex-valued components k_x, k_y and k_z corresponding to an near-field evanescent wave classified in a 2D array. i.e.,

$$e^{-jk_i i} = e^{-j\left(k_i' + jk_i''\right)i} = \underbrace{e^{-jk_i' i}}_{\text{oscillation}} \cdot \underbrace{e^{k_i'' i}}_{\text{attenuation}}, \tag{29}$$

where $i = \{x, y, z\}$. This indicates the presence of transversely decaying and longitudinally oscillatory evanescent waves of 3D arrays in near-field, a feature absent in 2D structures. On one hand, this structure imparts a confining attenuation to k_x and k_y, suppressing lateral energy spread. On the other hand, it grants k_z a tunable oscillatory component, breaking the exponential decay limit of 2D systems and enabling dynamic wavefront steering. These unique properties may present promising prospects for future investigations.

V. Numerical Results

This section presents a performance evaluation of the proposed codebook design scheme through simulations of beam and system performance based on the constructed 3D array models. Beam performance is evaluated by beam gain and beamwidth, where beam gain is defined as the magnitude of the correlation between the beamforming vectors at the target and non-target positions. System performance is assessed via spectral efficiency, the number of RF chains is set to $N_{\text{RF}} = K = 3$, and zero-forcing precoding is employed. The operating frequency is 1.5 GHz, the wavelength is $\lambda = 0.2$m, the antenna spacing is $d_x = d_y = \frac{\lambda}{2}$, with a offset of $\Delta d_x = d_x/2, \Delta_y = d_y/2$. Unless otherwise stated, the number of antennas in simulations is configured as $N_x = N_y = 9$. For clearer presentation of simulation results, different antenna configurations are employed in certain cases. Comparative fairness is ensured by keeping the total antenna elements consistent for all array configurations (2D and 3D) in the simulations.



(a) Angular domain. (b) Distance domain.

Fig. 5. Codebook scheme in angular and distance domain.

TABLE I
Comparison of Near-Field and Far-Field Codebook

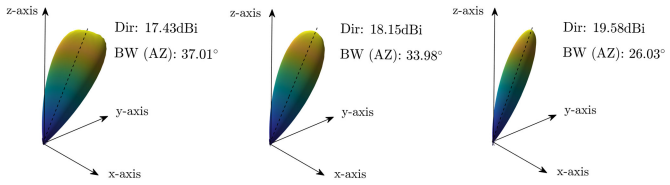
Comparison	Directivity	BW(EL)	BW(AZ)
near-field codebook ($r = d_F$)	20.347dBi	23.59°	12.75°
near-field codebook ($r = \frac{d_F}{3}$)	19.799dBi	23.88°	13.88°
far-field codebook	20.400dBi	23.6°	12.62°

A. Codebook Design Scheme

The performance of generated beams in angular and distance domains is illustrated in Fig. 5 (a) and Fig. 5 (b), respectively. In Fig. 5 (a), the antenna number is $N_x = N_y = 5$ and the user distance is set to the Rayleigh distance $d_F \approx 2.5$m. The angular sampling scheme demonstrates excellent differentiation at each sampled angle, which indicates that the proposed codebook scheme effectively supports angular discrimination in communication systems. This performance allows for the assignment of minimally interfering beams to users at different angles, leading to enhanced signal-to-interference-plus-noise ratio for each user.

In Fig. 5 (b), the antenna configuration is set with $N_x = N_y = 25$, the layer spacing is $h = \frac{\lambda}{2}$, and the angles are fixed at $\theta = 0°$ and $\phi = 0°$, where the radiative near-field interval is $[5\text{m}, 62.5\text{m}]$. It can be shown that the correlation between adjacent beams maintains a relatively low interference level (around 0.12), exhibiting strong discrimination in the distance domain, but with the beam diverging as the distance extends. This observation indicates that the proposed codebook scheme achieves beam focusing but approximately fails beyond the half of near-field region. The variations of beam patterns demonstrate a seamless transition of the proposed near-field codebook across different regions. As distance decreases, its focusing traits become more prominent, while a growing distance leads to more pronounced steering characteristics. This clarifies why the focusing behaviors are not ubiquitous throughout the near-field region, as there is a gradual transition towards beam steering instead of an abrupt change.

Moreover, the beamwidth and directivity of angular-domain beam of the far-field codebook are compared with those of near-field codebook at the Rayleigh distance and one-third of Rayleigh distance in Table I, with $h = 0.5\lambda, \theta = 40°, \phi = 40°$.



(a) 2D baseline. (b) $h = 0.5\lambda$. (c) $h = \lambda$.

Fig. 6. Angular domain beams with different layer spacings.

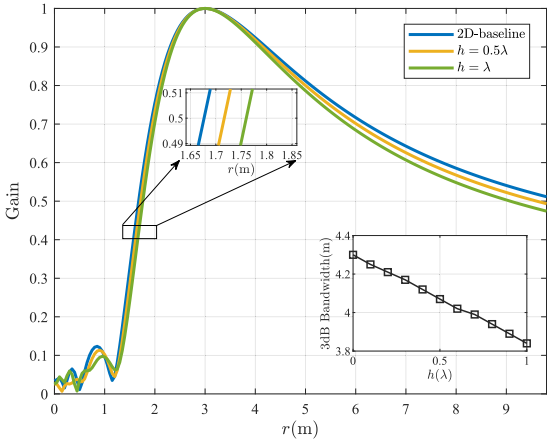


Fig. 7. Distance domain beams and bandwidth with varying layer spacings.

