Near-Field 2D Hierarchical Beam Training for Extremely Large-Scale MIMO

Yu Lu, *Student Member, IEEE*, Zijian Zhang, *Student Member, IEEE*, and Linglong Dai, *Fellow, IEEE*

Department of Electronic Engineering, Tsinghua University,
Beijing National Research Center for Information Science and Technology (BNRist), Beijing 100084, China
Email: y-lu19@mails.tsinghua.edu.cn, zhangzj20@mails.tsinghua.edu.cn, daill@tsinghua.edu.cn

***Abstract*—The evolution of multi-input multi-output (MIMO) will develop toward extremely large-scale MIMO (XL-MIMO) for future 6G communications. With the extension of the antenna array, the electromagnetic propagation change from far-field to near-field. Because of the near-field effect, the exhaustive near-field beam training scanning all angles and distances involves very high overhead. The existing fast near-field beam training scheme with extra time-delay circuits can reduce the overhead, but it suffers from very high hardware costs and energy consumption. To solve this issue, we propose a low-overhead near-field two dimension (2D) hierarchical beam training after carefully designing the near-field multi-resolution codebooks. Specifically, we first formulate the problem of designing near-field multi-resolution codewords, which have various angle coverage and distance coverage. Next, we propose a Gerchberg-Saxton (GS)-based algorithm to obtain the theoretical codeword by considering the ideal fully digital architecture, and an alternating optimization algorithm is then proposed to acquire the practical codeword by considering the hybrid digital-analog architecture. Finally, we generate multi-resolution codebooks and propose a near-field 2D hierarchical beam training scheme. Simulation results demonstrate that the proposed scheme can provide a tradeoff between the achievable rate performance and overhead in near-field XL-MIMO beam training.**

***Index Terms*—Extremely large-scale MIMO, beam training, codebook design.**

I. Introduction

With the emergence of new applications, such as digital twins, 6G is expected to achieve a 10-fold increase in spectrum efficiency than 5G [1]. The extremely large-scale MIMO (XL-MIMO) is a promising technique for 6G to achieve ultra-high spectrum efficiency [2]. In XL-MIMO systems, the base station (BS) deploys an extremely large-scale antenna (ELAA) to obtain the spatial multiplexing gain by reliable beamforming. Before beamforming, beam training should be conducted to search the optimal beamforming vector, i.e., codeword, in the predefined codebook.

There are two typical categories of beam training methods for MIMO, which are respectively far-field beam training and near-field beam training. For the first category, since the antenna number is not very large in 5G systems, the MIMO channel is modeled in the far-field region with the planar wave assumption, where the array response vector of the far-field channel is only related to the angle. In this case, the Discrete Fourier Transform (DFT) codebook can be utilized to conduct beam training to capture the physical angle information in the angle-domain of the channel paths. The size of the DFT codebook is proportional to the number of antennas at BS. Since XL-MIMO has a large number of antennas, the DFT codebook suffers from very high beam training overhead. Thus, to reduce the beam training overhead, some hierarchical beam training schemes were proposed [3], [4]. The basic idea is to search from the lowest-resolution sub-codebook to the highest-resolution sub-codebook in turn, where the angle range needed to be scanned reduces layer by layer gradually. With the help of hierarchical beam training, the overhead is only proportional to the logarithm of the antenna number at BS [4].

As the antenna number dramatically increases in 6G XL-MIMO systems, the near-field range will expand by orders of magnitude, which can be up to several hundreds of meters [5]. Thus, the XL-MIMO channel should be modeled in the near-field region with the spherical wave assumption. For the second category, i.e., near-field beam training, the array response vector of the near-field channel is not only related to the angle but also to distance. Thus, to capture the physical angle as well as distance information of the channel paths, a polar-domain codebook should be utilized instead of a DFT codebook [6]. Accordingly, the size of the polar-domain codebook is the product of the antenna number at BS and the number of sampled distances. Since only one angle and one distance can be measured in each time slot, the exhaustive search method for near-field beam training has a very high overhead [7]. To address this problem, we have proposed a fast time-delay based near-field beam training for XL-MIMO with low overhead [8], where each antenna requires a time-delay to provide frequency-dependent phase shift. In specific, thanks to the near-field beam split effect in a wide band, near-field beams can be flexibly controlled by extra time-delays hardware circuits and then focus on different angles and distances at different frequencies in one time slot. However, the time-delay based beamforming structure will lead to not only high cost but also very high energy consumption, especially for XL-MIMO systems with a large number of antennas.

Thus, to design a low-overhead beam training scheme, we propose a near-field two dimension (2D) hierarchical beam training scheme by designing the multi-resolution codebooks referring to the hierarchical beam training in the far-field scenario. Specifically, we first formulate the problem of near-field multi-resolution codeword design. Compared with the far-field case, the ideal beam pattern of near-field codeword

979-8-3503-1090-0/23/$31.00

should not only cover a certain angle range but also a certain distance range. By considering ideal fully digital architecture and practical hybrid digital-analog structure, we provide the design problem of the near-field theoretical codeword and practical codeword. Then, inspired by the Gerchberg–Saxton (GS) algorithm in phase retrieval problems in the digital hologram imaging, we first proposed a GS-based theoretical codeword design algorithm under a fully digital architecture assumption. Next, based on the designed theoretical codeword, an alternating optimization algorithm is proposed to acquire the practical codeword. With the aid of multi-resolution codebooks with different angle coverages and distance coverages, we propose a near-field two dimension (2D) hierarchical beam training scheme. where codewords are searched in multi-resolution codebooks layer by layer. Finally, we provide numerical simulation results to illustrate that the proposed beam training scheme can reach sub-optimal achievable rate performance with low overhead.

