Codebook Design for Extremely Large-Scale MIMO Systems: Near-field and Far-field

Xiangyu Zhang, *Student Member, IEEE,* Haiyang Zhang, *Member, IEEE,* Jianjun Zhang, *Member, IEEE,* Chunguo Li, *Senior Member, IEEE,* Yongming Huang, *Senior Member, IEEE,* Luxi Yang, *Senior Member, IEEE*

Abstract

Extremely large-scale multiple-input multiple-output (XL-MIMO) communication systems introduce a new communication paradigm called near-field communications, which identifies users' location within the near-field (Fresnel's region). In the near-field, beams can be steered in the angle and distance dimensions, resulting in an enormous codebook and a prolonged two-dimension beam alignment (BA) process. To keep low BA overhead while achieving low BA error, in this paper, we design a novel hierarchical codebook and a BA scheme for near-field XL-MIMO systems. Specifically, we first reveal the angle-offset effect of near-field beams and propose a novel spatial partition by considering the angle-offset effect. Based on the partition, we design distance-coarse and focusing beams. Distance-coarse beams are leveraged to construct the high level of the codebook for angle dimension alignment, while focusing beams are utilized to construct the last level codebook for distance dimension alignment. Corresponding to the proposed codebook structure, our BA scheme is a tree search consisting of two stages, namely the angle aligning stage and the distance aligning stage. Next, we formulate the desired codebook design problem as difference convex optimization problems, where three beam design guidelines are considered to minimize the BA error rate raised by the near-field angle-offset effect. After that, the proposed optimization problem is solved by the constrained concave-convex procedure. Numerical simulations verify the angle offset effect and our designed near-field beam. Furthermore, we

X. Zhang, C. Li, Y. Huang, and L. Yang are with the School of Information Science and Engineering, and the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China, and also with the Pervasive Communications Center, Purple Mountain Laboratories, Nanjing 211111, China (e-mail: xyzhang@seu.edu.cn; chunguoli@seu.edu.cn; huangym@seu.edu.cn; lxyang@seu.edu.cn); H. Zhang is with the School of Communication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 211111, China. (e-mail: 20220142@njupt.edu.cn); J. Zhang is with the College of Electronic and Information Engineering/College of Integrated Circuits, Nanjing University of Aeronautics and Astronautics, Nanjing, China (jianjun.zhang@nuaa.edu.cn).

show that our BA scheme only utilizes one percent of overhead but achieves a lower BA error rate compared with exhaustive searching.

Index Terms

Extremely large-scale MIMO, near-field communication, beam alignment, codebook design, and beam optimization.

I. Introduction

Over the past few years, large-scale antenna arrays and higher transmission bands have been treated as key technologies in the 5G area [1]. As a further step to achieve higher transmission rates, multiple-input multiple-output (MIMO) systems will use extremely large antenna arrays, known as Extremely large-scale MIMO (XL-MIMO) [2], [3]. Due to the expansion of the array aperture, the classical Rayleigh distance of XL-MIMO systems will be larger than the transmission distance, which means the user may be located within the near-field (Fresnel region). In the near-field, the commonly used transmission modeling assumption for the far-field, i.e., the planar wavefront assumption, no longer holds. Instead, the spherical wavefront must be considered to better model the signal phase difference among array elements [4]–[6].

Due to the spherical wavefront assumption, many fundamental techniques for present MIMO systems, including channel estimation, precoding, and combining, will be invalid for near-field communications. Some preliminary efforts have been devoted in this area [5]–[14]. In literature [5]–[7], the fundamental communication model of XL-MIMO systems considering near-field effects is proposed and analyzed. The communication model of extremely large-scale reconfigurable intelligent surfaces (RIS) is presented in literature [8], [9], in which the near-field/far-field effects of RIS are revealed. From the signal process aspect, the literature [10]–[12] study near-field channel estimation algorithms. In addition, the authors of [13], [14] propose the beamfocusing technology for near-field communications, which is capable of focusing signal energy around target locations, leading to high spectrum efficiency [15]–[17].

In XL-MIMO communication systems, aligned beamforming is required to overcome the large path loss of high-frequency signals [18]. During the initial communication phase, the transmit and receive beams should be aligned. An alignment procedure, also known as beam alignment (BA), is to find the best transmission and reception beam pair in a beam training codebook [19], [20]. The conventional far-field codebooks and beam alignment schemes have been extensively developed

in academic research and standardization. For example, literature [21] construct a random vector quantization (RVQ) codebook with vector quantization technologies, which generate the beams by mutually sampling independent vectors from a uniform distribution on the complex unit sphere to achieve orthogonality. Since the beams in RVQ do not have special physical meanings, the exhaustive search scheme is the only BA scheme that can be applied, which is inefficient. In millimeter-wave communication systems, a common choice is to give physical meaning to the beams to assist the BA, such as the Discrete Fourier Transform (DFT)-based codebook [22] that steers each beam to an angle. The authors in [23], [24] construct a hierarchical codebook with multi-resolution beams and employ a tree search algorithm as a BA scheme to determine the best beam's angle quickly.

However, the existing codebooks and beam alignment schemes [19], [20], [22]–[24] are designed based on the far-field plane wave assumption, which cannot be utilized in near-field XL-MIMO systems. For XL-MIMO systems, the codebook and beam alignment scheme design faces the following new challenges compared to the far-field codebook [25]. Firstly, the codebook for the XL-MIMO system must contain both near-field and far-field beams, which inevitably results in an expanded codebook. Searching for the best beam in the expanded codebook in a short coherent duration is a challenging task. Secondly, as a result of the beamfocusing, the BA must be processed in both angle and distance dimensions, which is more complex than searching only in the angle domain in the far-field.

Up to now, there are only very limited works of literature studying the codebook design for near-field XL-MIMO scenarios [25]–[27]. Literature [25] proposes a codebook for the RIS scenario, where the authors adopt the hierarchical architecture combined with a hierarchical BA scheme. However, their training beams at the high-level codebook may be sub-optimal. Literature [26] proposes a two-phase beam training method in which they first utilize the far-field beam to infer the users' angle roughly, and then they utilize the codebook in [25] to determine the proper beam. However, applying a far-field beam to a near-field user will result in the loss of beam gain, which may lead to BA error. Besides, literature [27] proposes a BA algorithm that reduces the BA overhead by simultaneously searching different distance areas with different frequency signals. However, the beam alignment errors may be raised due to the different frequencies.

Although an abundance of far-field codebooks and the works in [25]–[27] can be helpful examples, constructing an XL-MIMO codebook still has several challenges. They are: how to design an elementary codebook that covers both the near and far fields and ensures the beams

keep orthogonality; how to generate near-field coarse beams and utilize them to reduce the BA overhead; how to organize the BA in the distance and angle dimensions and design a reasonable hierarchical architecture to keep low BA overheads.

To address these challenges, we construct a novel near-field hierarchical codebook and a two-stage BA scheme to keep low BA overhead while achieving low BA error. Specifically, we first reveal the angle-offset effect of near-field beams and then propose a near-field orthogonal spatial partition. According to the beams' covering spatial partition, we define two kinds of beams, namely distance-coarse beams and focusing beams. We utilize distance-coarse beams to construct the high-level codebook for angle dimension alignment and focusing beams to construct the last-level codebook for distance dimension alignment. Then, we present our tree search BA algorithm, which consists of two-stage, i.e., the angle alignment stage and the distance alignment stage. Next, we formulate the desired codebook design as difference convex optimization problems, in which three beam design guidelines are considered to minimize the BA error rate raised by near-field angle-offset effects. After that, the proposed optimization problem is solved by the constrained concave-convex procedure (CCCP).

The main contributions of this paper can be summarized as follows

- Firstly, we reveal the angle-offset effect of the near-field beam, defined as that the beams' coverage angle has a distance-related offset when out of focus point. Based on that, we propose a novel spatial partition in the polar domain. According to the beams' covering spatial partition, we define two kinds of beams, namely distance-coarse beams and focusing beams. The distance-coarse beams are performed for angle alignment, while the focusing beams respond to distance alignment.
- Our proposed hierarchical codebook is constituted of several levels of subcodebooks. The high level of subcodebooks, named the angular alignment part (AAP), is constituted by the distance-coarse beams. The last level subcodebook, named the distance alignment part (DAP), consists of focusing beams. Accordingly, we propose a two-stage tree-searching BA algorithm, in which the beams at the high level and last level are successively performed to align the beam along the angle and distance dimensions, respectively. Our BA algorithm considerably reduces the BA overhead because two sequential linear searches are applied to the two-dimension search problem.
- We analyze the beam alignment error rate (BAER) arising from the angle-offset effect and propose three corresponding guidelines to minimize the BAER. Combining the three

guidelines with the beams' coverage partition constraint, we formulate the desired distance-coarse and focusing beams as difference convex optimizations. After that, we solve them with the constrained concave-convex procedure (CCCP).

Besides, we provide extensive numerical results to show the performance of our designed beam and BA. Particularly, we first show the beam pattern of the distance-coarse and focusing beams to verify the angle-offset effect and our codebook. Then, we compare our BA scheme with other near-field BA schemes in terms of BAER, and spectral efficiency, which illustrates the superiority of our algorithms.

The remaining part of this paper is organized as follows. Section II introduces our considered XL-MIMO wireless communication system and emphasizes the near-field channel model. Then, in Section III, we detail our spatial partition, codebook structure, BA algorithm, and beams design method. After that, numerical results are presented in Section IV. Finally, Section V concludes the whole paper.

Notations— The bold type $\mathbf{a}$, $\mathbf{A}$ denote the vector and matrix, respectively. $\mathbf{A}^{\mathrm{T}}$ and $\mathbf{A}^{\mathrm{H}}$ denote the transpose and the conjugate transpose of $\mathbf{A}$. Besides, $|\cdot|, \|\cdot\|_2$ represent the the norm of complex number and ℓ^2-norm, respectively. $\mathrm{Re}\{\cdot\}$ represents the real part of a complex number.

II. System and Signal Model

We consider an XL-MIMO wireless communication system as shown in Fig. 1. The base station (BS) equips an extremely large linear array with N elements, which are connected to N_{RF} RF chains. Accounting for practical limitations on the high-frequency hardware, we assume that the analog phase shifters have a constant modulus and the number of RF chains is less than the number of antenna elements, i.e., $N_{\mathrm{RF}} < N$. We suppose the antenna array is set at the X-axis, and the origin of the coordinate is located at the first antenna. Then, the location of the n-th element is $\mathbf{d}_n = (n\Delta d, 0)$.

In conventional MIMO, the transmission distance is generally larger than Rayleigh distance, denoted by $r_{\mathrm{Ray}} = \frac{2D^2}{\lambda}$, where λ is the wavelength, and $D = (N-1)\Delta d$ is the antenna aperture with $\Delta d = \frac{\lambda}{2}$, the antenna spacing. But in XL-MIMO, r_{Ray} may be larger than the transmission distance. For example, for a communication system operating in 28 GHz, the Rayleigh distance is 0.17 meters for the array with 8 antenna elements and 87.77 meters for the array with 128 antenna elements. Some users may locate in the near-field of XL-MIMO systems when the transmission distance is dozens of meters. In the far-field, the electromagnetic wave is the planar

wave, and the beam will direct in a single direction, which is represented as the red beam in Fig. 1. The near-field wavefront will be spherical, and the beam will be focused on a small area, which is represented as the blue beam in Fig. 1.

We represent the user's location as $\mathbf{u} = (r, \theta)$ in the polar coordinate system, where the angle $\theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ is measured counterclockwise with normal. We focus on the single-user case as it is easy to understand and illustrate the principle of our codebook design. The results of this work are easily extended to the multi-user case by utilizing the method in [28].