BW(EL) and BW(AZ) in Table I refer to elevation beamwidth and azimuth beamwidth, respectively. The data reveals that at Rayleigh distance, the pointing angle, beamwidth and directivity are nearly identical, confirming the feasibility of the proposed codebook for serving users across the entire spatial domain. However, at one-third of the Rayleigh distance, the beam experiences a performance degradation due to the limited availability of resources, with most having been allocated to achieve beam focusing in the distance domain.

B. Impact of Layer Spacing

In 3D antenna arrays, the layer spacing h is a key parameter distinguishing 3D arrays from 2D ones and is crucial for measuring the DoF in the vertical dimension. In this subsection, we analyze the impact of layer spacing h on the performance of codebook.[3]

Since achieving a separation of λ distance in real antenna arrays may be challenging when λ is large, the analysis of h is constrained to the range from 0 to λ.[4]

Fig. 6 (a), (b) and (c) illustrate the angular domain beams with $h = \{0, 0.5\lambda, \lambda\}$, respectively. The target angle of $\theta = 60°, \phi = 90°$. It is evident that beamwidth decreases and directivity strengthens with the increase of layer spacing, where Dir in the figure denotes directivity. For example, a 29% reduction in beamwidth and a roughly 3dBi increase in directivity are observed. The distance domain beam patterns for h of $\{0, 0.5\lambda, \lambda\}$ are presented in Fig. 7, with the angles

[3]Note that $h = 0$ corresponds to the single-layer 2D array case, it serves as the baseline for fair comparison with 3D arrays of varying layer spacings.

[4]If high-frequency electromagnetic waves are employed, separations beyond λ might be possible, yet it remains challenging in practical engineering.

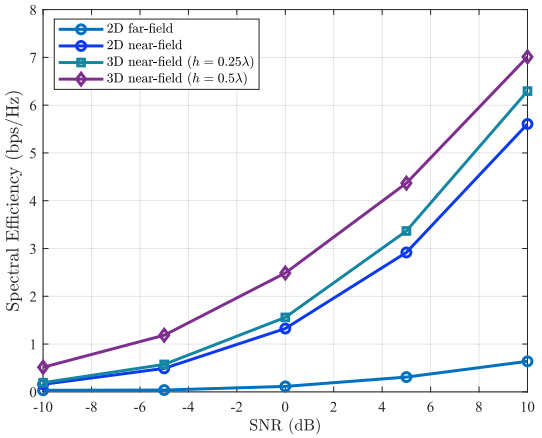


Fig. 8. Spectral efficiency comparison versus SNR with different codebooks.

$\theta = 0°, \phi = 0°$ and the focal distance $r = 3$m. It is observed that as h increases, the beam becomes narrower and the calculated 3dB bandwidth is decreased by approximately 12% over the range of layer spacing from 0 to λ. Furthermore, Fig. 8 provides a system-level performance evaluation of the codebook, comparing the spectral efficiency over a distance range of $[1, 6.56]$m. The near-field codebook exhibits a superior performance in the near-field region, which is strongly supported by the findings in [37]. Additionally, a consistent enhancement in spectral efficiency is observed with increasing layer spacing. Specifically, at a 10 dB SNR, the configuration with $h = 0.5\lambda$ yields a performance gain of approximately 27% compared to the case of $h = 0$.

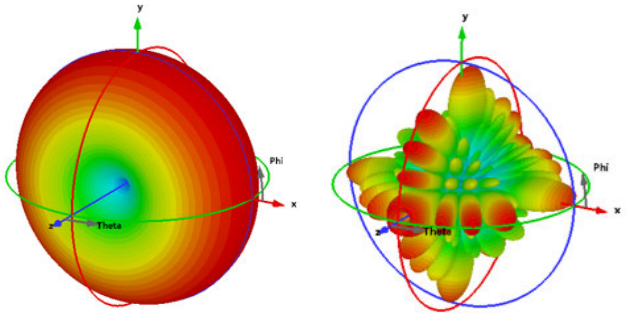
These results collectively demonstrate that increasing the layer spacing enables more precise beam targeting with reduced energy leakage to other channels, thereby effectively enhancing overall system performance.

C. Effects of Mutual Coupling

In this subsection, the impact of MC is investigated. Attributing to the distorted radiation pattern and reduced radiation efficiency in the closely spaced antennas, the resulted MC effect cannot be ignored in the codebook design [56], [57]. Thus, for practical applications and system optimization, it is essential to incorporate MC into codebook design [58].

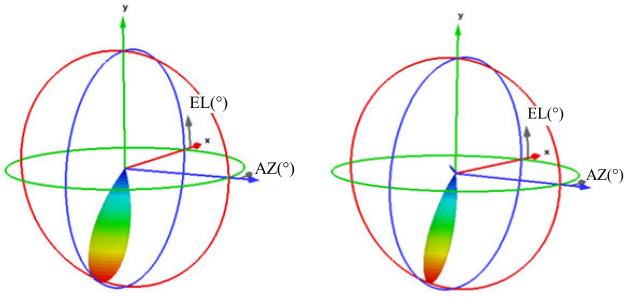
Specifically, the electric field with MC, $\mathbf{E}^{\text{MC}}$, is obtained via full-wave simulation of the array configuration, which accounts for the electromagnetic interactions between elements. In contrast, the ideal field $\mathbf{E}^{\text{Ideal}}$ is obtained by replicating the radiation field of a single isolated element at each array position and superimposing them according to the array geometry, thereby forming a coupling-free benchmark. Based on these two fields, the MC matrix can be calculated as $\mathbf{C} = \mathbf{E}^{\text{MC}}(\mathbf{E}^{\text{Ideal}})^{-1}$ [59].