Notations: Lower-case and upper-case boldface letters $\mathbf{a}$ and $\mathbf{A}$ denote a vector and a matrix, respectively; $\mathbf{a}^H$ and $\mathbf{A}^H$ denote the conjugate transpose of vector $\mathbf{a}$ and matrix $\mathbf{A}$, respectively; $\|\mathbf{a}\|_2$ denotes the l_2 norm of vector $\mathbf{a}$; $\|\mathbf{a}\|_F$ denotes the Frobenius norm of vector $\mathbf{a}$. Finally, $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the probability density function of complex multi-variate Gaussian distribution with mean $\boldsymbol{\mu}$ and variance $\boldsymbol{\Sigma}$. $\mathcal{U}(-a, a)$ denotes the uniform distribution on $(-a, a)$.

II. System Model

In this section, we will first introduce the signal model of the XL-MIMO system. Then, the existing near-field channel model will be briefly reviewed.

A. Signal Model

We consider the scenario where the BS employs a N-element ELAA to communicate with a single-antenna user. Let $\mathbf{h}^H \in \mathbb{C}^{1\times N}$ denote the channel from the BS to the user. Take downlink transmission as example, the received signal y can be represented by

$$y = \mathbf{h}^H \mathbf{v} s + n, \tag{1}$$

where $\mathbf{v} \in \mathbb{C}^{N\times 1}$ represents the beamforming vector at the BS, which is essentially a codeword, s represents the symbol transmitted by the BS, and $\mathbf{n} \sim \mathcal{CN}\left(\mathbf{0}_N, \sigma^2 \mathbf{I}_N\right)$ represents the received noise with σ^2 representing the noise power.

Next, we will briefly review existing near-field XL-MIMO channel models for existing near-field beam training schemes.

B. Near-Field XL-MIMO Channel Model

When the distance between the BS and the UE is smaller than the Rayleigh distance [9], the near-field XL-MIMO channel should be modeled with the spherical wave assumption, which can be expressed by

$$\mathbf{h} = \sqrt{N}\alpha \mathbf{b}(\theta, r). \tag{2}$$

where α is the complex path gain. $\mathbf{b}(\theta, r)$ in (2) represents the near-field array response vector, which can be represented by [6]

$$\mathbf{b}(\theta, r) = \frac{1}{\sqrt{N}}[e^{-j\frac{2\pi}{\lambda}(r^{(1)}-r)}, \cdots, e^{-j\frac{2\pi}{\lambda}(r^{(N)}-r)}]^H, \tag{3}$$

where r represents the distance from the UE to the center of the antenna array, $r^{(n)} = \sqrt{r^2 + \delta_n^2 d^2 - 2r\delta_n d\theta}$ represents the distance from the UE to the nth BS antenna, and $\delta_n = \frac{2n-N-1}{2}$ with $n = 1, 2, \cdots, N$.

Before data transmission, the beam training should be applied to directly estimate the physical angles and distances of near-field channel paths. The beam training is to measure the power of y to find the best codewords from the codebook. The existing near-field beam training scheme is conduct exhaustive search in the polar-domain codebook [6], which can be represented as

$$\mathbf{A} = [\mathbf{b}(\theta_1, r_1^1), \cdots, \mathbf{b}(\theta_1, r_1^{S_1}), \cdots, \mathbf{b}(\theta_N, r_N^{S_N})], \tag{4}$$

The near-field response vector $\mathbf{b}(\theta, r)$ implies that the optimal beam training codeword should focus on the spatial angle θ and BS-UE distance r. Thus, in the XL-MIMO system, the size of codebooks should not only relate to the sample number of angle but also distance, which leads to a large-size codebook and unfordable beam training overhead. In next Section III, we design the multi-resolution near-field codebooks.

III. Near-Field Multi-Resolution Codewords Design and 2D Hierarchical Beam training

In this section, we will firstly formulate the problem of codeword design in the near-field scenario, and then propose a corresponding GS-based theoretical codeword design scheme and alternating optimization for practical codeword design. Then, we generate multi-resolution codebooks and propose a near-field 2D hierarchical beam training scheme.