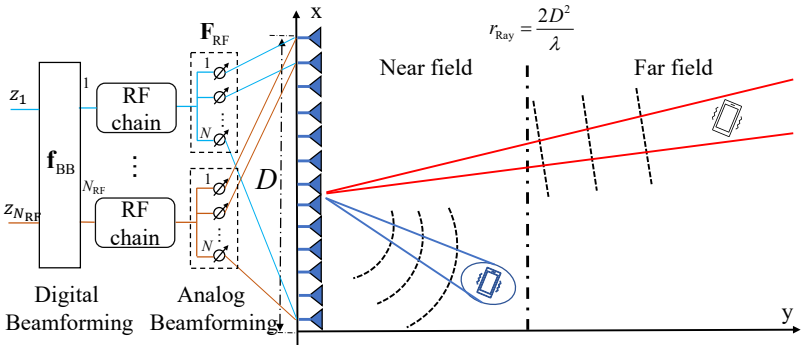


Fig. 1: The system model of XL-MIMO system

A. Near-Field Channel Model

We adopt a near-field channel model that has been extensively used in the literature [5], [10], [29]–[31], where the channel between the user and the n-th antenna element is

$$h(\mathbf{d}_n) = \frac{\lambda}{4\pi \left|\mathbf{d}_n - \mathbf{u}\right|} e^{-jk|\mathbf{d}_n - \mathbf{u}|} \sqrt{F\left(\theta\right)}, \tag{1}$$

where $\frac{\lambda}{4\pi|\mathbf{d}_n - \mathbf{u}|}$ is the free space path loss and $k = \frac{2\pi}{\lambda}$. $F\left(\theta\right)$ is the antenna radiation patterns at the BS. The channel vector for the antenna array will be $\mathbf{h} = [h(\mathbf{d}_1), \cdots, h(\mathbf{d}_N)]^{\mathrm{T}} \in \mathbb{C}^{N\times 1}$.

In the far-field, the electromagnetic wave is the planar wave, meaning that the gain and the incident angle of received signals are the same for each antenna. Thus, the distance can be approximated by $|\mathbf{d}_n - \mathbf{u}| \approx r - n\Delta d \sin\theta$. Ignoring the path loss difference raised by the array, the far-field channel is represented as

$$\mathbf{h}(\theta) = \frac{\lambda}{4\pi r} e^{-jkr} \mathbf{a}\left(\theta\right), \tag{2}$$

where $\mathbf{a}\left(\theta\right) \in \mathbb{C}^{N\times 1}$ is the far-field array response vector, writing as

$$\mathbf{a}\left(\theta\right) = \left[1, e^{jk\Delta d \sin\theta}, \ldots, e^{jk(N-1)\Delta d \sin\theta}\right]^{\mathrm{T}}. \tag{3}$$

In the near-field, the wavefront is spherical. A common simplification of the distance for near-field channels is

$$|\mathbf{d}_n - \mathbf{u}| \simeq r + \frac{(n\Delta d)^2}{2r} - n\Delta d \sin\theta. \tag{4}$$

The approximation is produced by a Maclaurin series expansion [32], [33]. Ignoring the path loss difference and only considering the phase variations among antenna elements, the near-field channel can be rewritten as

$$\mathbf{h}(r,\theta) = \frac{\lambda}{4\pi r} e^{-jkr} \mathbf{a}\left(r,\theta\right), \tag{5}$$

where $\mathbf{a}\left(r,\theta\right)$ is the near-field array response vector, given by

$$\mathbf{a}\left(r,\theta\right) = \left[1, e^{-jk\left(\frac{(\Delta d)^2}{2r} - \Delta d \sin\theta\right)}, \cdots, e^{-jk\left(\frac{((N-1)\Delta d)^2}{2r} - (N-1)\Delta d \sin\theta\right)}\right]^{\mathrm{T}}. \tag{6}$$

The difference between (3) and (6) is the second-order item on the phase.

B. Signal Model

In the BA phase, the BS performs the beam selected from a designed training beam codebook and sends the orthogonal pilot $\mathbf{z} \in \mathbb{C}^{1\times\tau}$ with τ symbols, following $\mathbf{z}\mathbf{z}^{\mathrm{H}} = 1$. The received pilot signals at the user side can be formulated as

$$\mathbf{y} = \sqrt{p}\mathbf{h}^{\mathrm{H}}\mathbf{f}\mathbf{z} + \mathbf{n}, \tag{7}$$

where p is the transmit power, $\mathbf{n}$ is the additive Gaussian noise. The elements of $\mathbf{n}$ follow complex Gaussian distribution, i.e., the i-th elements $\mathbf{n}_i \sim \mathcal{CN}(0,\sigma^2)$ with σ^2 representing the noise variance. $\mathbf{f}$ represents the beamforming vector, which is combined with analog beamforming $\mathbf{F}_{\mathrm{RF}} \in \mathbb{C}^{N\times N_{\mathrm{rf}}}$ and digital beamforming $\mathbf{f}_{\mathrm{BB}} \in \mathbb{C}^{N_{\mathrm{rf}}\times 1}$, i.e., $\mathbf{f} = \mathbf{F}_{\mathrm{RF}}\mathbf{f}_{\mathrm{BB}}$. Because $\mathbf{F}_{\mathrm{RF}}$ is realized by phase shifters, the n-th row and n_{RF} column element follows $e^{j\varphi_{(n,n_{\mathrm{RF}})}}$, with $\varphi \in [0, 2\pi]$. We assume $\mathbf{f}$ is normalized, i.e., $\|\mathbf{f}\|_2 \leq 1$. As [34], [35], we focus on the design of $\mathbf{f}$, because the $\mathbf{f}_{\mathrm{RF}}$ and $\mathbf{f}_{\mathrm{BB}}$ can be well recovered once $\mathbf{f}$ is available .

III. Codebook Structure and Beam Alignment Method

In this section, we construct our XL-MIMO codebook and present our BA algorithm. Particularly, in subsection III-A and III-B, we introduce the codebook structure and the spatial partition, respectively. Then, in III-C, we formulate the beam design problem, and an iterative algorithm is proposed for the corresponding problem. Finally, in Subsection III-D, we detail our beam alignment method.

A. Codebook Structure

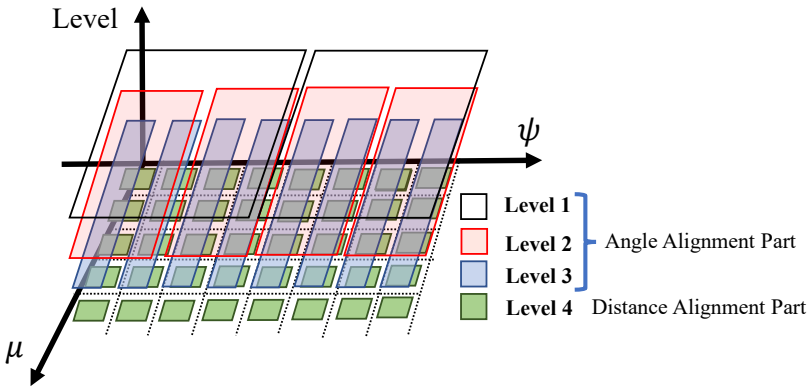


Fig. 2: The illustration of our codebook structure

It is widely accepted that the XL-MIMO system has several new characteristics, including high-resolution beams, an enormous beamspace, the near-field user, and the beamfocusing effect. An ad hoc XL-MIMO codebook needs to be designed to adapt these characteristics. Specifically, as the users of the XL-MIMO system may locate in both near-field and far-field, the XL-MIMO codebook should include massive beams for near-field beamfocusing and far-field beamsteering vectors. The near-field beamfocusing and far-field beamsteering vectors in the XL-MIMO codebook will be organized along distance and angle dimensions according to their coverage areas. Accordingly, beams need to be aligned in the distance and angle dimensions. However, it is challenging to search for the best beam within a short coherent time.

In response to this challenge, we propose a hierarchical multi-layer codebook. We show our codebook structure in Fig. 2, in which we define $\mu = \frac{1}{r}$ and $\psi = \sin\theta$ for easy expression. Our codebook contains two kinds of beams, distance-coarse beams and focusing beams. The distance-coarse beams cover both near-field and far-field, whose coverage area is shown by the black, red, and blue rectangles in Fig. 2. The focusing beams own the highest resolution in angle and distance, which is represented as the green square in Fig. 2. We utilize the distance-coarse beams to construct the high-level sub-codebook, termed the angular alignment part (AAP). It can also be observed that the distance-coarse beams at the high level of AAP cover a wider angle range for roughly determining the angle of the best beam, as shown with red rectangles. The low-level AAP will utilize more beams to cover the same angular range for more refined determinations, as illustrated by the blue rectangles. As its name implies, the AAP is only used

for angle alignments. In contrast to AAP, we utilize the focusing beam to construct the last level subcodebook, named the distance alignment part (DAP). The DAP aligns the beam along the distance because the focusing beam can be treated as the child node of distance-coarse beams that subdivides the coverage area along the distance dimension.

To illustrate the structure of our codebook more clearly, we show the covering relationship of the beams in mathematical ways. We assume the codebook has $L+1$ levels of subcodebook. The first L level is AAP, and the last is DAP. The DAP contains $M_D = Q \times B$ beams, corresponding to Q distance partitions and $B = 2^L$ angle partitions. We represent $\mathbf{f}_{q,b}^{(\mathrm{D})}$ as the focusing beam focusing at q-th distance and b-th angle partition. The coverage area of $\mathbf{f}_{q,b}^{(\mathrm{D})}$ is represented as $\mathbb{M}_{q,b}^{(\mathrm{D})}$. Meanwhile, we represent $\mathbf{f}_{l,b}^{(\mathrm{A})}$ as the distance-coase beam at the l-th layer of AAP and direct to b-th angle partition. The coverage area of $\mathbf{f}_{l,b}^{(\mathrm{A})}$ is represented as $\mathbb{M}_{l,b}^{(\mathrm{A})}$. According to the description before, the covering relationship between $\mathbb{M}_{L,b}^{(\mathrm{A})}$ and $\mathbb{M}_{q,b}^{(\mathrm{D})}$ can be represented as

$$\mathbb{M}_{L,b}^{(\mathrm{A})} = \bigcup_{q=1}^{Q} \mathbb{M}_{q,b}^{(\mathrm{D})}. \tag{8}$$

The $\mathbb{M}_{l,b}^{(\mathrm{A})}$ can be represented as

$$\mathbb{M}_{l,b}^{(\mathrm{A})} = \bigcup_{b'=(b-1)\mathcal{B}}^{b\mathcal{B}} \mathbb{M}_{L,b'}^{(\mathrm{A})}, \tag{9}$$

where $\mathcal{B} = 2^{(L-l)}$.

In the BA stage, the subcodebooks are successively performed from up to the bottom. The binary tree search is first applied at AAP to align the beam along the angular dimension. After the angle is aligned, the distance searching beam at the corresponding angle is performed to determine the user's distance. The detail of our BA approach will be presented in Subsection III-C.

B. Spatial Partition

In this section, we reveal the angle-offset effect of the near-field beam and illustrate our codebook by presenting the spatial partition.