Fig. 9 (a) presents the beam pattern of the adopted antenna while Fig. 9 (b) shows the pattern of the whole array with $h = \lambda$ without a codebook. The beam patterns of the designed codebook with $h = 0$ and λ are depicted in Fig. 9 (c) and Fig. 9 (d) when focal distance is set $r = 3$m, and target angles are $\theta = 60°, \phi = 90°$. The dipole antenna is printed on a substrate with dimensions of 12.2mm × 78mm while each



(a) Beam pattern of a single antenna.

(b) Beam pattern of the whole array.



(c) Front view of the beam with $h = 0$.

(d) Front view of the beam with $h = \lambda$.

Fig. 9. Beam pattern with MC in different cases.

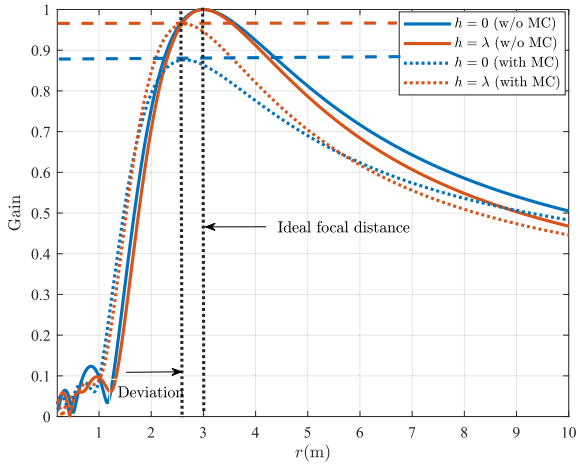


Fig. 10. Gain of beam with and w/o MC versus varying layer spacings.

antenna element sized at $1\text{mm} \times 75\text{mm}$. It can be seen that the desired angle is misaligned due to MC, which manifests as a 1° azimuth and a 4° elevation deviation when the layer spacing is $h = 0$, while the azimuth deviation becomes 3° when $h = \lambda$. Moreover, even in the presence of MC, increasing the layer spacing from 0 to λ still results in a narrower beamwidth and higher directivity, achieving a beamwidth reduction of around 34% and a directivity improvement of about 23%. While Fig. 10 presents the beam gain comparison for cases $h = \{0, \lambda\}$ with and without MC. It is observed that although the focal distance is initially set to $r = 3\text{m}$, integrating the

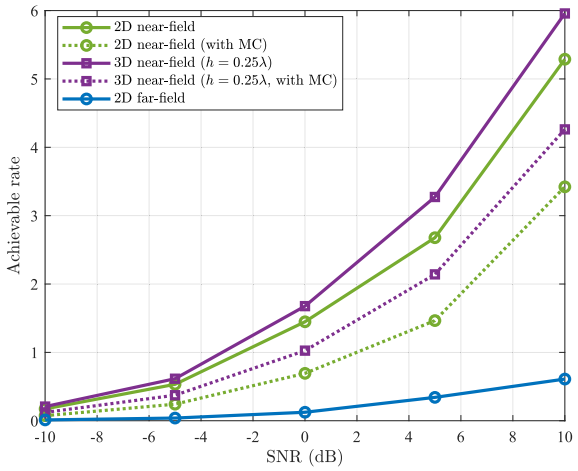


Fig. 11. Spectral efficiency versus SNR with different codebooks considering MC.

MC matrix into the codebook shifts the focal distance to approximately 2.6m with reduced beam gain. As illustrated in Fig. 11, the spectral efficiency is presented for both the MC-aware and ideal scenarios, with a distance range of $[1, 6.56]\text{m}$ and layer spacings of $h = \{0, 0.25\lambda\}$. At 10 dB SNR, the spectral efficiency for $h = 0.25\lambda$ exhibits a reduction of approximately 28.5% compared to the ideal case, while the 2D case experiences a decrease of about 35.3%.

The simulation results reveal several notable trends, such as the affected beam gain and wider bandwidth due to the consideration of MC. While an increase in h decreases beamwidth under both circumstances, the introduction of MC exacerbates the beamwidth gap between different layer spacings. This can be attributed to that the increase of h partly relieve the MC effect, leading to a smaller degradation in spectral efficiency.

VI. Conclusion

In this paper, a novel codebook design scheme based on a double-layer 3D antenna array is proposed. The scheme aims to minimize the coherence among codewords, which enhances the overall performance of the communication system. Although the codebook is primarily designed for near-field scenarios, far-field users can also be effectively served with appropriate adjustments to distance parameters. Additionally, the evanescent codewords are derived to address the issue of codebook expansion, following an analysis which reveals the unique wavenumber properties in such 3D structures. Simulation results show narrower angular and distance beamwidth, and a superior spectral efficiency in near-field scenarios compared to 2D arrays, leading to enhanced overall system performance. Furthermore, practical MC inducing beam deviation and broadening is considered.

Future endeavors will delve into joint codebook design that actively compensates MC and other practical problems. Specifically, the MC matrix obtained via our electric-field inversion method, as well as through impedance or S-parameter measurements, could provide a foundation for future codebook optimization. Potential directions include the

pre-compensation of ideal steering vectors to recover beam directivity and the integration of coupling-aware constraints into optimization algorithms to preserve beam orthogonality. Additionally, recognizing the importance of layer spacing for performance improvement, potential methods to increase layer spacing in actual 3D array structures will be explored. These research directions are anticipated to propel the continuous evolution of MIMO communication systems.