A. Near-Field Multi-Resolution Codeword Design

1) Problem formulation: To evaluate the effectiveness of the codeword $\mathbf{v}$, we reference $G(\mathbf{v}, \theta, r)$ as the beamforming gain of $\mathbf{v}$ in the angle θ and the distance r. The $G(\mathbf{v}, \theta, r)$ can be represented as

$$G(\mathbf{v}, \theta, r) = \sqrt{N}\mathbf{b}(\theta, r)^H \mathbf{v}. \tag{5}$$

Suppose the angle coverage and distance coverage of codeword $\mathbf{v}$ are $\mathbf{B}_{\mathbf{v},\theta} \triangleq [\theta, \theta + B_\theta]$ and $\mathbf{B}_{\mathbf{v},r} \triangleq [r, r + B_r]$, where B_θ and B_r are the angle sampled step and distance sampled step. The ideal beam pattern is denote as

$$\mathbf{g}_{\mathbf{v}} = \left[g_{\mathbf{v}}(\theta_1, r_1^1), \cdots, g_{\mathbf{v}}(\theta_N, r_N^1), \cdots, g_{\mathbf{v}}(\theta_N, r_N^{S_N})\right], \tag{6}$$

where $g_{\mathbf{v}}(\theta, r) = |g_{\mathbf{v}}(\theta, r)| e^{jf(\theta, r)}$ is the beamforming gain. $|g_{\mathbf{v}}(\theta, r)|$ can be represented by

$$|g_{\mathbf{v}}(\theta, r)| = \begin{cases} \sqrt{C_{\mathbf{v}}}, & \theta \in \mathbf{B}_{\mathbf{v},\theta},\ r \in \mathbf{B}_{\mathbf{v},r} \\ 0, & \theta \notin \mathbf{B}_{\mathbf{v},\theta},\ r \notin \mathbf{B}_{\mathbf{v},r} \end{cases}. \tag{7}$$

For the ideal beam pattern in (6), the amplitude information

$|g_{\mathbf{v}}(\theta, r)|$ in (7) of beamforming gains in target angle coverage and distance coverage are flattened while other beamforming gains are zero. Meanwhile, the phase information $f(\theta, r)$ of the beamforming gains can be designed flexibly. Compared to a far-field codeword, the near-field codeword should not only cover a certain angle range but also a certain distance range.

The aim of designing a codeword is to make the beam pattern $\mathbf{A}^H\mathbf{v}$ obtained by beamforming with the codeword as close as possible to the ideal beam pattern $\mathbf{g}_{\mathbf{v}}$. Thus, the theoretical codeword $\mathbf{v}$ design problem can be expressed as

$$\min_{\mathbf{v}, f(\theta,r)} \left\|\mathbf{A}^H\mathbf{v} - \mathbf{g}_{\mathbf{v}}\right\|_2^2. \tag{P1}$$

In (P1), the ideal theoretical codeword $\mathbf{v}$ can only be realized by the fully digital XL-MIMO system, where each antenna requires one dedicated radio frequency (RF) chain. However, fully digital XL-MIMO system results in unaffordable energy consumption. In fact, hybrid digital-analog structure is usually preferred in XL-MIMO system to improve the energy efficiency. In this structure, we need to design practical codewords considering the hardware constraints in terms of phase shifter resolution and number of radio frequency (RF) chains N_{RF}. Based on the theoretical codeword $\mathbf{v}$, the design of the practical codeword $\mathbf{v}_{\mathrm{p}} \triangleq \mathbf{F}_{\mathrm{RF}}\mathbf{f}_{\mathrm{BB}}$ can be denotes as

$$\begin{aligned} \min_{\mathbf{F}_{\mathrm{RF}}, \mathbf{f}_{\mathrm{BB}}} \quad & \|\mathbf{v} - \mathbf{F}_{\mathrm{RF}}\mathbf{f}_{\mathrm{BB}}\|_2 \\ \text{s.t.} \quad & \|\mathbf{F}_{\mathrm{RF}}\mathbf{f}_{\mathrm{BB}}\|_2 = 1, \\ & [\mathbf{F}_{\mathrm{RF}}]_{n,i} = e^{j\delta}, \delta \in \Phi_b \\ & n = 1, 2, \ldots, N_{\mathrm{t}}, i = 1, 2, \ldots, N_{\mathrm{RF}}, \end{aligned} \tag{P2}$$

where the $\mathbf{F}_{\mathrm{RF}} \in \mathbb{C}^{N \times N_{\mathrm{RF}}}$ and $\mathbf{f}_{\mathrm{BB}} \in \mathbb{C}^{N_{\mathrm{RF}} \times 1}$ are the analog beamforming matrix and the digital beamforming vector. $\mathbf{\Phi}_b = \left[\pi\left(-1 + \frac{1}{2^b}\right), \pi\left(-1 + \frac{3}{2^b}\right), \ldots \pi\left(1 - \frac{1}{2^b}\right)\right]$ is the set of quantized phase shifters with b bits.

Next we will introduce the design method of the theoretical codeword $\mathbf{v}$ and practical codeword $\mathbf{v}_{\mathrm{p}}$.

2) Design of the theoretical codeword $\mathbf{v}$*:* In order to solve the (P1), we draw the experience from the **Gerchberg–Saxton (GS) algorithm**, which is widely applied in phase retrieval problem for digital hologram imaging [10], [11]. In the phase retrieval problem, the phase information needed to be obtained with the fixed amplitude information, which is same as the phase information $f(\theta, r)$ design of the ideal beam pattern in problem (P1). Specifically, the proposed GS-based near-field codeword design procedure is shown in **Algorithm 1**.