The traditional DFT codebook partitions the angular dimension by minimizing the maximum correlation coefficient of any two array response vectors of spatial partitions, which will ensure the beams are as orthogonal as possible. We extend this idea to the near-field and obtain our

spatial partition. The correlation coefficient of any two near-field array response vectors is given by

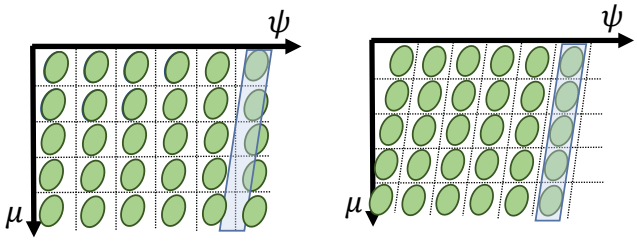
$$\begin{aligned}\sigma(\mu_1,\mu_2,\psi_1,\psi_2) =&\mathbf{a}\left(\mu_1,\psi_2\right)^{\mathrm{H}}\mathbf{a}\left(\mu_2,\psi_2\right)\\ \simeq&\sum_{n=0}^{N-1}\exp\left(jk\left(\frac{1}{2}(n\Delta d)^2(\mu_1-\mu_2)-n\Delta d(\psi_1-\psi_2)\right)\right),\end{aligned} \tag{10}$$

where (μ_1,ψ_1), (μ_2,ψ_2) are two coordinate of partition. It can be observed that the maximum $\sigma(\mu_1,\mu_2,\psi_1,\psi_2)$ of all partitions is achieved at the two nearest neighboring spatial partitions. To minimize the maximum $\sigma(\mu_1,\mu_2,\psi_1,\psi_2)$ of any two spatial partitions, we need to maximize the distance of any two neighboring partitions. Thus, the near-field partition can be obtained by evenly dividing $\psi\in[-1,1]$ into B parts and $\mu=\frac{1}{r}\in[0,\frac{1}{r_{\text{Fre}}}]$ into Q parts, where r_{Fre} is the Fraunhofer distance. Mathematically, the spatial partition can be formulated as

$$\begin{aligned}\mathbb{M}_{q,b}=\left\{(\mu,\psi)\,|(q+1)\Delta\mathcal{D}<\mu<q\Delta\mathcal{D},-1+(b-1)\Delta\mathcal{A}<\psi<-1+b\Delta\mathcal{A}\right\},\\ b=[1,\cdots,B],q=[0,\cdots,Q],\end{aligned} \tag{11}$$

where $\Delta\mathcal{D}=\frac{1}{Qr_{\text{Fre}}}$ and $\Delta\mathcal{A}=\frac{2}{B}$ are the distance interval and the angle interval, respectively.

We present the spatial partition in Fig. 3(a). However, the partition in (11) may not support our codebook design because the distance-coarse and focusing beam have the angle offset effect. The angle offset effect is defined as that the coverage angle have a distance-related offset when out of focus point. In the following, we present the angle offset effect by revealing the coverage area of the near-field beam, which is defined as the area where the magnitude response is large than -3 dB. Then, we propose the adjusted spatial partition.



(a) The spatial partition without considering angle offset

(b) The spatial partition with considering angle offset

Fig. 3: The illustration of spatial partition with/without angle offset

Theorem 1 (The angle-offset effect of focusing beam): For a focusing beam $\mathbf{f}$ that focuses at one point with the angle $\psi_{\mathbf{f}}$ and distance $\mu_{\mathbf{f}}$, the coverage area of can be approximated by

$$\mathbb{M}_{\mathbf{f}} = \left\{ (\mu, \psi) \,\middle|\, |\mu - \mu_{\mathbf{f}}| < \hat{\mu}, |\psi - \psi_{\mathbf{f}} + \Delta\psi| < \hat{\psi} \right\}, \tag{12}$$

where $\Delta\psi$ is the angle offset, which can be represented as

$$\Delta\psi = \frac{1}{2}(N-1)\Delta d(\mu - \mu_{\mathbf{f}}). \tag{13}$$

$\hat{\mu}$ and $\hat{\psi}$ is the 3 dB boundary, which is bounded by $\hat{\mu} < \frac{2}{(N-1)^2\Delta d}$ and $\hat{\psi} < \frac{1}{N-1}$.

Proof: Please refer to Appendix A. ■

Theorem 1 reveals the coverage area of near-field focusing beams. The actual coverage area of the beam is an ellipse with respect to μ and ψ. But for easy to partition the space, we approximate the coverage area as a diamond. To illustrate Theorem 1, we show the beam patterns of $\mathbf{f}$ that focus at point $\mu = 0.23\ m^{-1}$ and angle $\sin\theta = 0$ in Fig. 4. $\mathbf{f}$ is generated according to [14]. The antenna array has 128 elements mm-wave with $D = \frac{\lambda}{2}$ antenna spacing working in the 28GHz. We mark the focus point with a red-filled diamond and the line (13) with a red dashed line. It can be easily found that the main lobe is systematic to the red line, which is not perpendicular to the axis. And the angle of the main lobe has a distance-related offset when out of focus point.

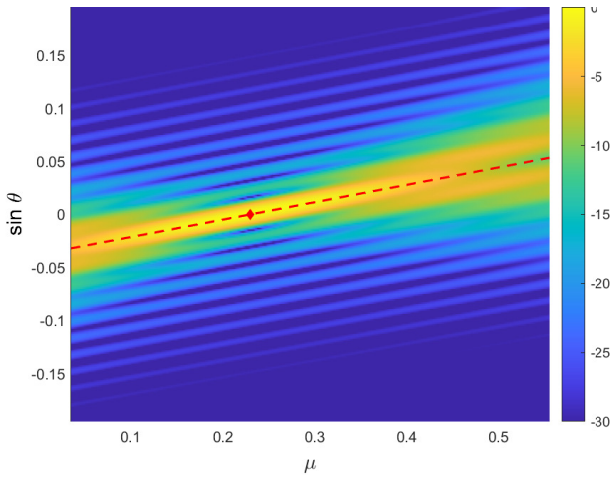


Fig. 4: The illustration of the angle-offset effect with focusing beam

Corollary 1: According to the coverage area of the focusing beam in (12), we can have the upper boundary of distance interval, i.e., $\Delta\mathcal{D} < \frac{2}{(N-1)^2\Delta d}$ and the upper boundary angle interval, i.e., $\Delta\mathcal{A} < \frac{1}{N-1}$.

Theorem 2 (The angle-offset effect of distance-coarse beam): Similar to the focusing beam, the distance-coarse beams covering the whole distance area and covering angle in far-field is $\psi_l < \psi < \psi_h$, the coverage area is

$$\mathbb{M}_{\mathbf{f}} = \left\{ (\mu, \psi) \left| 0 < \mu < \frac{1}{r_{\text{Fre}}}, \Delta\psi_d + \psi_l < \psi < \Delta\psi_d + \psi_h \right. \right\}. \tag{14}$$

$\Delta\psi_d$ is the angle offset, given by

$$\Delta\psi_d = \kappa(N-1)\Delta d\mu, \tag{15}$$

where κ is the solution of the following equation

$$\frac{1}{N}\sum_{p=0}^{\Delta p}\sum_{n=0}^{N-1}\exp\left(j\pi\left(\frac{\kappa^2}{N}n^2 - (2\kappa - \frac{2p}{N})n\right)\right) = 1/\sqrt{2}, \tag{16}$$

where Δp represents the coverage angle range. Certainly, κ has the boundary $0 < \kappa < 1/2$.

Proof: Please refer to Appendix B. ■

Theorem 2 reveals the coverage area of distance-coarse beams, which are beams contained by AAP. Similar to Theorem 1, (14) show that the coverage angle has a linear offset, which is $\Delta\psi_d = \kappa(N-1)\Delta d\mu$. Then, we have the coverage area for the distance-coarse beam at the last level of AAP, which is

$$\mathbb{M}_{L,b}^{(\text{A})} = \left\{ (\mu, \psi) \left| 0 < \mu < \frac{1}{r_{\text{Fre}}}, -1 + (b-1)\Delta\mathcal{A} + \Delta\psi_d < \psi < -1 + b\Delta\mathcal{A} + \Delta\psi_d \right. \right\}, \quad b = [1, \cdots, B]. \tag{17}$$

The coverage area of a distance coarse beam is shown as the blue parallelogram in Fig .3. It can be observed that due to the angle offset, the spatial partition in (11) may not support covering the coverage area of the distance-coarse beam with focusing beams, i.e., (8). Hence, we add the angle offset in the partition and obtain the following adjusted spatial partition, which is:

$$\mathbb{M}_{q,b}^{(\text{D})} = \left\{ (\mu, \psi) \left| (q+1)\Delta\mathcal{D} < \mu < q\Delta\mathcal{D}, -1 + (b-1)\Delta\mathcal{A} + q\Delta\psi_a < \psi < -1 + b\Delta\mathcal{A} + q\Delta\psi_a \right. \right\}, \quad b = [1, \cdots, B], q = [0, \cdots, Q], \tag{18}$$

where $\Delta\psi_a = \frac{\Delta\psi_d}{Q}$.

C. Beam Alignment Method

According to our codebook structure, we propose a two stages BA scheme. The first stage corresponds to the angle alignment stage, in which binary tree searching is performed on the

AAP to align the beam in the angle dimension. After the beam is aligned in the angle dimension, BS will perform the focusing beam in the DAP along the distance dimension. Our BA algorithm can significantly save BA overhead because it decouples the two-dimensional BA process into a tree search and a linear search.

Our BA algorithm is detailed in Algorithm 1, where $\mathbf{f}_{l,b}^{(\mathrm{A})}$ denotes the angle alignment beam in the l-th layer, b-th angle partition and $\mathbf{f}_{q,b}^{(\mathrm{D})}$ denotes the distance alignment beam corresponding to k-th distance partition and b-th angle partition. $y = \left|\mathbf{y}^{\mathrm{H}}\mathbf{z}\right|$ represents the received signal power of beams. Steps 2 to 14 of Algorithm 1 is the angle alignment stage. In each angle alignment time, BS transmits the training beam pair $\{\mathbf{f}_{l,2b+1}^{(\mathrm{A})}, \mathbf{f}_{l,2b+2}^{(\mathrm{A})}\}$. According to the received signal power of the beam pair denoted as $\{y_{l,2b+1}^{(\mathrm{A})}, y_{l,2b+2}^{(\mathrm{A})}\}$, the user sends b' back to BS, where $b' = 2b$ if $|y_{l,2b+1}^{(\mathrm{A})}| > |y_{l,2b+2}^{(\mathrm{A})}|$, otherwise $b' = 2b+1$. After that, BS selects the child node beam pair in the next layer, i.e., $\{\mathbf{f}_{l+1,2b'+1}^{(\mathrm{A})}, \mathbf{f}_{l+1,2b'+2}^{(\mathrm{A})}\}$. The BS repeats this process until the L level is reached and the beam along the angle dimension is aligned. Then, the BS sequentially performs the beams at the aligned angle of the last level subcodebook, as shown in steps 15 to 20 of Algorithm 1. Finally, the beam with the highest received signal power is the aligned beam.

In the above BA algorithm, an error may be generated at the transition of two stages due to the slope difference between the focusing beam and the distance-coarse beam, i.e., $\kappa \neq \frac{1}{2}$. Considering this, we broaden the search partition to eliminate it, as shown in steps 10 and 12.

Next, we compare the BA overhead of our BA scheme with several BA schemes in [25], [26], as shown in Table I. In Table I, L represents the number of child nodes for each root node [25], and c represents the number of candidate beams in [27]. From this table, we can conclude that the BA scheme adopting hierarchical searches has significantly lower BA overheads due to the $\log$ operation. In contrast with our BA scheme, the BA scheme in [25] adopts a hierarchical architecture to align the distance dimension as well. However, as we analyze in the Theorem 1, the number of distance partitions Q should be a small number, which will support $Q \leq L\log_L(Q)$ to be held in most scenarios.