APPENDIX
PROOF OF THE LEMMA 1

Through variable substitution, the coherence of a double-layer 3D array defined in (8) can be simplified as

$$\begin{aligned}&\lim_{N_x,N_y\to\infty} f(\mathbf{a}_i^{\text{near}},\mathbf{a}_j^{\text{near}})\\ &\approx \lim_{N_x,N_y\to\infty}\frac{1}{2N_xN_y}\left|\sum_{n_x}\sum_{n_y}(1+\Gamma)\cdot e^{j(\xi_1 n_x+\xi_2 n_x^2)}\right.\\ &\quad\left.\times e^{j(\xi_3 n_y+\xi_4 n_y^2+\xi_5 n_x n_y)}\right|, \end{aligned}\tag{30}$$

where Γ and ξ_i $(i\in\{1,2,\cdots,5\})$ are constants depending on the spatial locations and array parameters. Therefore, the proof the convergence reduces to

$$\begin{aligned}&\lim_{N_x,N_y\to\infty} f(\mathbf{a}_i^{\text{near}},\mathbf{a}_j^{\text{near}}) \simeq \lim_{N_x,N_y\to\infty}\frac{1}{N_xN_y}\\ &\left|\sum_{n_x}\sum_{n_y} e^{j(\xi_1 n_x+\xi_2 n_x^2+\xi_3 n_y+\xi_4 n_y^2+\xi_5 n_x n_y)}\right|\\ &\overset{(a)}{\approx} \lim_{N_x,N_y\to\infty}\frac{1}{N_xN_y}\left|\int_{-\tilde{N}_x}^{\tilde{N}_x}\int_{-\tilde{N}_y}^{\tilde{N}_y} e^{j(\xi_2 n_x^2+\xi_4 n_y^2+\xi_5 n_x n_y)}\right.\\ &\left.\times e^{j(\xi_1 n_x+\xi_3 n_y)}dn_x dn_y\right|, \end{aligned}\tag{31}$$

where approximation (a) is obtained by replacing the summation symbol with an integral. To eliminate the cross-term, $\alpha=\frac{\xi_5}{2\xi_2}$ and $\beta=\frac{\xi_4}{\xi_2}-\frac{\xi_5^2}{4\xi_2^2}$ are set, then we can obtain

$$\begin{aligned}&\lim_{N_x,N_y\to\infty} f(\mathbf{a}_i^{\text{near}},\mathbf{a}_j^{\text{near}})\\ &\approx \lim_{N_x,N_y\to\infty}\frac{1}{N_xN_y}\left|\int_{-\tilde{N}_x}^{\tilde{N}_x}\int_{-\tilde{N}_y}^{\tilde{N}_y} e^{j(\xi_2[(n_x+\alpha n_y)^2+\beta n_y^2]}\right.\\ &\qquad\left.\times\, e^{j\xi_1 n_x+\xi_3 n_y)}dn_x dn_y\right|\\ &\overset{(b)}{=} \lim_{N_x\to\infty}\frac{1}{N_xN_y}\left|\int_{\overline{N}_{-y}}^{\overline{N}_y}\left[\int_{\overline{N}_{-x}}^{\overline{N}_x} e^{j\xi_2 t_1^2}dt_1\right] e^{j\xi_2\beta t_2^2}dt_2\right|. \end{aligned}\tag{32}$$

where $t_1=n_x+\alpha n_y+\sqrt{\frac{\xi_1}{4\xi_3}}, t_2=n_y+\sqrt{\frac{\xi_3-\xi_1\alpha}{2\xi_2\beta}}, \overline{N}_{-x}=-\tilde{N}_x+\alpha t_2+\sqrt{\frac{\xi_1}{4\xi_2}}, \overline{N}_x=\tilde{N}_x+\alpha t_2+\sqrt{\frac{\xi_1}{4\xi_2}}, \overline{N}_{-y}=-\tilde{N}_y+\sqrt{\frac{\xi_3-\xi_1\alpha}{2\xi_2\beta}}, \overline{N}_y=\tilde{N}_y+\sqrt{\frac{\xi_3-\xi_1\alpha}{2\xi_2\beta}}$. Step (b) is derived by setting $\alpha=\frac{\xi_5}{2\xi_2}$ and $\beta=\frac{\xi_4}{\xi_2}-\frac{\xi_5^2}{4\xi_2^2}$. Considering the Fresnel functions $C(x)=\int_0^x\cos\left(\frac{\pi}{2}t^2\right)dt, S(x)=\int_0^x\sin\left(\frac{\pi}{2}t^2\right)dt$, coherence can be approximated as

$$\begin{aligned}&\lim_{N_x,N_y\to\infty} f(\mathbf{a}_i^{\text{near}},\mathbf{a}_j^{\text{near}})\\ &\approx \lim_{N_x,N_y\to\infty}\frac{1}{N_xN_y}\left|\int_{\overline{N}_{-y}}^{\overline{N}_y}\left[\int_{-\infty}^{\infty} e^{j\xi_2 t_1^2}dt_1\right] e^{j\xi_2\beta t_2^2}dt_2\right|.\\ &= \lim_{N_x,N_y\to\infty}\frac{1}{N_xN_y}\left|\int_{-\infty}^{\infty}\sqrt{\frac{\pi}{2\xi_2}}(1+j)e^{j\xi_2\beta t_2^2}dt_2\right|.\\ &= \lim_{N_x,N_y\to\infty}\frac{1}{N_xN_y}\frac{\pi}{\xi_2}\sqrt{\frac{1}{\beta}}. \end{aligned}\tag{33}$$

If $\xi_2=\frac{1}{2r_i}\left(1-\cos^2\theta_i\sin^2\phi_i\right)-\frac{1}{2r_j}\left(1-\cos^2\theta_j\sin^2\phi_j\right)\neq 0$, (33) converges to 0.