Algorithm 1: GS-based theoretical codeword design

Inputs: $|\mathbf{g}_{\mathbf{v}}|$, $C_{\mathbf{v}}$, I_{max}, $\mathbf{A}^H$, $B_{\mathbf{v},\theta}$, $B_{\mathbf{v},r}$.
Initialization: randomly generate $f(\theta, r)$ and obtain the ${\mathbf{g}_{\mathbf{v}}}^0$.
1. **for** $i = 1, 2, \cdots, I_{\max}$ **do**
2. calculate the $\hat{\mathbf{v}}^i$ based on ${\mathbf{g}_{\mathbf{v}}}^{i-1}$ by (8)
3. ${\mathbf{g}_{\mathbf{v}}}^i = |\mathbf{g}_{\mathbf{v}}|\angle(\mathbf{A}^H\hat{\mathbf{v}}^i)$
4. **end for**
5. $\mathbf{v} = \left(\mathbf{A}\mathbf{A}^H\right)^{-1}\mathbf{A}{\mathbf{g}_{\mathbf{v}}}^{I_{max}}$
Output: Theoretical codeword $\mathbf{v}$.

Base on the (P1), in i-th iteration, given $\mathbf{g}_{\mathbf{v}}^{i-1}$, the $\hat{\mathbf{v}}^i$ can be obtained by least square algorithm as

$$\hat{\mathbf{v}}^i = \left(\mathbf{A}\mathbf{A}^H\right)^{-1}\mathbf{A}\mathbf{g}_{\mathbf{v}i-1}. \tag{8}$$

Then, based on $\hat{\mathbf{v}}^i$, we can obtain the current beam pattern is $\mathbf{A}^H\hat{\mathbf{v}}^i$. In order to maintain the amplitude information of the $\mathbf{g}_{\mathbf{v}}$ to approach the ideal beam pattern, we only assign the phase information of current beam pattern $\mathbf{A}^H\hat{\mathbf{v}}^i$ to $\mathbf{g}_{\mathbf{v}}^i$. After the iteration number reaches I_{max}, we utilize $\mathbf{g}_{\mathbf{v}}^{I_{max}}$ to obtain the designed theoretical codeword $\mathbf{v}$.

3) Design of the practical codeword $\mathbf{v}_{\mathrm{p}}$*:* Based on the theoretical codeword $\mathbf{v}$ obtained by **Algorithm 1**, we solve the practical codeword $\mathbf{v}_{\mathrm{p}}$ design problem (P2) by iteratively optimize the digital beamforming vector $\mathbf{f}_{\mathrm{BB}}$ and the analog beamforming matrix $\mathbf{F}_{\mathrm{RF}}$ considering the hardware constraints. **Algorithm 2** provides the specific procedure to design the practical codeword.

Algorithm 2: Practical codeword design

Inputs: $\mathbf{v}$, $\mathbf{T}_{max}$, $\mathbf{P}_{max}$, $\mathbf{\Phi}_b$, N, N_{RF}.
Initialization: randomly generate $\mathbf{F}_{\mathrm{RF}}^0$.
1. **for** $t = 1, 2, \cdots, T_{\max}$ **do**
// Design the digital beamforming vector.
2. calculate the $\mathbf{f}_{\mathrm{BB}}^t$ by (9)
// Design the analog beamforming matrix.
3. **for** $p = 1, 2, \cdots, P_{max}$ **do**
4. **for** $n = 1, 2, \cdots, N$ **do**
5. **for** $i = 1, 2, \cdots, N_{\mathrm{RF}}$ **do**
6. Search $\delta_{n,i}$ to satisfy (10)
7. **end for**
8. **end for**
9. **if** $\delta_{n,i}^{p-1} = \delta_{n,i}^p$ **then**
10. Jump to Step2
11. **end if**
12. **end for**
13. obtain the $\mathbf{F}_{\mathrm{RF}}^t$ by utilizing (11)
14. **end for**
Output: $\mathbf{f}_{\mathrm{BB}} = \mathbf{f}_{\mathrm{BB}}^{T_{max}}$, $\mathbf{F}_{\mathrm{RF}} = \mathbf{F}_{\mathrm{RF}}^{T_{max}}$.

For the given analog beamforming matrix $\mathbf{F}_{\mathrm{RF}}$, the optimization problem of the digital beamforming vector $\mathbf{f}_{\mathrm{BB}}$ can be expressed as

$$\min_{\mathbf{f}_{\mathrm{BB}}} \|\mathbf{v} - \mathbf{F}_{\mathrm{RF}}\mathbf{f}_{\mathrm{BB}}\|_2, \tag{P2.1}$$

which can be solved least square as

$$\hat{\mathbf{f}}_{\mathrm{BB}} = \left(\mathbf{F}_{\mathrm{RF}}^H\mathbf{F}_{\mathrm{RF}}\right)^{-1}\mathbf{F}_{\mathrm{RF}}^H\mathbf{v} \tag{9}$$