Algorithm	Exhaustive search	Hierarchical BA in [25]	Two-stage BA scheme in [27]	our BA scheme
BA overhead	$\mathcal{O}(QB)$	$\mathcal{O}(L\log_L(QB))$	$\mathcal{O}(B+cQ)$	$\mathcal{O}(2\log_2(B)+2Q)$

TABLE I: The BA overhead comparison

Algorithm 1 Two Stage Tree Searching Algorithm

Initialization $b = 0$.
for $l = 1, 2 \ldots, L$ **do**
  BS performs $\{\mathbf{f}_{l,2b+1}^{(\mathrm{A})}, \mathbf{f}_{l,2b+2}^{(\mathrm{A})}\}$ and the user observe $\{y_{l,2b+1}^{(\mathrm{A})}, y_{l,2b+2}^{(\mathrm{A})}\}$;
  $b^* = \underset{q \in \{2b+1, 2b+2\}}{\arg\max} |y_{l,b}^{(\mathrm{A})}|$
  if $b^* = 2b + 1$ and $l \neq L$ **then**
    User sends back 0, and BS updates the index of the beam $b' = 2b$;
  else if $b^* = 2b + 2$ and $l \neq L$ **then**
    User sends back 1, and BS updates the index of the beam $b' = 2b + 1$;
  else if $b^* = 2b + 1$ and $l = L$ **then**
    Obtain the possible aligned user's angle partition index set $\mathbb{P} = [2b, 2b + 1]$;
  else if $q^* = 2b + 2$ and $l = L$ **then**
    Obtain the possible aligned user's angle partition index set $\mathbb{P} = [2b + 2, 2b + 3]$;
  end if
  $b = b'$;
end for
for $q = 1, 2 \ldots, Q$ **do**
  for $\hat{b} \in \mathbb{P}$ **do**
    BS use $\mathbf{f}_{q,\hat{b}}^{(\mathrm{D})}$ and the the user obtain $y_{q,\hat{b}}^{(\mathrm{D})}$;
  end for
end for
Obtain the aligned user's beams index $\{q^*, b^*\} = \underset{q \in [1,Q], b \in \mathbb{P}}{\arg\max} |y_{q,b}^{(\mathrm{D})}|$.

D. Beams Design

In this subsection, we propose the guidelines for designing the beams and formulate the training codebook design as an optimization problem. Then, an iterative algorithm is proposed to address the formulated non-convex optimization problem.

1) Beam design guidelines: The distance-coarse and focusing beams are designed to minimize the BA Error Rate (BAER). The BA error is defined as the scenario when the user is not located at the coverage area of the aligned beam, and the BA error rate is the ratio between the BA error times and the total BA times. Next, we will present the guidelines for designing the magnitude response of the coverage area, the ripple of the coverage area, and the transition band to minimize the BAER.

Proposition 1 (The guidelines of the magnitude response): The magnitude response in the beam's coverage area should be as large as possible, and the magnitude response in other areas should be as low as possible.

Proof: In the angle alignment stage, BA Error is mainly caused by the misjudgment (Steps 4 in Algorithm 1), which is occurred when the user is inside $\mathbb{M}_{l,2b+1}^{(\mathrm{A})}$ but $y_{l,2b+1}^{(\mathrm{A})} < y_{l,2b+2}^{(\mathrm{A})}$ or the user is inside $\mathbb{M}_{l,2b+2}^{(\mathrm{A})}$ but $y_{l,2b+1}^{(\mathrm{A})} > y_{l,2b+2}^{(\mathrm{A})}$. The probability of such two situation can be represented as $\mathcal{P}_{l,2b+1}(\mathbf{f}_{l,2b+1}^{(\mathrm{A})}, \mathbf{f}_{l,2b+2}^{(\mathrm{A})})$ and $\mathcal{P}_{l,2b+2}(\mathbf{f}_{l,2b+1}^{(\mathrm{A})}, \mathbf{f}_{l,2b+2}^{(\mathrm{A})})$. The probability $\mathcal{P}$ have the follows properties [36]

$$\mathcal{P}_{l,2b+1}(\mathbf{f}_{l,2b+1}^{(\mathrm{A})}, \mathbf{f}_{l,2b+2}^{(\mathrm{A})}) \leq \exp\left(-(v_{l,2b+1} - v_{l,2b+2})^2/2\right), \tag{19}$$

$$\mathcal{P}_{l,2b+2}(\mathbf{f}_{l,2b+1}^{(\mathrm{A})}, \mathbf{f}_{l,2b+2}^{(\mathrm{A})}) \leq \exp\left(-(v_{l,2b+2} - v_{l,2b+1})^2/2\right), \tag{20}$$

where $v_{l,b} = \left|\mathbf{a}^{\mathrm{H}}(\mu,\psi)\mathbf{f}_{l,b}^{(\mathrm{A})}\right|$; μ, ψ are the distance and angle of user.

Besides, in the distance alignment stage, the misjudgment in the steps 21 of Algorithm 1 may also raised BA error, which is occurred when $|y_{q',b'}^{(\mathrm{D})}| < \max\limits_{q\in[1,Q],b\in[1,2^L]}\{|y_{q,b}^{(\mathrm{D})}|\}$. Thereby, we can represent the probability of the misjudgment of beams $\mathbf{f}_{q',b'}^{(\mathrm{D})}$ is represented as

$$P_{q',b'}(\mathbf{f}_{q,b}^{(\mathrm{D})}) = 1 - \prod_{q=1}^{Q}\prod_{b=1}^{2^L} Q\left(\frac{v_{q,b} - v_{q',b'}}{\delta_n}\right), \tag{21}$$

where $v_{q,b} = \left|\mathbf{a}^{\mathrm{H}}(\mu,\psi)\mathbf{f}_{q,b}^{(\mathrm{D})}\right|$. $Q = \frac{1}{\sqrt{2\pi}}\int_x^\infty \mathrm{e}^{-\frac{t^2}{2}}\,\mathrm{d}t$ represents the Q-function and δ_n is the power of noise.

From (19), (20) and (21), we have the conclusion that the lower magnitude response in the coverage area $v_{q,b}$ and the higher magnitude response in other coverage areas $v_{q,b}$ results in a higher BA Error as described by papers [37], [38]. ■

Proposition 1 is suitable for both AAP and DAP, which is the most general guideline for codebook design that is also utilized in the conventional codebook design. Next, we state the guidelines for the beams in the DAP.

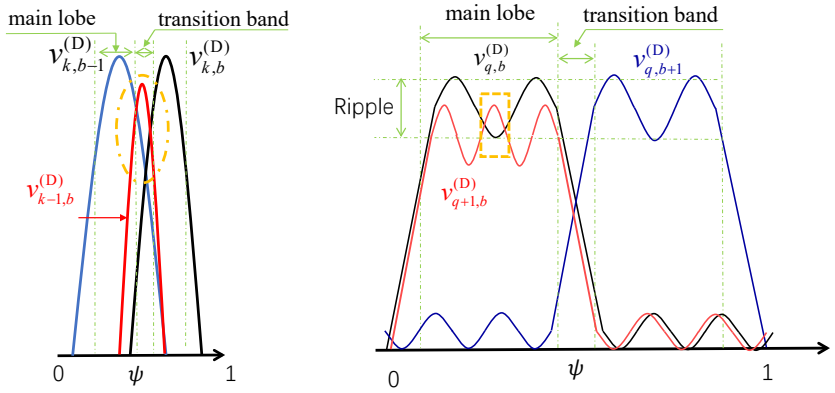
Proposition 2 (The guidelines of the transition band): The transition band should be as narrow as possible.

Proof: For easy illustration, we present the situation that the BAER raised by the transition band in Fig. 5(a). The dark and red lines represent the beam pattern of $\mathbf{f}_{q,b}^{(\mathrm{D})}$ and $\mathbf{f}_{q+1,b}^{(\mathrm{D})}$ at the distance partition q. Due to the offset in Theorem 1, the peak of $\mathbf{v}_{q-1,b}^{(\mathrm{D})}$ may located between the transition band between $\mathbf{v}_{q,b+1}^{(\mathrm{D})}$ and $\mathbf{v}_{q,b}^{(\mathrm{D})}$, as the red line shown. It is obvious that the error

alignment will occur inside the yellow circle. We eliminate this negative effect by adjusting the width of the transition band. Hence, we have the guideline of the transition band. ■

Proposition 3 (The guidelines of the ripple): The ripple in the coverage area should be as small as possible.

Proof: We present the BAER raised by the ripple of the coverage area in Fig. 5(b). In this figure, the dark and red lines represent the beam pattern of $\mathbf{f}_{q,b}^{(\mathrm{D})}$ and $\mathbf{f}_{q+1,b}^{(\mathrm{D})}$ at the distance area q. In normal, the beam pattern of $\mathbf{f}_{q,b}^{(\mathrm{D})}$ is higher than $\mathbf{f}_{q+1,b}^{(\mathrm{D})}$. However, the contrary may be raised by the ripple. As shown in the yellow rectangle, the peak magnitude of the mismatched beam pattern can be greater than the valley magnitude of the matched beam pattern. Thereby, the huge ripple in the mainlobe results in the BA error. ■



(a) The BAER raised by the transition band

(b) The BAER raised by the ripple

Fig. 5: The illustration of BAER raised by the transition band and the ripple.

Next, we formulate the problem of training codebook design as an optimization problem according to three guidelines and the converging area constraint. We denote $\mathbb{M}$ as the coverage area and $\mathbb{S}$ as the area out of the converging area. Then, the design of distance-coarse $\mathbf{f}_{l,b}^{(\mathrm{A})}$ can be formulated as following optimization problem

$$\begin{aligned} \max_{\mathbf{f}_{l,b}^{(\mathrm{A})}} \quad & \zeta_M - \zeta_S \\ \text{s.t.} \quad & \zeta_M = \min_{(r_m,\psi_m)\in\mathbb{M}} \left|\mathbf{a}(r_m,\psi_m)^{\mathrm{H}}\mathbf{f}_{l,b}^{(\mathrm{A})}\right| \\ & \zeta_S = \max_{(r_s,\psi_s)\in\mathbb{S}} \left|\mathbf{a}(r_s,\psi_s)^{\mathrm{H}}\mathbf{f}_{l,b}^{(\mathrm{A})}\right| \\ & \|\mathbf{f}_{l,b}^{(\mathrm{A})}\|_2 \leq 1 \end{aligned} . \tag{22}$$

The design of focusing beam $\mathbf{f}_{q,b}^{(\mathrm{D})}$ can be formulated as

$$\begin{aligned}
\max_{\mathbf{f}_{q,b}^{(\mathrm{D})}} \quad & \zeta_M - \zeta_S \\
\text{s.t.} \quad & \zeta_M = \min_{(r_m,\psi_m)\in\mathbb{M}} \left|\mathbf{a}(r_m,\psi_m)^{\mathrm{H}}\mathbf{f}_{q,b}^{(\mathrm{D})}\right| \\
& \zeta_S = \max_{(r_s,\psi_s)\in\mathbb{S}} \left|\mathbf{a}(r_s,\psi_s)^{\mathrm{H}}\mathbf{f}_{q,b}^{(\mathrm{D})}\right| \\
& \left|\mathbf{a}(r_m,\psi_m)^{\mathrm{H}}\mathbf{f}_{q,b}^{(\mathrm{D})}\right| < \zeta_M + r_p, (r_m,\psi_m) \in \mathbb{M} \\
& \|\mathbf{f}_{q,b}^{(\mathrm{D})}\|_2 \le 1
\end{aligned}, \tag{23}$$

where r_p is tolerated ripple boundary.

In Problem (22) and (23), our objective is to maximize the difference between the main lobe and the side lobe, which follows the guideline of magnitude response, i.e., Proposition 1. The guideline of the transition band, i.e., Proposition 2, is guaranteed by our spatial partition. The guideline of the ripple is guaranteed by the third constraint in (23). By comparing the formulated optimization problem (22) and (23), the problem (22) can be treated as the special case of Problem (23), i.e., $r_p = \infty$. Thereby, we only focus on Problem (23).