Otherwise, if $x_2=0$, since ξ_2,ξ_4 and ξ_5 cannot be zero simultaneously, where $\xi_4=\frac{1}{2r_i}\left(1-\sin^2\theta_i\right)-\frac{1}{2r_j}\left(1-\sin^2\theta_j\right), \xi_5=-\frac{1}{r_i}(\cos\theta_i\sin\theta_i\sin\phi_i)+\frac{1}{r_j}(\cos\theta_j\sin\theta_j\sin\phi_j)$, we define

$$\begin{aligned}&I(N_x,N_y)\\ &=\left|\int_{-\tilde{N}_x}^{\tilde{N}_x}\int_{-\tilde{N}_y}^{\tilde{N}_y} e^{j(\xi_x n_y+\xi_y n_x)}e^{j(\xi_x n_x+\xi_y n_y)}dn_x dn_y\right|,\\ &\lim_{N_x,N_y\to\infty} f(\mathbf{a}_i^{\text{near}},\mathbf{a}_j^{\text{near}})=\lim_{N_x,N_y\to\infty}\frac{1}{N_xN_y}I(N_x,N_y). \end{aligned}\tag{34}$$

and the following is a case-by-case analysis.

A. $\xi_5=0,\xi_4\neq 0$

First, integrating over n_x by separating the cases $\xi_1=0$ and $\xi_1\neq 0$, which yields

$$\begin{aligned}&I(N_x,N_y)\\ &=\begin{cases}\left|\dfrac{2\sin(\xi_1\tilde{N}_x)}{\xi_1}\displaystyle\int_{-\tilde{N}_y}^{\tilde{N}_y} e^{j(\xi_4 n_y^2+\xi_3 n_y)}dn_y\right| & \text{if } \xi_1\neq 0,\\ \left|2N_x\displaystyle\int_{-\tilde{N}_y}^{\tilde{N}_y} e^{j(\xi_4 n_y^2+\xi_3 n_y)}dn_y\right| & \text{if } \xi_1=0.\end{cases} \end{aligned}\tag{35}$$

The integral term $\int_{-\tilde{N}_y}^{\tilde{N}_y} e^{j(\xi_4 n_y^2+\xi_3 n_y)}dn_y$ is of the Fresnel type and can be evaluated via the variable substitution $t=n_y+\sqrt{\frac{\xi_3}{2\xi_4}}$, leading to the following result

$$\begin{aligned}&\lim_{N_x,N_y\to\infty} f(\mathbf{a}_i^{\text{near}},\mathbf{a}_j^{\text{near}})\\ &\overset{(c)}{\leq}\begin{cases}\dfrac{2}{|\xi_1|N_xN_y}\cdot\left|\sqrt{\dfrac{\pi}{2\xi_4}}(1+j)\right| & \text{if } \xi_1\neq 0,\\ \dfrac{N_y}{2}\cdot\left|\sqrt{\dfrac{\pi}{2\xi_4}}(1+j)\right| & \text{if } \xi_1=0,\end{cases}\\ &=0, \end{aligned}\tag{36}$$

where the inequality (c) is derived from the bounds $\left|\sin\left(\xi_1\tilde{N}_x\right)\right|\leq 1$ and $|a\cdot b|\leq|a|\cdot|b|$.

B. $\xi_5\neq 0$

First, integrating over n_x, then we have

$$\begin{aligned}&I(N_x,N_y)\\ &=\left|\int_{-\tilde{N}_y}^{\tilde{N}_y}\left[\int_{-\tilde{N}_x}^{\tilde{N}_x} e^{jn_x(\xi_5 n_y+\xi_1)}dn_x\right]e^{j\left(\xi_4 n_y^2+\xi_3 n_y\right)}dn_y\right|\\ &=\left|\int_{-\tilde{N}_y}^{\tilde{N}_y}\left[\frac{2\sin(N_x(\xi_5 n_y+\xi_1))}{\xi_5 n_y+\xi_1}\right]e^{j\left(\xi_4 n_y^2+\xi_3 n_y\right)}dn_y\right|. \end{aligned}\tag{37}$$

1) $\xi_4 = 0$: In the case of $\xi_4 = 0$, we perform a change of variables $v = \xi_5 n_y + \xi_1$, yielding $dv = \xi_5 dn_y$, then we have

$$I(N_x, N_y) = \left| \frac{1}{\xi_5} e^{-j\frac{\xi_3}{\xi_5}} \int_a^b e^{j\frac{\xi_3}{\xi_5}u} \frac{2\sin(N_x u)}{u} du \right|, \quad (38)$$

where $a = -\xi_5 \tilde{N}_y + \xi_1$, $b = \xi_5 \tilde{N}_y + \xi_1$. Given that when $N_x \to \infty$, $\int_a^b e^{j\frac{\xi_3}{\xi_5}u} \frac{\sin(N_x u)}{u} du \to \pi$, then it follows that

$$\lim_{N_x, N_y \to \infty} f(\mathbf{a}_i^{\text{near}}, \mathbf{a}_j^{\text{near}}) \approx \lim_{N_x, N_y \to \infty} \frac{2\pi}{|\xi_5| N_x N_y} = 0. \quad (39)$$