Then, for the given analog beamforming vector $\mathbf{f}_{\mathrm{BB}}$, the optimization problem of $\mathbf{F}_{\mathrm{RF}}$ can be expressed as

$$\begin{aligned} \min_{\mathbf{F}_{\mathrm{RF}}} \quad & \|\mathbf{v} - \mathbf{F}_{\mathrm{RF}}\mathbf{f}_{\mathrm{BB}}\|_2 \\ \text{s.t.} \quad & \|\mathbf{F}_{\mathrm{RF}}\mathbf{f}_{\mathrm{BB}}\|_2 = 1, \\ & [\mathbf{F}_{\mathrm{RF}}]_{n,i} = e^{j\delta}, \delta \in \Phi_b, \\ & n = 1, 2, \ldots, N_{\mathrm{t}}, i = 1, 2, \ldots, N_{\mathrm{RF}}, \end{aligned} \tag{P2.2}$$

The optimization of $\mathbf{F}_{\mathrm{RF}}$ problem (P2.2) can be converted to the minimization absolute value of each entry of the vector

$\mathbf{v} - \mathbf{F}_{\mathrm{RF}}\mathbf{f}_{\mathrm{BB}}$. Hence, the problem (P2.2) can be transformed into N sub-problems, which can be optimized one by one. The n-th sub-problem is rewritten as

$$\min_{\theta_1,\theta_2,\ldots,\theta_{N_{\mathrm{RF}}}} \left| [\mathbf{v}]_n - \sum_{i=1}^{N_{\mathrm{RF}}} [\mathbf{f}_{\mathrm{BB}}]_i \, e^{j\delta_{n,i}} \right| \quad \text{s.t. } \delta_{n,i} \in \Phi_b, i = 1, 2, \ldots, N_{\mathrm{RF}}. \tag{10}$$

To obtain the solution to (10), the exhaustive search is a obvious choice, where all the combination of $\delta_{n,1}, \cdots, \delta_{n,N_{\mathrm{RF}}}$ are test to minimize the objective. However, the number of combination is $2^{bN_{\mathrm{RF}}}$, which has prohibitively high computational complexity. For example, if $b = 4, N_{\mathrm{RF}} = 32$, the $2^{bN_{\mathrm{RF}}} \approx 7.9 \times 10^{28}$! Thus, we need to investigate near-optimal search method to reduce complexity.

In this case, we propose a high efficient individual search method, where each $\delta_{n,i}$ is determined separately in each iteration. The specific procedures are summarized in **Algorithm 2**. We firstly initialize the $\delta^0_{n,1}, \cdots, \delta^0_{n,N_{\mathrm{RF}}}$ by choosing the entry from the $\mathbf{\Phi}_b$ and generate $\mathbf{F}^0_{\mathrm{RF}}$. In p-th iteration, we find best $\delta_{n,1}, \cdots, \delta_{n,N_{\mathrm{RF}}}$ one by one. In step 6, for $\delta_{n,i}$, we search through the $\mathbf{\Phi}_b$ to find the optimal choice to satisfy the (10). This iterative process performs stop until the number of iterations reaches predetermined figure or $\delta^{p-1}_{n,i} = \delta^p_{n,i}$. Then the n-th row of the designed $\hat{\mathbf{F}}_{\mathrm{RF}}$ can be expressed as

$$\left[\hat{\mathbf{F}}_{\mathrm{RF}}\right]_{n,:} = \left[e^{j\hat{\delta}_{n,1}}, e^{j\hat{\delta}_{n,2}}, \ldots, e^{j\hat{\delta}_{n,N_{\mathrm{RF}}}}\right] \tag{11}$$

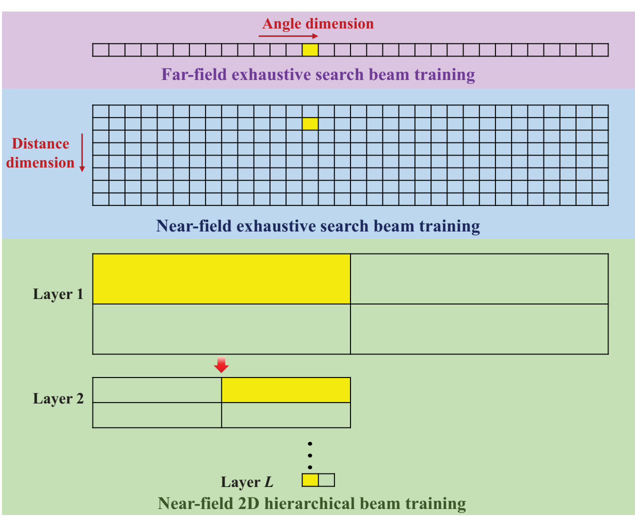


Fig. 1. Comparison between the far-field exhaustive search, near-field exhaustive search and the near-field 2D hierarchical beam training.

B. Near-field 2D Hierarchical Beam Training

In order to obtain the tradeoff between the near-field beam training overhead and the performance, one of the method is to apply a hierarchical near-field codebook, which consists of multi-resolution codebooks. The sizes of codebooks are determined by the angle sample step and distance sample step, i.e., B_θ and B_r in (7). Specifically, as the increase of B_θ and B_r, the corresponding codeword has a lower resolution, and the corresponding codebook size becomes smaller. As mentioned before, we can generate near-field multi-resolution codebooks with different angle coverages and distance coverages based on the **Algorithm 1** and **Algorithm 2**.