However, in Problem (23), $\mathbb{M}$ and $\mathbb{S}$ are continuous and uncountable, resulting in prohibitive computational complexity and frequently intractable beam design issues. We sample the constrained point inside the $\mathbb{M}$ and $\mathbb{S}$. Then, we drop the subscripts and superscripts, and the problem (23) after sampling is

$$\begin{aligned}
\max_{\mathbf{f}} \quad & \zeta_M - \zeta_S \\
\text{s.t.} & \zeta_M = \min_{m=1,\cdots,V_m} \left|\mathbf{a}_m^{\mathrm{H}}\mathbf{f}\right| \\
& \zeta_S = \max_{s=1,\cdots,V_s} \left|\mathbf{a}_s^{\mathrm{H}}\mathbf{f}\right| \\
& \left|\mathbf{a}_m^{\mathrm{H}}\mathbf{f}\right| < \zeta_M + r_p, (m = 1,\cdots,V_m) \\
& \|\mathbf{f}\|_2 \le 1,
\end{aligned} \tag{24}$$

where $\mathbf{a}_m^{\mathrm{H}}$ is array responds vector corresponding to the sampled points from $\mathbb{M}$ and $\mathbf{a}_s^{\mathrm{H}}$ is array responds vector corresponding to the sampled points from $\mathbb{S}$. V_m and V_s is total sample number in the coverage are and out of the coverage area.

Here, we discuss the discretization error because of great differences between $\left|\mathbf{a}^{\mathrm{H}}(r,\psi)\mathbf{f}\right|$ and $\left|\mathbf{a}^{\mathrm{H}}(r+o_r,\psi+o_a)\mathbf{f}\right|$, where o_r and o_a are two real number that is smaller than sample distance. To illustrate the influence by o_a, we expand $\left|\mathbf{a}^{\mathrm{H}}(r,\psi+o_a)\mathbf{f}\right|$ by Taylor series as follows

$$\left|\mathbf{a}^{\mathrm{H}}(r,\psi+o_a)\mathbf{f}\right| = \left|\mathbf{a}^{\mathrm{H}}(r,\psi)\mathbf{f}\right| + o_a \mathrm{Re}\left\{\mathbf{a}^{\mathrm{H}}(\mathrm{r},\psi)\mathbf{f}\mathbf{f}^{\mathrm{H}}\frac{\partial\mathbf{a}(\mathrm{r},\psi)}{\partial\psi}\right\} + \mathcal{O}(\mathrm{o_a}). \tag{25}$$

In (25), it can be observed that the difference is mainly raised by the second item on the right side of the equation. As $\|\mathbf{f}\|_2 \leq 1$, $\mathrm{Re}\{\mathbf{a}^{\mathrm{H}}(\mathrm{r},\psi)\mathbf{f}\mathbf{f}^{\mathrm{H}}\frac{\partial \mathbf{a}(\mathrm{r},\psi)}{\partial \psi}\} \leq \mathrm{N}$ holds. Then, we can constrain the sampling interval to eliminate the error.

2) Iterative Algorithm: Problem (24) is a typical Difference of Convex (DC) function optimization problem [39] that we can transform into a convex problem to solve. Due to a mass of constrain in (24), we leverage one of the most efficient methods, constrained concave-convex procedure (CCCP) [40] to tackle this problem. The first step of CCCP is to approximate the concave part of our problem, i.e., $-\zeta_M$ with one-order Taylor expansion, and turn the DC problem into a convex problem. By the convexity of the constrain of ζ_M and ζ_S, we obtain

$$\|\mathbf{a}^{\mathrm{H}}\mathbf{f}\|_2 = \|\mathbf{a}^{\mathrm{H}}\mathbf{f}_t\|_2 + \mathrm{Re}\left\{\mathbf{f}_{\mathrm{t}}^{\mathrm{H}}\mathbf{a}\mathbf{a}^{\mathrm{H}}(\mathbf{f}-\mathbf{f}_{\mathrm{t}})\right\}. \tag{26}$$

The solution of Problem (24) can be obtained by iteratively solving the following convex problem,

$$\begin{aligned}
\max_{\mathbf{f}} \ & \zeta_M - \zeta_S \\
\text{s.t. } & \zeta_M = \max_{m=1,\cdots,V_m} \|\mathbf{a}_m^{\mathrm{H}}\mathbf{f}_t\|_2 + \mathrm{Re}\left\{\mathbf{f}_{\mathrm{t}}^{\mathrm{H}}\mathbf{a}_{\mathrm{m}}\mathbf{a}_{\mathrm{m}}^{\mathrm{H}}(\mathbf{f}-\mathbf{f}_{\mathrm{t}})\right\} \\
& \zeta_S = \max_{s=1,\cdots,V_s} \|\mathbf{a}_s^{\mathrm{H}}\mathbf{f}_t\|_2 + \mathrm{Re}\left\{\mathbf{f}_{\mathrm{t}}^{\mathrm{H}}\mathbf{a}_{\mathrm{s}}\mathbf{a}_{\mathrm{s}}^{\mathrm{H}}(\mathbf{f}-\mathbf{f}_{\mathrm{t}})\right\} \\
& \|\mathbf{a}_m^{\mathrm{H}}\mathbf{f}_t\|_2 < \zeta_M + r_p, (m = 1, \cdots, V_m) \\
& \|\mathbf{f}\|_2 \leq 1.
\end{aligned} \tag{27}$$

Problem (27) is a convex problem that can be easily solved. As a result, a KKT solution of (24) can be obtained monotonically using the CCCP algorithm [40]. The problem (22) has the same solving process, which only drops the second convex constraint. The initialization of the CCCP algorithm affects its convergence performance significantly. We set the initialization as $\mathbf{f}_0 = \frac{1}{V_m}\sum_{m=1}^{V_m}\mathbf{a}_m$. We conclude the CCCP algorithm in Algorithm 2.

IV. NUMERICAL RESULT

This section presents numerical results to present the beam pattern of our beams and the BA performance of our proposed codebook. Specifically, we first show the designed beam in Subsection IV-A and IV-B. Then, we show the BA accuracy and the achievable rate of our codebook in Subsection IV-C.

Throughout our simulations, we consider a ULA array with 128 elements working at $f = 28$ GHz, whose antenna spacing Δd is $\lambda/2$. Under this setting, $\kappa \simeq 1/8$ in Theorem 2 when

Algorithm 2 Beam design for hierarchical XL-MIMO codebook

1: **Input:** $\mathbf{a}_m$, $\mathbf{a}_s$, r_p
2: **Initialization:** $\mathbf{f}_0 = \frac{1}{V_m}\sum_{m=1}^{V_m}\mathbf{a}_m$, $e_p = \infty$, and $t = 0$.
3: **while** $e_p < E_b$ **do**
4: Solving the optimization problem (27) and obtain $\mathbf{f}_t$;
5: Calculating the iterative error $e_p = |\zeta_M(t) - \zeta_M(t-1)| + |\zeta_S(t) - \zeta_S(t-1)|$;
6: $t = t + 1$;
7: **end while**
8: **Output:** designed training beamforming vector $\mathbf{f}_{\text{opt}} = \mathbf{f}_t$.

$\Delta p \geq 2$. The Rayleigh distance is $r_{Ray} = 87.77$ m, and the Fresnel distance is $r_{Fre} = 1.26$ m. We assume the user is located within the area from 2.22 m to 100 m and from $-90°$ to $90°$ corresponding to $\mu \in [0.01, 0.477]$ and $\psi \in [-1, 1]$, respectively. Then, we have 128 angle partitions as $\Delta\mathcal{A} = \frac{2}{N} \simeq 0.016$ and eight distance partitions as $\Delta\mathcal{D} = 1/(10\sqrt{2}) < \frac{2}{(N-1)^2\Delta d}$. We define the far-field region as $\mu \in [0.01, \frac{3}{4}\Delta\mathcal{D}]$, which is 18.8 m to 100 m. The rest distance partition is $\mu_p \in [(p - \frac{1}{4})\Delta\mathcal{D}, (p + \frac{3}{4})\Delta\mathcal{D}], p = 1, \ldots, 7$. Based on such a spatial partition, our codebook contains eight levels of subcodebook. The upper seven levels are the AAP, and the last is DAP. The beam will search the AAP for seven iterations and the DAP for eight iterations.

The transition band in (23) is set as 0 for both angle and distance dimension, which means there is no blank area between the coverage areas of neighboring beams. It will enable the codebook to achieve high transmission rates. The sampling interval in (24) is defined as $0.1/N$ in angle and $\frac{1}{200\sqrt{2}}$ in distance. Besides, the beam is demonstrated with the normalized magnitude response given by

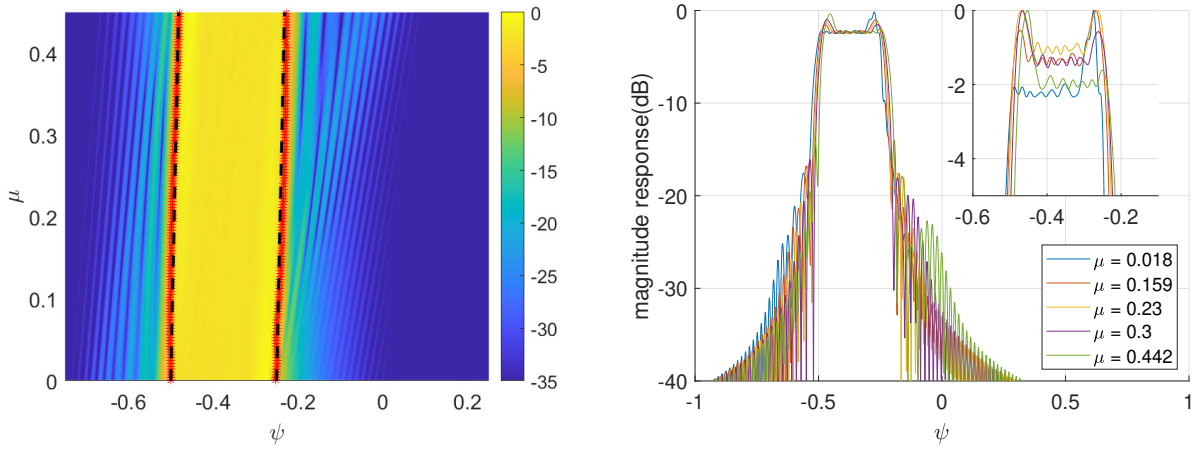
$$B_{\mathbf{f}}(\mu, \psi) = 20\log\left(\frac{\left|\mathbf{a}^H(\mu, \psi)\mathbf{f}\right|}{\max\limits_{\mu\in[0,0.45],\psi\in[-1,1]} |\ \mathbf{a}^H(\mu, \psi)\mathbf{f}\ |}\right). \tag{28}$$

A. Distance-coarse Beam

In this subsection, we illustrate our designed distance-coarse beams in the AAP by taking $\mathbf{f}_{3,3}^{(\mathrm{A})}$ as an example, whose coverage area is $\psi \in [-0.5, -0.25]$ at the far-field and show it's beam pattern in Fig. 6.

Fig. 6(a), we show the beam pattern with a surface plot. The dark dashed lines represent the constrained mainlobe area in (14), and the red asterisks represent the -3dB width. From this pic,

one can observe that the beam's 3 dB width satisfies the boundary, which verifies the accuracy of Theorem 2. But, the 3 dB beam width at μ =0.25 has slightly exceeded the constrained boundary but gradually narrowed when $\mu = 0$ and $\mu = 0.45$. This indicates that the beams broaden the mainlobe area at $\mu = 0.23$ to keep the magnitude response in every distance as high as possible. To more clearly illustrate the beam pattern, we plot the beam pattern of $\mathbf{f}_{3,3}^{(\mathrm{A})}$ at distances $\mu = 0.018, 0.159, 0.23, 0.3, 0.442$ in Fig. 6(b). It can be observed that the main lobe keeps higher than -3 dB. Meanwhile, the magnitude response of the main lobe has an obvious Gibbs phenomenon. The ripple at the area, except the Gibbs peak, is less than 0.1 dB.