2) $\xi_4 \neq 0$: Similarly, $v = \xi_5 n_y + \xi_1$ is set, then (37) converts to

$$\begin{aligned} I(N_x, N_y) &= \left| \frac{1}{\xi_5} \int_a^b e^{j\left(\xi_4\left(\frac{u-\xi_1}{\xi_5}\right)^2 + \xi_3\left(\frac{u-\xi_1}{\xi_5}\right)\right)} \frac{2\sin(N_x u)}{u} du \right| \\ &= \left| \frac{1}{\xi_5} \int_a^b e^{j\alpha(u)} \frac{2\sin(N_x u)}{u} du \right| \\ &\overset{(d)}{\leq} \int_a^b \left| \frac{1}{\xi_5} e^{j\alpha(u)} \frac{2\sin(N_x u)}{u} \right| du \\ &= \int_a^b \left| \frac{1}{\xi_5} \frac{2\sin(N_x u)}{u} \right| du \\ &\overset{(e)}{\leq} \int_a^b \left| \frac{1}{\xi_5} \frac{2}{u} \right| du, \end{aligned} \quad (40)$$

where step (d) is realized by $\left|\int f(u)du\right| \leq \int |f(u)|\, du$ and (e) is derived from $|\sin(N_x u)| \leq 1$. Considering

$$\begin{aligned} \int_a^b \frac{1}{|u|} du &= \ln\left((\xi_5 N_y)^2 - \xi_1^2\right) \\ &= 2\ln(\xi_5 N_y) + \ln\left(1 - \frac{\xi_1^2}{(\xi_5 N_y)^2}\right), \end{aligned} \quad (41)$$

then the coherence function can be simplified as

$$\begin{aligned} &\lim_{N_x, N_y \to \infty} f(\mathbf{a}_i^{\text{near}}, \mathbf{a}_j^{\text{near}}) \\ &= \lim_{N_x, N_y \to \infty} \frac{1}{N_x N_y} I(N_x . N_y) \\ &\leq \lim_{N_x, N_y \to \infty} \frac{1}{N_x N_y} \left| \frac{2}{\xi_5} \right| \\ &\quad \cdot \left(2\ln\left(|\xi_5 N_y|\right) + \ln\left(1 - \frac{\xi_1^2}{(\xi_5 N_y)^2}\right)\right), \end{aligned} \quad (42)$$

Given that $\frac{\ln|\xi_5| N_y}{N_x N_y} = \frac{1}{N_x} \cdot \frac{\ln(|\xi_5| N_y)}{N_y}$, $\frac{\ln|\xi_5| N_y}{N_x N_y}$ goes to $\frac{1}{N_x} \cdot 0 = 0$ when $N_x, N_y \to \infty$, (42) converges to 0.

Therefore, Lemma 1 holds universally. This ensures that as the array dimensions grow, different codewords achieve orthogonality, a vital feature for reducing interference and improving direction-of-arrival estimation accuracy in double-layer 3D systems.

REFERENCES

[1] T. L. Marzetta, "Massive MIMO: An introduction," *Bell Labs Tech. J.*, vol. 20, pp. 11–22, 2015.

[2] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bolcskei, "An overview of MIMO communications—A key to gigabit wireless," *Proc. IEEE*, vol. 92, no. 2, pp. 198–218, Feb. 2004.

[3] K. Zheng, L. Zhao, J. Mei, B. Shao, W. Xiang, and L. Hanzo, "Survey of large-scale MIMO systems," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 3, pp. 1738–1760, 3rd Quart., 2015.

[4] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5G wireless networks: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1617–1655, 3rd Quart., 2016.

[5] F. Lara, D. Altamirano, A. R. Narváez, F. Grijalva, and H. B. Molina, "Interference reduction factor evaluation for cylindrical and parallelepiped antenna arrays in massive MIMO systems," in *Proc. IEEE Colombian Conf. Commun. Comput. (COLCOM)*, Aug. 2024, pp. 1–5.

[6] C. A. Balanis, *Antenna Theory: Analysis and Design*, 4th ed., Hoboken, NJ, USA: Wiley, 2016.

[7] P. Hannan, "The element-gain paradox for a phased-array antenna," *IEEE Trans. Antennas Propag.*, vol. AP-12, no. 4, pp. 423–433, Jul. 1964.

[8] S. S. A. Yuan et al., "Breaking the degrees-of-freedom limit of holographic MIMO communications: A 3-D antenna array topology," *IEEE Trans. Veh. Technol.*, vol. 73, no. 8, pp. 11276–11288, Aug. 2024.

[9] S. S. A. Yuan and W. E. I. Sha, "Exceeding Hannan's limit by a 3-D antenna array," in *Proc. IEEE INC-USNC-URSI Radio Sci. Meeting (Joint AP-S Symp.)*, Florence, Italy, Jul. 2024, p. 182.

[10] R. Ji et al., "Exploring Hannan limitation for 3D antenna array," 2024, *arXiv:2409.01566*.

[11] Y. Wang, J. Li, L. Huang, Y. Jing, A. Georgakopoulos, and P. Demestichas, "5G mobile: Spectrum broadening to higher-frequency bands to support high data rates," *IEEE Veh. Technol. Mag.*, vol. 9, no. 3, pp. 39–46, Sep. 2014.