Then, these multi-resolution codebooks are applied to conduct near-field 2D hierarchical beam training. Compared with far-field scenario, the near-field 2D hierarchical beam training need to reduce the search range of angle and distance at the same time as shown in Fig. 1. The specific near-field beam training procedure is summarized in **Algorithm 3**.

Algorithm 3: Near-field 2D hierarchical beam training

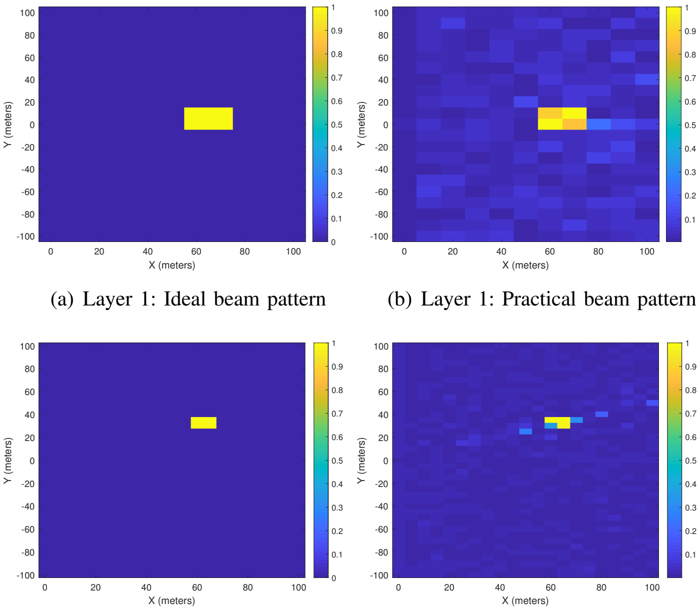
Inputs: L,$\{B^1_\theta, B^2_\theta, \cdots, B^L_\theta\}$, $\{B^1_r, B^2_r, \cdots, B^L_r\}$, $y_{opt} = 0$, $s_{opt} = 0$
// Generate L sub-codebooks
1. **for** $l = 1, 2, \cdots, L$ **do**
2. generate the collection of $\mathbf{B}^l_{\mathbf{v}_{l,k},\theta}$ and $\mathbf{B}^l_{\mathbf{v}_{l,k},r}$ based on B^l_θ and B^l_r
3. generate $|g_{\mathbf{v}}(\theta, r)|$ for based on (7)
4. obtain the practical codewords in l-th sub-codebook $\mathbf{W}_l$ based on **Algorithm 1** and **Algorithm 2**.
5. **end for**
6. $\mathbf{W} = \mathbf{W}_1$
// Conduct beam training
7. **for** $l = 1, 2, \cdots, L$ **do**
8. **for** $\mathbf{v}_{l,k}$ in $\mathbf{W}$ **do**
9. $y^l_k = \mathbf{h}^H \mathbf{v}_{l,k} s + n$
10. **if** $y^l_k > y_{opt}$ **then**
11. $k_{opt} = k$
12. **end if**
13. **end for**
14. choose $\mathbf{v}_{l+1,k}$ in $\mathbf{W}_{l+1}$ satisfied $\mathbf{B}^{l+1}_{\mathbf{v}_{l+1,k},\theta} \in \mathbf{B}^l_{\mathbf{v}_{l,k_{opt}},\theta}$ and $\mathbf{B}^{l+1}_{\mathbf{v}_{l+1,k},r} \in \mathbf{B}^l_{\mathbf{v}_{l,k_{opt}},r}$
15. the chosen codewords $\mathbf{v}_{l+1,k}$ compose the $\mathbf{W}$
16. **end for**

Output: The feedback optimal codeword index k_{opt} from the user.

First, as shown in Step2, for l-th codebook generation, we need to divide the angle coverage $\mathbf{B}^l_{\mathbf{v}_k,\theta}$ and distance coverage $\mathbf{B}^l_{\mathbf{v}_k,r}$ based on angle samples step B^l_θ and distance samples step B^l_r for each codeword. Then, in Steps 3-4, the codewords design scheme based on **Algorithm 1** and **Algorithm 2** is applied to obtain the l-th codebook $\mathbf{W}_l$. Then, Steps 7-16 are operated to search the optimal codeword in each codebook layer by layer.

C. Comparison of the beam training overhead

Beam training overhead refers to the number of time slots used for beam training. Generally, the beam training overhead is determined by the spatial resolutions of an antenna array on the angle and distance, i.e., the number of sampled angles U and the number of sampled distances S. It is worth pointing



(a) Layer 1: Ideal beam pattern (b) Layer 1: Practical beam pattern

(c) Layer 2: Ideal beam pattern (d) Layer 2: Practical beam pattern

Fig. 2. Comparison of the beam patterns of different layers of the hierarchical codebook.

out that U is usually set as the same as the number of antennas on the array. The training overhead of the exhaustive near-field beam training scheme is US. Meanwhile, the training overhead of the time-delay based beam training is only related to the number of sampled distances S. For the proposed 2D hierarchical beam training method, the beam training overhead can be represented as $\mathcal{O}\left(log\left(U\right)+log\left(S\right)\right)$. It is obvious that, the training overhead of the proposed 2D hierarchical beam training is much less than that of the exhaustive near-field beam training. Since the number of sampled angles U is usually large than the number of sampled distances S [8], the training overhead of the proposed 2D hierarchical beam training is larger than that of the time-delay based beam training. However, the performance of the time-delay based beam training heavily depends on the extra hardware overhead and wideband condition, which will be further verified by simulation results in Section IV.