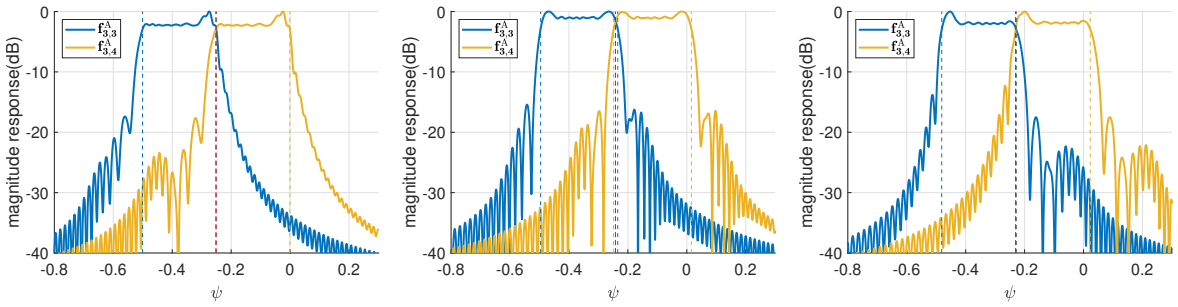


(a) The beam pattern in the service area

(b) The AAP in several selected distance

Fig. 6: Instance for AAP: the beam pattern of angle alignment beam $\mathbf{f}_{3,3}^{(\mathrm{A})}$

Next, in Fig. 7, we show the normalized beam patterns of beam pair $\mathbf{f}_{3,3}^{(\mathrm{A})}$, $\mathbf{f}_{3,4}^{(\mathrm{A})}$ to illustrate of the transition band in AAP. $\mathbf{f}_{3,3}^{(\mathrm{A})}$ covers $\psi \in [-0,5,-0.25]$ and $\mathbf{f}_{3,4}^{(\mathrm{A})}$ covers $\psi \in [-0,25,0]$. The three plots correspond to the beam pattern when $\mu = 0.018$ (far-field), $\mu = 0.23$ and $\mu = 0.37$ (near-field). In these figures, the beam pattern of $\mathbf{f}_{3,3}^{(\mathrm{A})}$ and $\mathbf{f}_{3,4}^{(\mathrm{A})}$ are shown in the blue and yellow line, respectively. The 3 dB width of the two beams is shown in dashed lines with the corresponding color. The red dashed line represents the dividing line of two beams. In Fig. 7 (a) and Fig. 7(c), the dividing line is overlapped with two 3 dB boundaries. In Fig. 7(b), the dividing line is inside two beams' coverage area. Hence, it can be concluded that the designed beams perfectly cover their area and leave no blank area in the transition band. Meanwhile, we can observe that the beam patterns in Fig. 7(a) and Fig. 7(c) are not symmetrical. Namely, the decays at the transition band for the two beams are not the same.



(a) The AAP of beam pair $\mathbf{f}^{(\mathrm{A})}_{3,3}$, $\mathbf{f}^{(\mathrm{A})}_{3,4}$ with $\mu = 0.018$

(b) The AAP of beam pair $\mathbf{f}^{(\mathrm{A})}_{3,3}$, $\mathbf{f}^{(\mathrm{A})}_{3,4}$ with $\mu = 0.23$

(c) The AAP of beam pair $\mathbf{f}^{(\mathrm{A})}_{3,3}$, $\mathbf{f}^{(\mathrm{A})}_{3,4}$ with $\mu = 0.37$

Fig. 7: The illustration of transition band

Also, in Fig. 8, we show the normalized magnitude response of beams with different resolutions. It is well known that an antenna array reaches high resolution and high gain when it focuses on a small area. Observing this figure, the coarse beams allocate the total energy evenly in the coverage area, which makes the higher-level beams have lower gain compared to the lower-level beams. Furthermore, it can be observed from this figure that the two child beams can cover the entire area of their parent beam.

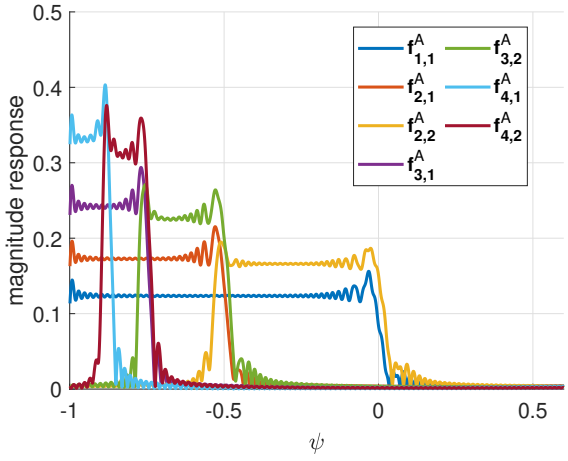
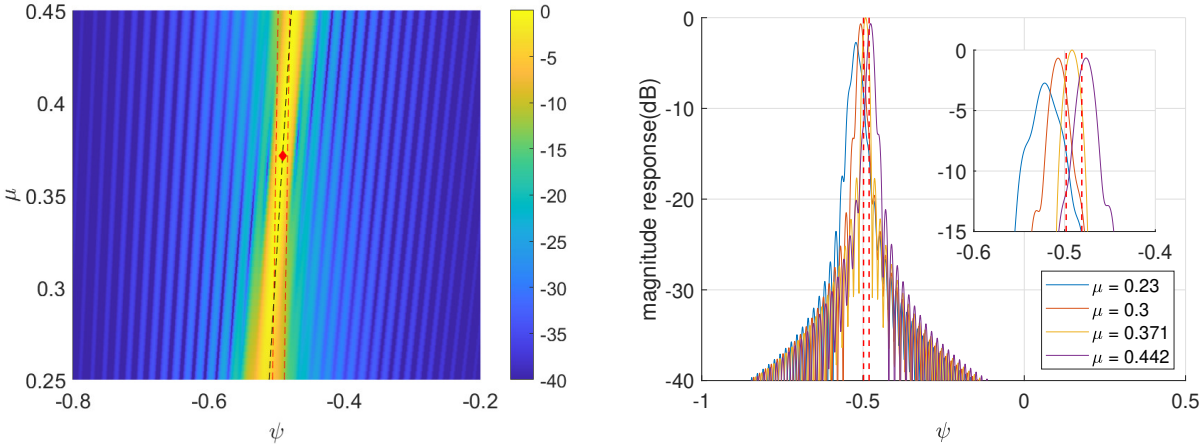


Fig. 8: The magnitude response of normalized beams with different resolutions

B. Focusing Beams

In this subsection, we present the beam pattern of focusing beams by taking an example of $\mathbf{f}^{(\mathrm{D})}_{6,32}$, which focuses on $\mu = 0.37$ and $\psi = -0.49$. Fig .9(a) shows the normalized beam pattern

of the whole area. In this figure, we show the coverage area of $\mathbf{f}_{6,32}^{(\mathrm{A})}$ with the red dashed line, which is the root node of $\mathbf{f}_{6,32}^{(\mathrm{D})}$. The symmetrical line in (13) is marked with the black dashed line, and the maximum peaks are marked with the red-filled diamond. It can be observed that the black dashed line is in the center of the main lobe, which verifies Theorem 1. Meanwhile, the red and black lines have different slop, which coincides with (13) and (14). In Fig. 9(b), we show the beam pattern at $\mu = 0.23, 0.3, 0.371$ and 0.442, which are the distance focusing point of the fourth to the seventh distance partition. From this figure, it is obvious that the 3 dB beam width covers the given angle. The maximum peak is between the two boundaries.



(a) The beam pattern in the service area

(b) The AAP in several selected distance

Fig. 9: Instance for DAP: beam pattern of distance searching beam $\mathbf{f}_{7,32}^{(\mathrm{D})}$

Then, we show the coverage relation between the last level of AAP (the seventh level subcodebook) and the DAP (the eighth level subcodebook) by presenting the beam pattern of focusing beams $\mathbf{f}_{i,32}^{(\mathrm{A})}$ $(i = 1, \cdots, 7)$ that are under cover of $\mathbf{f}_{7,32}^{(\mathrm{A})}$ in Fig. 10. In this figure, the beam patterns are spliced together. The white dashed lines are the boundary of the distance partitions, and the coverage area of $\mathbf{f}_{7,32}^{(\mathrm{A})}$ is between the red dashed lines. It can be observed that the designed focusing beams perfectly cover the area of the upper-level root node.

We also present the magnitude response at the line $\psi = \frac{1}{2}(N-1)\Delta d\mu + 0.5$ with a distance-magnitude response plot in Fig. 11. In this figure, we show the normalized magnitude response of $\mathbf{f}_{7,32}^{(\mathrm{A})}$ with the red dashed line. The distance boundary is represented by the dark dashed lines. The normalized magnitude responses of focusing beams are distinguished with different colors. From this figure, we can observe that the focusing beams perfectly cover the distance-coarse

beam's area. Under their coverage distance area, the magnitude response is higher than -1 dB. Their boundaries perfectly match the partitioned line except for the far-field area. The far-field area beam shows a wider mainlobe than the others because it has more sample points when $\mu \to 0$.

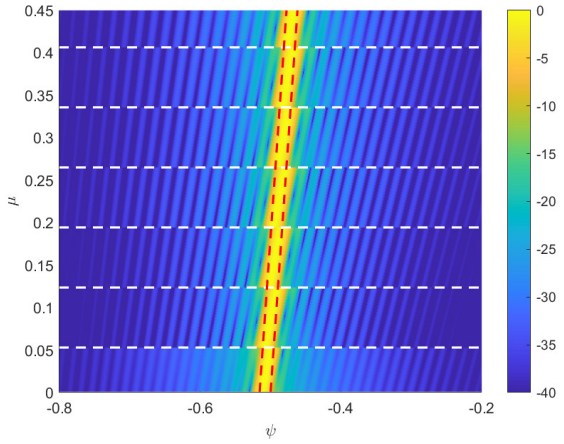


Fig. 10: The beam pattern of focusing beam set $\{\mathbf{f}_{k,32}^{(\mathrm{D})}\ (k = 1, \cdots, 7)\}$ under $\mathbf{f}_{7,32}^{(\mathrm{A})}$

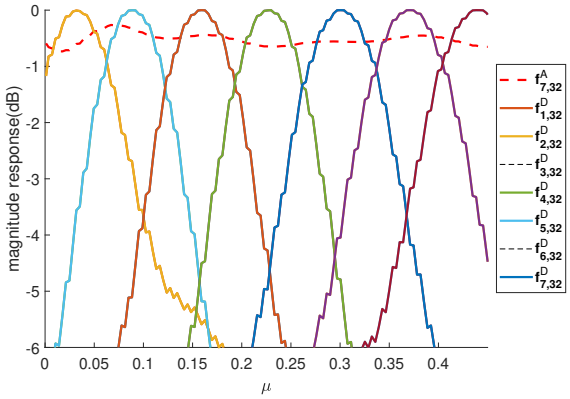


Fig. 11: The distance-magnitude response of the beams that are under cover of $\mathbf{f}_{7,32}^{(\mathrm{A})}$

C. BA Performance

In this subsection, the BAER and spectral efficiency are plotted in the scenario that the user is located in $\mu \in [0.01, 0.477]$ and $\psi \in [-1, 1]$ and the SNR is ranged from -10 dB to 20 dB. The presented BAER and the achievable rate are the average results of the 15000 times Monte Carlo experiment.