[12] W. Liu, C. Pan, H. Ren, J. Wang, and R. Schober, "Near-field multiuser beam-training for extremely large-scale MIMO systems," *IEEE Trans. Commun.*, vol. 73, no. 4, pp. 2663–2679, Apr. 2025.

[13] C. Weng, X. Guo, and Y. Wang, "Near-field beam training with hierarchical codebook: Two-stage learning-based approach," *IEEE Trans. Veh. Technol.*, vol. 73, no. 9, pp. 14003–14008, Sep. 2024.

[14] J. Mo et al., "Beam codebook design for 5G mmWave terminals," *IEEE Access*, vol. 7, pp. 98387–98404, 2019.

[15] S. Mabrouki, I. Dayoub, Q. Li, and M. Berbineau, "Codebook designs for millimeter-wave communication systems in both low- and high-mobility: Achievements and challenges," *IEEE Access*, vol. 10, pp. 25786–25810, 2022.

[16] D. Yang, L.-L. Yang, and L. Hanzo, "DFT-based beamforming weight-vector codebook design for spatially correlated channels in the unitary precoding aided multiuser downlink," in *Proc. IEEE Int. Conf. Commun.*, May 2010, pp. 1–5.

[17] Y. T. Wu, Y. Y. Zhao, and F. Yu, "Comparison of codebooks for beamforming in limited feedback MIMO systems," in *Proc. IEEE Int. Conf. Comput. Sci. Autom. Eng. (CSAE)*, May 2012, pp. 32–36.

[18] T. Kim, J. Park, J.-Y. Seol, S. Jeong, J. Cho, and W. Roh, "Tens of Gbps support with mmWave beamforming systems for next generation communications," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Atlanta, GA, USA, Dec. 2013, pp. 3685–3690.

[19] J. Suh, C. Kim, W. Sung, J. So, and S. W. Heo, "Construction of a generalized DFT codebook using channel-adaptive parameters," *IEEE Commun. Lett.*, vol. 21, no. 1, pp. 196–199, Jan. 2017.

[20] Y. Xie, S. Jin, J. Wang, Y. Zhu, X. Gao, and Y. Huang, "A limited feedback scheme for 3D multiuser MIMO based on Kronecker product codebook," in *Proc. IEEE 24th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Sep. 2013, pp. 1130–1135.

[21] J. Li et al., "Codebook design for uniform rectangular arrays of massive antennas," in *Proc. IEEE 77th Veh. Technol. Conf. (VTC Spring)*, Jun. 2013, pp. 1–5.

[22] Y. Wang, L. Jiang, and Y. Chen, "Kronecker product-based codebook design and optimisation for correlated 3D channels," *Trans. Emerg. Telecommun. Technol.*, vol. 26, no. 11, pp. 1225–1234, Nov. 2015.

[23] M. Arshad, I. Khan, J. Lloret, and I. Bosch, "A novel multi-user codebook design for 5G in 3D-MIMO heterogeneous networks," *Electronics*, vol. 7, no. 8, p. 144, Aug. 2018.

[24] Z. Peng and W. Li, "Two-stage 3D codebook design and fast beam search algorithm for millimeter-wave massive MIMO systems," *Electronics*, vol. 9, no. 2, p. 302, Feb. 2020.

[25] Y. Yuan, Y. Wang, W. Zhang, and F. Peng, "Separate horizontal & vertical codebook based 3D MIMO beamforming scheme in LTE—A networks," in *Proc. IEEE 78th Veh. Technol. Conf. (VTC Fall)*, Las Vegas, NV, USA, Sep. 2013, pp. 1–5.

[26] K. Xu, F.-C. Zheng, P. Cao, H. Xu, and X. Zhu, "Fast beam training for FDD multi-user massive MIMO systems with finite phase shifter resolution," *IEEE Trans. Veh. Technol.*, vol. 70, no. 1, pp. 459–473, Jan. 2021.

[27] Y. Han, S. Jin, X. Li, Y. Huang, L. Jiang, and G. Wang, "Design of double codebook based on 3D dual-polarized channel for multiuser MIMO system," *EURASIP J. Adv. Signal Process.*, vol. 2014, no. 1, pp. 1–13, Jul. 2014.

[28] M. Cai et al., "Effect of wideband beam squint on codebook design in phased-array wireless systems," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2016, pp. 1–6.

[29] M. Cai, J. N. Laneman, and B. Hochwald, "Beamforming codebook compensation for beam squint with channel capacity constraint," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2017, pp. 76–80.

[30] H. Yu, P. Guan, Y. Wang, and Y. Zhao, "Performance analysis and codebook design for mmWave beamforming system with beam squint," *IEEE Wireless Commun. Lett.*, vol. 10, no. 9, pp. 2013–2016, Sep. 2021.

[31] E. Björnson et al., "Towards 6G MIMO: Massive spatial multiplexing, dense arrays, and interplay between electromagnetics and processing," 2024, *arXiv:2401.02844*.

[32] H. Zhang, N. Shlezinger, F. Guidi, D. Dardari, and Y. C. Eldar, "6G wireless communications: From far-field beam steering to near-field beam focusing," *IEEE Commun. Mag.*, vol. 61, no. 4, pp. 72–77, Apr. 2023.