IV. Simulation Results

For simulations, we assume that the number of BS antennas and RF chains are $N=512$ and $N_{\mathrm{RF}}=100$. The wavelength is set as $\lambda=0.005$ meters, corresponding to the 60 GHz frequency. The quantified bits number of phase shifters is set as $b=5$. The path gain α, angle θ and distance r are generated as following: $\alpha\sim\mathcal{CN}\left(0,1\right)$, $\theta\sim\mathcal{U}\left(-1,1\right)$, and $r\sim\mathcal{U}\left(20,100\right)$ meters. The SNR is defined as $1/\sigma^2$.

Fig. 2 shows the comparison of the ideal beam pattern and the normalized practical beam pattern obtained by the designed practical codeword. In these heat maps, the brighter the color, the greater the beamforming gain at this position.

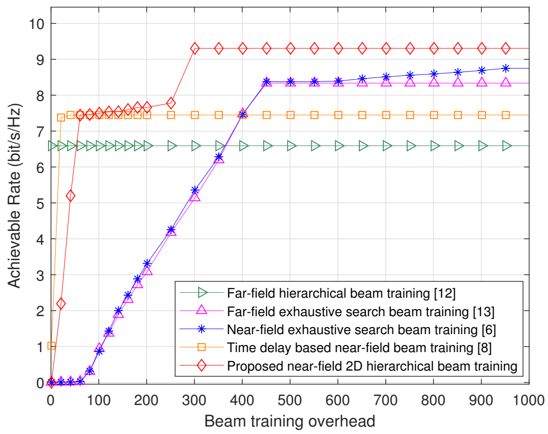
TABLE I
Comparison of Beam Training Overhead

Method	Overhead	Value
Far-field hierarchical scheme [12]	$\sum_l^L U^{(l)}$	40
Far-field exhaustive search scheme [13]	U	512
Near-field exhaustive search scheme [6]	US	8192
Time-delay based near-field scheme [8]	S	16
Proposed near-field 2D hierarchical scheme	$\sum_l^L U^{(l)}S^{(l)}$	268

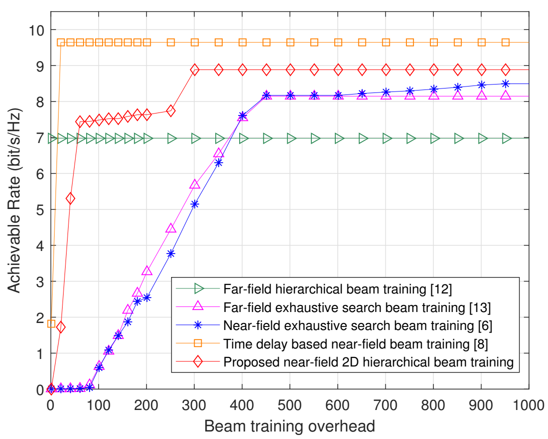
It is worth noting that, in order to show the beam pattern more clearly, we utilize the rectangular coordinate system to present the beamforming gains of the locations in two-dimension space, where the coordinates of X-axis and Y-axis satisfy $x=r\cos(\theta)$, and $y=r\cos(\theta)$. Fig. 2 (a) presents an ideal beam pattern of the sub-codebook of layer 1, where the beam should focus on the target location, i.e., $x=r\cos(\theta)$, and $y=r\cos(\theta)$. After we conduct beamforming with the designed practical codeword, we can obtain Fig. 2 (b), which presents the beamforming gains of different locations in space with the designed practical codeword. From Fig. 2 (b) we can see that the target location has the largest beamforming gain and other locations have much lower beamforming gains. Moreover, from Fig. 2 (c) and (d), for the codeword in the codebook of layer 2, the designed practical codeword can also realize the beam pattern close to the ideal beam pattern.

Table. I presents the comparison of beam training overhead for different methods. We compare the proposed near-field 2D hierarchical beam training algorithm with the existing far-field hierarchical beam training scheme [12], far-field exhaustive search beam training scheme [13], the near-field exhaustive search beam training scheme [6], and time-delay based near-field beam training scheme [8]. We set the number of angle and distance samples as $U=512$ and $S=16$, respectively. The overhead of the far-field exhaustive search is set as the same as the number of sampled angles, i.e., 512. For the near-field exhaustive search scheme, which relates to not only the number of sampled angles but also the number of sampled distances. Thus, the overhead of the near-field exhaustive search beam training scheme is set as $512\times16=8192$. The overhead of time-delay based near-field beam training relates to the number of sampled distances, which is set as 16. For the far-field hierarchical beam training scheme, $U^{(l)}$ is the number of sampled angles in the l-th layer, where $U^{(1)}=4, U^{(2)}=4, U^{(2)}=32$. Thus, the overhead of far-field hierarchical beam training is $\sum_l^L U^{(l)}=4+4+32=40$. For the proposed near-field 2D hierarchical beam training algorithm, we use three-layer codebooks. The size of codebook layer 1 can be calculated as $64\times4=256$, where the number of sampled angles and distances is set as 64 and 4. For the codebook layer 2 and layer 3, we only need to search 8 and 4 codewords. Thus the overhead of the proposed near-field 2D hierarchical beam training algorithm is 268, which is almost half of 512 and only 3.3 % of 8192.