We compare our algorithm with the following benchmark algorithms:

- Evenly Distance Partition Codebook (EDPC): The space is partitioned evenly in ψ and r dimensions. The beams are the array response vector steering to each partition. The BA scheme utilizes an exhausting search.
- Unevenly Distance Partition Codebook (UDPC): The codebook is constructed similarly to EDPC, but the space partition follows (11).
- Hierarchy Structure Codebook (HSC): The codebook and BA scheme were proposed in [25]. We adopt the binary tree in the angle dimension but still adopt an exhausting search in the distance dimensions.
- Two Stage Codebook (TSC): The BA scheme was proposed in [27], in which they utilized the codebook proposed in [25].
- Distance Searching Codebook (DSC): This codebook is the distance alignment part of our proposal codebook with the exhausting search algorithm.
- XL-MIMO Hierarchy Codebook (XL-HC): Our proposal codebook with hierarchy searching.

First, we present the BAER performance of different codebooks in Fig. 12. We define the BA error as the scenario when the user is not located at the coverage area of the chosen beam. Then, the BAER is defined as the ratio between the BA error times and the total BA times. It can be observed from Fig. 12 that EDPC shows higher BAERs than UDPC, which indicates our partition's effectiveness. DSC has the lowest angle BAER in all codebooks. XL-HC shows lower BAER than EDPC, TSC, and HSC. Meanwhile, XL-HC has better performance than UDPC when SNR$>$5 dB. Due to the hierarchical architecture, a slight performance loss can be observed by comparing the DSC and XL-HC.

Finally, we compare the spectral efficiency of different codebooks in Fig. 13. It is observed from Fig. 13 that the TSC achieves poor performance because it does not align the beam. Except for TSC, the spectral efficiency of other algorithms are very close, although they have significant differences in BAER. As expected, the EDPC shows a small performance loss compared to the UDPC, which is coincident with the conclusion in Fig. 12. Our proposed DSC achieves the best spectral efficiency among all the algorithms.

V. Conclusion

In this paper, we proposed a hierarchical codebook and an efficient BA algorithm for extremely large-scale antenna array MIMO wireless communication systems, which are capable of achiev-

ing a low BA error rate and maintaining low BA overheads. Specifically, we revealed the angle-offset effect of the near-field beam and proposed a novel near-field spatial partition. Based on this, we designed distance-coarse beams covering all distance-dimension partitions and focusing beams covering only one partition. We utilized the distance-coarse beams to construct the high level of codebooks for angle alignment and the focusing beams to construct the last level of codebooks for distance alignment. Accordingly, we proposed a two-stage tree search BA scheme. We formulated the distance-coarse and focusing beams design as difference convex optimization problems taking into account the three beam design guidelines. The constrained concave-convex procedure (CCCP) was applied to solve the corresponding non-convex problems. Numerical results showed that our proposed codebook outperformed the existing benchmark schemes in terms of BAER performance with significantly reduced overheads.

APPENDIX A
PROOF OF THEOREM 1

Proof: Theorem 1 presents the coverage area effect of a focusing beam $\mathbf{f}$ that focus on the points $(\mu_{\mathbf{f}}, \psi_{\mathbf{f}})$. Firstly, we define the array response matrix as

$$\mathbf{A}(\mu) = [\mathbf{a}(\mu, -1), \mathbf{a}(\mu, -1 + 2/N), \cdots, \mathbf{a}(\mu, 1)] \in \mathbb{C}^{N \times N}, \tag{29}$$

Then, we define the beam pattern vector at $\mu_{\mathbf{f}}$ as $\mathbf{B}^{(\mu_{\mathbf{f}})} = \mathbf{A}(\mu_{\mathbf{f}})^{\mathrm{H}}\mathbf{f}$, which represents the beam pattern is sampled in the $\boldsymbol{\psi} = [-1, -1 + 2/N, \cdots, 1]$. As $\mathbf{A}(\mu)$ is a unitary matrix, i.e.,

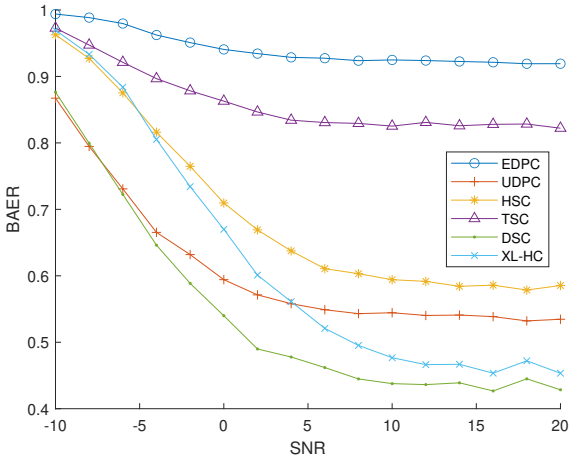


Fig. 12: The BAER performances of different codebook

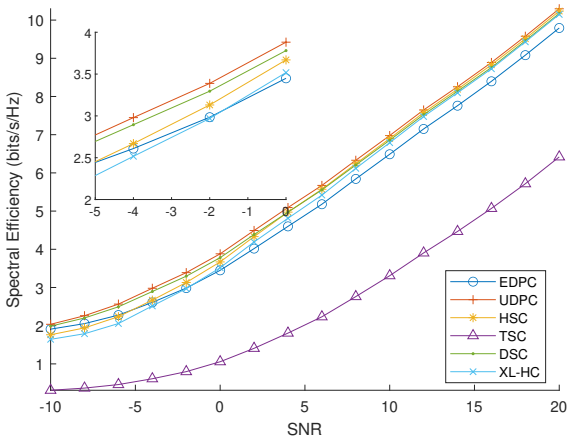


Fig. 13: The spectral efficiency of different codebook

$\mathbf{A}(\mu)\mathbf{A}(\mu)^{\mathrm{H}} = \mathbf{I}$ [41], the beam pattern can be represented as

$$B(\mu,\psi) = \frac{1}{N}\mathbf{a}(\mu,\psi)^{\mathrm{H}}\mathbf{A}(\mu_{\mathbf{f}})\mathbf{A}(\mu_{\mathbf{f}})^{\mathrm{H}}\mathbf{f} = \frac{1}{N}\mathbf{T}^{\mathrm{H}}(\mu,\psi)\mathbf{B}^{(\mu)}, \tag{30}$$

where $\mathbf{T}(\mu,\psi) = \mathbf{A}(\mu_{\mathbf{f}})^{\mathrm{H}}\mathbf{a}(\mu,\psi)$. We denote $\mathbf{T}_v(\mu,\psi)$ as the v-th element of $\mathbf{T}(\mu,\psi)$ given by

$$\mathbf{T}_v(\mu,\psi) = \mathbf{a}\left(\mu,\psi\right)^{\mathrm{H}}\mathbf{a}\left(\mu_{\mathbf{f}},\psi_v\right) \simeq \sum_{n=0}^{N-1}\exp\left(jk\left(\frac{1}{2}(n\Delta d)^2\Delta\mu - n\Delta d\Delta\psi\right)\right), \tag{31}$$

where $\Delta\mu = \mu - \mu_{\mathbf{f}}$, and $\Delta\psi = \psi - \psi_v$. $\psi_v = -1 + \frac{2(v-1)}{N}$. It can be simplified by:

$$\begin{aligned}
\mathbf{T}_v(\mu,\psi) &\simeq \sum_{l=0}^{\inf}\sum_{n=0}^{N-1}\frac{\left(jk\left(\frac{1}{2}\left(n\Delta d\right)^2\Delta\mu - n\Delta d\Delta\psi\right)\right)^l}{l!} \\
&= \sum_{l=0}^{\inf}\frac{(jk)^l\sum_{n=0}^{N-1}\sum_{m=0}^{l}C_l^m\left(\frac{1}{2}\Delta\mu\right)^m(-\Delta\psi)^{l-m}(n\Delta d)^{l+m}}{l!} \\
&\overset{(a)}{\simeq} \sum_{l=0}^{\inf}\frac{(jk)^l\sum_{m=0}^{l}\frac{1}{\Delta d}\int_0^{(N-1)\Delta d}C_l^m\left(\frac{1}{2}\Delta\mu\right)^m(-\Delta\psi)^{l-m}x^{l+m}dx}{l!} \\
&= \sum_{l=0}^{\inf}\frac{(jk)^l((N-1))^{l+1}\Delta d^l}{l!}\cdot\sum_{m=0}^{l}\frac{C_l^m\left(\frac{(N-1)\Delta d\Delta\mu}{2}\right)^m(-\Delta\psi)^{l-m}}{l+m+1} \\
&\overset{(b)}{\simeq} \sum_{l=0}^{\inf}\frac{(jk)^l((N-1))^{l+1}\Delta d^l(\frac{(N-1)\Delta d\Delta\mu}{2} - \Delta\psi)^l}{l!(l+1)}
\end{aligned} \tag{32}$$

where the approximation (a) utilizes the discrete sum to approximate the continuous integration and the approximation (b) is achieved by $\sum_{m=0}^{l}\frac{C_l^m(l+1)}{l+m+1} = 1$. It should be noticed that the two

approximations only hold in lower-order items. Hence, we have $\mathbf{T}_v(\mu,\psi) \simeq 1$ when $|\Delta\mu| \ll 1$, $|\Delta\psi| \ll 1$ and

$$\Delta\psi = \frac{1}{2}(N-1)\Delta d\Delta\mu. \tag{33}$$

Without loss of generality, we assume $\mathbf{B}^{(\mu_\mathbf{f})}$ only have one nonzero element at v_f, i.e., $\mathbf{B}^{(\mu_\mathbf{f})}[v_f] = 1$. Thereby, we have

$$\begin{aligned} B(\mu,\psi) &= \frac{1}{N}\mathbf{T}^{\mathrm{H}}(\mu,\psi)\mathbf{B}^{(\mu_\mathbf{f})} = \frac{1}{N}\mathbf{T}_{v_f}(\mu,\psi) \\ &\simeq \begin{cases} 1 & \psi = \psi_{v_f} + \frac{1}{2}(N-1)\Delta d(\mu-\mu_0) \\ 0 & \text{other} \end{cases}, \end{aligned} \tag{34}$$

which represents that the main lobe is symmetrical to the line (33).

Next, we show the coverage area of the focusing beam. By substituting (33) into $\mathbf{T}_{v_f}(\mu,\psi)$, we have

$$\mathbf{T}_{v_f}(\mu,\psi) \simeq \sum_{n=0}^{N-1} \exp\left(jk\frac{(n^2-n(N-1))\Delta d^2\Delta\mu}{2}\right) = \sum_{n=0}^{N-1} \exp\left(jk(\frac{n^2}{N-1}-n)\Delta d\Delta\psi\right). \tag{35}$$

In the above equation, it is hard to find the solution for $|\mathbf{T}_\mathbf{f}| = \frac{1}{\sqrt{2}}$. But it is definitely that $|\mathbf{T}_\mathbf{f}| < \frac{1}{\sqrt{2}}$, when $\Delta\mu = \frac{2}{(N-1)^2\Delta d}$ and $\Delta\psi = \frac{1}{N-1}$. Thereby, we have the coverage area for a given $\mathbf{f}$ that focus on the points $(\mu_\mathbf{f},\psi_\mathbf{f})$ have the coverage area, given by

$$\mathbb{M}_\mathbf{f} = \left\{(\mu,\psi)||\mu-\mu_\mathbf{f}| < \hat{\mu}, |\psi-\psi_\mathbf{f}+\frac{1}{2}(N-1)\Delta d(\mu-\mu_\mathbf{f})| < \hat{\psi}\right\}, \tag{36}$$

where $\hat{\mu} < \frac{2}{(N-1)^2\Delta d}$ and $\hat{\psi} < 1/(N-1)$. Theorem 1 is proven.