[33] M. Cui, Z. Wu, Y. Lu, X. Wei, and L. Dai, "Near-field MIMO communications for 6G: Fundamentals, challenges, potentials, and future directions," *IEEE Commun. Mag.*, vol. 61, no. 1, pp. 40–46, Jan. 2023.

[34] X. Wei, L. Dai, Y. Zhao, G. Yu, and X. Duan, "Codebook design and beam training for extremely large-scale RIS: Far-field or near-field?," *China Commun.*, vol. 19, no. 6, pp. 193–204, Jun. 2022.

[35] Y. Chen, Y. Wang, Z. Wang, and Z. Han, "Angular-distance based channel estimation for holographic MIMO," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 6, pp. 1684–1702, Jun. 2024.

[36] Y. Chen, X. Guo, G. Zhou, S. Jin, D. W. K. Ng, and Z. Wang, "Unified far-field and near-field in holographic MIMO: A wavenumber-domain perspective," *IEEE Commun. Mag.*, vol. 63, no. 1, pp. 30–36, Jan. 2025.

[37] 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.

[38] H. Zhang, N. Shlezinger, F. Guidi, D. Dardari, M. F. Imani, and Y. C. Eldar, "Beam focusing for near-field multiuser MIMO communications," *IEEE Trans. Wireless Commun.*, vol. 21, no. 9, pp. 7476–7490, Sep. 2022.

[39] Y. Xu et al., "Hashing beam training for near-field communications," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Denver, CO, USA, Jun. 2024, pp. 732–737.

[40] K. T. Selvan and R. Janaswamy, "Fraunhofer and Fresnel distances: Unified derivation for aperture antennas," *IEEE Antennas Propag. Mag.*, vol. 59, no. 4, pp. 12–15, Aug. 2017.

[41] J. Sherman, "Properties of focused apertures in the Fresnel region," *IRE Trans. Antennas Propag.*, vol. 10, no. 4, pp. 399–408, Jul. 1962.

[42] 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, 2023.

[43] I. Ahmed et al., "A survey on hybrid beamforming techniques in 5G: Architecture and system model perspectives," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 3060–3097, 4th Quart., 2018.

[44] Y. Liu, C. Ouyang, Z. Wang, J. Xu, X. Mu, and A. L. Swindlehurst, "Near-field communications: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 27, no. 3, pp. 1687–1728, Jun. 2025.

[45] J. Zhu, Z. Wan, L. Dai, M. Debbah, and H. V. Poor, "Electromagnetic information theory: Fundamentals, modeling, applications, and open problems," *IEEE Wireless Commun.*, vol. 31, no. 3, pp. 156–162, Jun. 2024.

[46] J. Boersma, "Computation of Fresnel integrals," *Math. Comput.*, vol. 14, no. 72, p. 380, 1960.

[47] M. D. Renzo and M. D. Migliore, "Electromagnetic signal and information theory," *IEEE BITS Inf. Theory Mag.*, vol. 4, no. 1, pp. 25–39, Mar. 2024.

[48] M. Milosevic, "On the nature of the evanescent wave," *Appl. Spectrosc.*, vol. 67, no. 2, pp. 126–131, Feb. 2013.

[49] A. Pizzo, A. D. J. Torres, L. Sanguinetti, and T. L. Marzetta, "Nyquist sampling and degrees of freedom of electromagnetic fields," *IEEE Trans. Signal Process.*, vol. 70, pp. 3935–3947, 2022.

[50] J. Yang et al., "Revealing the evanescent components in Kronecker product-based codebooks: Insights and applications," *IEEE Trans. Wireless Commun.*, vol. 25, pp. 4006–4020, 2026.

[51] J. A. Kong, *Theory of Electromagnetic Waves*. New York, NY, USA: Wiley-Interscience, 1975.

[52] R. Ji et al., "Extra DoF of near-field holographic MIMO communications leveraging evanescent waves," *IEEE Wireless Commun. Lett.*, vol. 12, no. 4, pp. 580–584, Apr. 2023.

[53] X. Chen, *Computational Methods for Electromagnetic Inverse Scattering*. Hoboken, NJ, USA: Wiley, 2018.

[54] G. Lerosey, J. de Rosny, A. Tourin, and M. Fink, "Focusing beyond the diffraction limit with far-field time reversal," *Science*, vol. 315, no. 5815, pp. 1120–1122, Feb. 2007.

[55] C. Ouyang, Z. Wang, B. Zhao, X. Zhang, and Y. Liu, "On the impact of reactive region on the near-field channel gain," *IEEE Commun. Lett.*, vol. 28, no. 10, pp. 2417–2421, Jul. 2024.

[56] X. Chen, S. Zhang, and Q. Li, "A review of mutual coupling in MIMO systems," *IEEE Access*, vol. 6, pp. 24706–24719, 2018.

[57] T. Gong et al., "Holographic MIMO communications: Theoretical foundations, enabling technologies, and future directions," *IEEE Commun. Surveys Tuts.*, vol. 26, no. 1, pp. 196–257, 1st Quart., 2024.

[58] S. S. A. Yuan, X. Chen, C. Huang, and W. E. I. Sha, "Effects of mutual coupling on degree of freedom and antenna efficiency in holographic MIMO communications," *IEEE Open J. Antennas Propag.*, vol. 4, pp. 237–244, 2023.

[59] R. Ji et al., "Electromagnetic hybrid beamforming for holographic MIMO communications," *IEEE Trans. Wireless Commun.*, vol. 23, no. 11, pp. 15973–15986, Nov. 2024.