Fig. 3 presents the performance of achievable rate comparison against the beam training overhead under different



(a)



(b)

Fig. 3. Achievable sum-rate performance comparison with respect to the beam training overhead under different bandwidths. (a) 100 MHz; (b) 500 MHz.

bandwidths. From Fig. 3 (a), where the bandwidth is 100 MHz, we can observe that the proposed near-field 2D hierarchical beam training can achieve the best performance of all schemes with relatively lower overhead. For example, the proposed scheme outperforms the far-field angle-domain codebook with only half of the beam training overhead. The reason is that the existing far-field codebook can only capture the angle information of the channel path. Moreover, the time-delay based scheme has worse performance than the proposed scheme in this narrow-band condition. The principal reason is that the ability of time-delay circuits to control the beam split will decrease by reducing the bandwidth. Meanwhile, Fig. 3 (b) illustrates the wide-band situation, where the bandwidth is 500 MHz. It can be observed that the time-delay based beam training scheme has better performance than the proposed scheme. However, the proposed scheme has much lower hardware cost and is bandwidth-independent. Thus, we believe that the proposed scheme provides a tradeoff between the performance and overhead in near-field XL-MIMO beam training in a more general and cost-saving way.

V. Conclusions

In this paper, we proposed a low-overhead near-field 2D hierarchical beam training by designing the near-field multi-resolution codebooks. Specifically, we first formulate the problem of designing near-field codeword and generating multi-resolution codebooks. It is worth pointing out that the proposed Gerchberg–Saxton (GS) based algorithm can be utilized in designing codewords to realize arbitrary beam patterns. Then, a low-overhead near-field 2D hierarchical beam training scheme is proposed. Significantly, the proposed scheme can achieve sub-optimal performance without restriction to the hardware cost and wideband condition.

VI. Acknowledgment

This work was supported in part by the National Key Research and Development Program of China (Grant No. 2020YFB1807201), in part by the National Natural Science Foundation of China (Grant No. 62031019), and in part by the European Commission through the H2020-MSCA-ITN META WIRELESS Research Project under Grant 956256.

References

[1] M. Giordani, M. Polese, M. Mezzavilla, S. Rangan, and M. Zorzi, "Toward 6G networks: Use cases and technologies," *IEEE Commun.Mag.*, vol. 58, no. 3, pp. 55–61, Mar. 2020.

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

[3] A. Alkhateeb, O. El Ayach, G. Leus, and R. W. Heath, "Channel estimation and hybrid precoding for millimeter wave cellular systems," *IEEE J. Sel. Top. Signal Process.*, vol. 8, no. 5, pp. 831–846, Oct. 2014.

[4] S. Noh, M. D. Zoltowski, and D. J. Love, "Multi-resolution codebook based beamforming sequence design in millimeter-wave systems," in *IEEE Global Communications Conference (GLOBECOM'15)*, 2015, pp. 1–6.

[5] A. Pizzo, L. Sanguinetti, and T. L. Marzetta, "Fourier plane-wave series expansion for holographic mimo communications," *IEEE Trans. Wireless Commun.*, vol. 21, no. 9, pp. 237–246, Sep. 2022.

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

[7] X. Wei, L. Dai, Y. Zhao, G. Yu, and X. Duan, "Channel estimation for RIS assisted wireless communications: Part II - an improved method based on double-structured sparsity," *China Commun.*, vol. 19, no. 6, pp. 193–204, Jun. 2022.

[8] M. Cui, L. Dai, Z. Wang, S. Zhou, and N. Ge, "Near-field rainbow: Wideband beam training for XL-MIMO," *IEEE Trans. Wireless Commun.*, (early access) 2022.

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

[10] O. Bucci, G. Franceschetti, G. Mazzarella, and G. Panariello, "Intersection approach to array pattern synthesis," *IEEE Photonics Journal*, vol. 137, no. 6, pp. 349–357, Dec. 1990.

[11] W. Chen, "3D gerchberg-saxton optical correlation," *IEEE Photonics Journal*, vol. 10, no. 2, pp. 1–9, Apr. 2018.

[12] S. Noh, M. D. Zoltowski, and D. J. Love, "Multi-resolution codebook and adaptive beamforming sequence design for millimeter wave beam alignment," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 5689–5701, Sep. 2017.

[13] J. Lee, G. Gil, and Y. H. Lee, "Channel estimation via orthogonal matching pursuit for hybrid MIMO systems in millimeter wave communications," *IEEE Trans. Wireless Commun.*, vol. 64, no. 6, pp. 2370–2386, Jun. 2016.