■

APPENDIX B
PROOF OF THEOREM 2

We prove the effect for a distance-coarse beam $\mathbf{f}_w$ that covers all distance partitions. Without loss of generality, we assume $\mathbf{B}^{(0)}[p_f] = 1, (p_l < p_f < p_h)$, which means it covers the angle between ψ_{p_l} and ψ_{p_h} in far-field. Then, we represent the beam pattern at the other distance as

$$B(\mu,\psi) \simeq \sum_{p=p_l}^{p_h} \frac{1}{N}\mathbf{T}^{\mathrm{H}}_{(\mu,\psi)}\mathbf{B}^{(0)} = \frac{1}{N}\sum_{p=p_l}^{p_h}\sum_{n=0}^{N-1} \exp\left(jk\left(\frac{1}{2}(n\Delta d)^2(\mu-0) - n\Delta d(\psi-\psi_p)\right)\right), \tag{37}$$

We first focus on the boundary corresponding to ψ_{p_l}. Similar to Theorem 1, we assume the relationship between μ and ψ for 3 dB boundary can be approximated by a linear function, i.e.,

$\Delta\psi = \psi - \psi_l = \kappa(N-1)\Delta d(\mu - 0)$. Then, the beam response on the boundary line can be rewritten as

$$B(\Delta\psi) = \frac{1}{N}\sum_{p=0}^{\Delta p}\sum_{n=0}^{N-1}\exp\left(jk\Delta d\left(\frac{1}{2}\frac{\kappa\Delta\psi}{N-1}n^2 - (\Delta\psi - \frac{2p}{N})n\right)\right), \tag{38}$$

where $\Delta p = \frac{p_h - p_l}{2}$. (38) has the definitional domain $0 \leq \Delta\psi \leq \frac{\kappa(N-1)\Delta d}{r_{\mathrm{Fre}}}$, where $r_{\mathrm{Fre}} = \sqrt[3]{\frac{D^4}{8\lambda}} \simeq \frac{N\Delta d}{2}$ represents the Fresnel distance. We consider a common scenario that $\Delta d = \frac{\lambda}{2}$ and $k\Delta d = \pi$. In such a scenario, the minimum value is obtained at the boundary. Thus, we only need to ensure $\left|B(\frac{\kappa(N-1)\Delta d}{r_{\mathrm{Fre}}})\right| > 1/\sqrt{2}$ to let $|B(\mu,\psi)| > 1/\sqrt{2}$ in the defined area. Thereby, we have the solution of κ from the following equation

$$B(2\kappa) = \frac{1}{N}\sum_{p=0}^{\Delta p}\sum_{n=0}^{N-1}\exp\left(j\pi\left(\frac{\kappa^2}{N}n^2 - (2\kappa - \frac{2p}{N})n\right)\right) = 1/\sqrt{2}. \tag{39}$$

In summary, the coverage area of the distance-coarse beam

$$\mathbb{M}_{\mathbf{f}} = \left\{(\mu,\psi)|0 < \mu < \frac{1}{r_{\mathrm{Fre}}}, \kappa(N-1)\Delta d\mu + \psi_l < \psi < \kappa(N-1)\Delta d\mu + \psi_h\right\} \tag{40}$$

Theorem 2 is proved.

REFERENCES

[1] M. Shafi, A. F. Molisch, P. J. Smith, T. Haustein, P. Zhu, P. De Silva, F. Tufvesson, A. Benjebbour, and G. Wunder, "5g: A tutorial overview of standards, trials, challenges, deployment, and practice," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 6, pp. 1201–1221, 2017.

[2] E. Björnson, L. Sanguinetti, H. Wymeersch, J. Hoydis, and T. L. Marzetta, "Massive MIMO is a reality—What is next?" *Digital Signal Processing*, vol. 94, pp. 3–20, Nov. 2019.

[3] E. D. Carvalho, A. Ali, A. Amiri, M. Angjelichinoski, and R. W. Heath, "Non-stationarities in extra-large-scale massive MIMO," *IEEE Wireless Communications*, vol. 27, no. 4, pp. 74–80, Aug. 2020.

[4] H. Lu and Y. Zeng, "How does performance scale with antenna number for extremely large-scale MIMO?" in *ICC 2021 - IEEE International Conference on Communications*, Jun. 2021, pp. 1–6.

[5] Lu, Haiquan and Zeng, Yong, "Near-field modelling and performance analysis for multi-user extremely large-scale MIMO communication," *IEEE Communications Letters*, pp. 1–1, 2021.

[6] N. Decarli and D. Dardari, "Communication Modes With Large Intelligent Surfaces in the Near Field," *IEEE Access*, vol. 9, pp. 165 648–165 666, 2021.

[7] H. Lu and Y. Zeng, "Communicating with extremely large-scale array/surface: Unified modelling and performance analysis," *IEEE Transactions on Wireless Communications*, 2021.

[8] W. Tang, M. Z. Chen, X. Chen, J. Y. Dai, Y. Han, M. Di Renzo, Y. Zeng, S. Jin, Q. Cheng, and T. J. Cui, "Wireless communications with reconfigurable intelligent surface: Path loss modeling and experimental measurement," *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, pp. 421–439, 2021.

[9] D. Dardari, "Communicating with large intelligent surfaces: Fundamental limits and models," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 11, pp. 2526–2537, 2020.

[10] M. Cui and L. Dai, "Channel estimation for extremely large-scale MIMO: Far-Field or Near-Field?" *arXiv:2108.07581 [cs, math]*, Dec. 2021.

[11] Y. Han, S. Jin, C.-K. Wen, and X. Ma, "Channel estimation for extremely large-scale massive MIMO systems," *IEEE Wireless Communications Letters*, vol. 9, no. 5, pp. 633–637, 2020.

[12] X. Zhang, Z. Wang, H. Zhang, and L. Yang, "Near-field channel estimation for extremely large-scale array communications: A model-based deep learning approach," *arXiv preprint arXiv:2211.15440*, 2022.

[13] M. Cui, L. Dai, R. Schober, and L. Hanzo, "Near-field wideband beamforming for extremely large antenna arrays," *arXiv:2109.10054 [cs, eess, math]*, Dec. 2021.

[14] H. Zhang, N. Shlezinger, F. Guidi, D. Dardari, M. F. Imani, and Y. C. Eldar, "Beam focusing for near-field multi-user MIMO communications," *IEEE Transactions on Wireless Communications*, 2022.

[15] A. Kay, "Near-field gain of aperture antennas," *IRE Transactions on Antennas and Propagation*, vol. 8, no. 6, pp. 586–593, Nov. 1960.

[16] R. Hansen, "Focal region characteristics of focused array antennas," *IEEE Transactions on Antennas and Propagation*, vol. 33, no. 12, pp. 1328–1337, Dec. 1985.

[17] 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," *arXiv preprint arXiv:2203.13035*, 2022.

[18] I. F. Akyildiz, C. Han, Z. Hu, S. Nie, and J. M. Jornet, "Terahertz band communication: An old problem revisited and research directions for the next decade," *IEEE Transactions on Communications*, 2022.

[19] H. Shokri-Ghadikolaei, L. Gkatzikis, and C. Fischione, "Beam-searching and transmission scheduling in millimeter wave communications," in *2015 IEEE international conference on communications (ICC)*, 2015, pp. 1292–1297.

[20] T. Nitsche, A. B. Flores, E. W. Knightly, and J. Widmer, "Steering with eyes closed: mm-wave beam steering without in-band measurement," in *2015 IEEE Conference on Computer Communications (INFOCOM)*, 2015, pp. 2416–2424.

[21] C. K. Au-yeung and D. J. Love, "On the performance of random vector quantization limited feedback beamforming in a MISO system," *IEEE Transactions on Wireless Communications*, vol. 6, no. 2, pp. 458–462, 2007.

[22] A. W. Shaban, O. Damen, Y. Xin, and E. Au, "Statistically-Aided Codebook-Based Hybrid Precoding for mmWave Single-User MIMO Systems," in *2020 IEEE International Conference on Communications Workshops (ICC Workshops)*, Jun. 2020, pp. 1–6.

[23] J. Zhang, Y. Huang, Q. Shi, J. Wang, and L. Yang, "Codebook Design for Beam Alignment in Millimeter Wave Communication Systems," *IEEE Transactions on Communications*, vol. 65, no. 11, pp. 4980–4995, 2017.

[24] Z. Xiao, P. Xia, and X.-G. Xia, "Codebook Design for Millimeter-Wave Channel Estimation With Hybrid Precoding Structure," *IEEE Transactions on Wireless Communications*, vol. 16, no. 1, pp. 141–153, 2017.

[25] 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 Communications*, vol. 19, no. 6, pp. 193–204, 2022.

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

[27] M. Cui, L. Dai, Z. Wang, S. Zhou, and N. Ge, "Near-field rainbow: Wideband beam training for XL-MIMO," *IEEE Transactions on Wireless Communications*, pp. 1–1, 2022.

[28] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 831–846, 2014.

[29] J. Yang, Y. Zeng, S. Jin, C.-K. Wen, and P. Xu, "Communication and localization with extremely large lens antenna array," *IEEE Transactions on Wireless Communications*, vol. 20, no. 5, pp. 3031–3048, 2021.

[30] Y. Wang, M. Narasimha, and R. W. Heath, "MmWave Beam Prediction with Situational Awareness: A Machine Learning Approach," in *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. Kalamata: IEEE, Jun. 2018, pp. 1–5.

[31] W. Tang, M. Z. Chen, J. Y. Dai, Y. Zeng, X. Zhao, S. Jin, Q. Cheng, and T. J. Cui, "Wireless communications with programmable metasurface: New paradigms, opportunities, and challenges on transceiver design," *IEEE Wireless Communications*, vol. 27, no. 2, pp. 180–187, Apr. 2020.

[32] F. Bohagen, P. Orten, and G. E. Oien, "Design of optimal high-rank line-of-sight MIMO channels," *IEEE Transactions on Wireless Communications*, vol. 6, no. 4, pp. 1420–1425, 2007.

[33] Bohagen, Frode and Orten, Pal and Oien, Geir E., "On spherical vs. plane wave modeling of line-of-sight MIMO channels," *IEEE Transactions on Communications*, vol. 57, no. 3, pp. 841–849, 2009.

[34] I. Ahmed, H. Khammari, A. Shahid, A. Musa, K. S. Kim, E. De Poorter, and I. Moerman, "A survey on hybrid beamforming techniques in 5G: Architecture and system model perspectives," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 3060–3097, 2018.

[35] D. Zhang, Y. Wang, X. Li, and W. Xiang, "Hybridly connected structure for hybrid beamforming in mmwave massive MIMO systems," *IEEE Transactions on Communications*, vol. 66, no. 2, pp. 662–674, 2017.

[36] J. R. Barry, E. A. Lee, and D. G. Messerschmitt, *Digital communication*. Springer Science & Business Media, 2012.

[37] K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang, "On beamforming with finite rate feedback in multiple-antenna systems," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2562–2579, 2003.

[38] D. Love, R. Heath, and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Transactions on Information Theory*, vol. 49, no. 10, pp. 2735–2747, 2003.

[39] H. A. Le Thi and T. Pham Dinh, "Dc programming and dca: thirty years of developments," *Mathematical Programming*, vol. 169, no. 1, pp. 5–68, 2018.

[40] A. L. Yuille and A. Rangarajan, "The concave-convex procedure," *Neural computation*, vol. 15, no. 4, pp. 915–936, 2003.

[41] H. L. Van Trees, *Optimum array processing: Part IV of detection, estimation, and modulation theory*. John Wiley & Sons, 2002.