Near-Field Wideband Beam Training Based on Distance-Dependent Beam Split

Tianyue Zheng, Mingyao Cui, Zidong Wu, and Linglong Dai, *Fellow, IEEE*

***Abstract*—Near-field beam training is essential for acquiring channel state information in 6G extremely large-scale multiple input multiple output (XL-MIMO) systems. To achieve low-overhead beam training, existing method has been proposed to leverage the near-field beam split effect, which deploys true-time-delay arrays to simultaneously search multiple angles of the entire angular range in a distance ring with a single pilot. However, the method still requires exhaustive search in the distance domain, which limits its efficiency. To address the problem, we propose a distance-dependent beam-split-based beam training method to further reduce the training overheads. Specifically, we first reveal the new phenomenon of distance-dependent beam split, where by manipulating the configurations of time-delay and phase-shift, beams at different frequencies can simultaneously scan the angular domain in multiple distance rings. Leveraging the phenomenon, we propose a near-field beam training method where both different angles and distances can simultaneously be searched in one time slot. Thus, a few pilots are capable of covering the whole angle-distance space for wideband XL-MIMO. Theoretical analysis and numerical simulations are also displayed to verify the superiority of the proposed method on beamforming gain and training overhead.**

***Index Terms*—Beam training, extremely large-scale MIMO (XL-MIMO), near-field, wideband.**

I. Introduction

With the dramatically increasing demand for data transmission in future sixth-generation (6G) communications, a 10-fold increase in spectrum efficiency compared with 5G is anticipated in 6G [1], [2]. To fulfill this vision, 5G massive multiple input multiple output (MIMO) systems will evolve into 6G extremely large-scale MIMO (XL-MIMO) systems [3]–[5]. Thanks to the dramatically increased antenna aperture, the spatial multiplexing and beamforming gains could be significantly enhanced in XL-MIMO [6]. To harvest these gains, acquiring accurate channel state information efficiently by beam training is necessary. Specifically, beam training is realized by selecting an optimal beam codeword from a predefined beam codebook so as to assign directional beams to target users, which serves as a dominate CSI acquition scheme in millimeter wave (mmWave) and terahertz (THz) XL-MIMO systems [7], [8].

A. Prior Works

Compared with 5G massive MIMO, beam training for 6G XL-MIMO will induce overwhelming training overheads due to fundamental change in electromagnetic (EM) field property. Specifically, in 5G systems when the antenna number at the base station (BS) is not very large, channel model mainly considers far-field propagation where radiated EM waves can be approximated as planar waves. Thus, the array response vector of the far-field channel is solely related to the angle. Correspondingly, the orthogonal Discrete Fourier Transform (DFT) codebook can conduct an efficient scan in the angular domain [9]–[11], where each codeword in the codebook corresponds to a narrow beam directing to a specific angle.

As for the 6G XL-MIMO systems, however, the increased antenna aperture leads to more users located in spherical-wave-based near-field regions [12], [13]. Unlike the angle-dependent far-field channel, the angle and distance from BS to user jointly determine the array response vectors of near-field channels. Thereby, to perform near-field beam training, a polar-domain codebook was proposed in [12] to simultaneously obtain accurate angle and distance information. This codebook partitions the entire "angle-distance" space into multiple grids, where each codeword is construced by the array response vector aligned with one sampled grid. During beam training, all codewords are sequentially tested, where the angle and distance information are estimated based on the codeword yielding the highest received power. Unfortunately, although this beam training method can acquire the near-field CSI accurately, it consumes unacceptable pilot overheads due to the two-dimensional exhaustive search for all possible BS-to-user angles and distances.

To cope with this problem, two categories of near-field beam training have been proposed, namely narrowband and wideband methods. Consider the narrowband methods at first. They are realized either by hierarchical beam training [14]–[17] or inference from far-field beam training [18]–[23]. The near-field hierarchical beam training methods employ multi-resolution codebooks with different angle and distance coverages. Specifically, the spatial region covered by a lower-resolution codeword at any given layer is subdivided into several higher-resolution and non-overlapping spatial regions in the subsequent layer. During beam training, one can start from the lowest-resolution codebook and progress to the highest-resolution one layer by layer, and thereby gradually reduces the angle and distance estimation granularity by choosing the spatial region with larger received power in each layer [14]–[17]. As for the inference from far-field beam training, these methods attempt to recover near-field CSI with the aid of far-field codebook via a two-stage training strategy. In the first stage, the methods estimate the candidate angles based on the pattern of received power provided by far-field exhaustive

This work was funded in part by the National Science Fund for Distinguished Young Scholars (Grant No. 62325106), and in part by the National Key R&D Program of China (No. 2023YFB811503).

T. Zheng, Z. Wu, and L. Dai are with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China, and also with the Beijing National Research Center for Information Science and Technology (BNRist), Beijing 100084, China (e-mails: {zhengty22, wuzd19}@mails.tsinghua.edu.cn, {daill}@tsinghua.edu.cn). M. Cui is with the Department of Electrical and Electronic Engineering, the University of Hong Kong, Hong Kong (e-mail: cuimy23@connect.hku.hk).

beam training. Then the second stage uses a polar-domain codebook to search for distance within candidate angles or directly estimates the distance from far-field received beam pattern with deep learning-based methods [18]–[23]. Moreover, there have been works [24], [25] devoted to improving the accuracy of beam training by beam refinement methods, since the performance of above methods is limited by the resolution of the predefined codebook.

The above methods can to some extent reduce the training overheads. Nevertheless, the narrowband methods only measures the received power of one subcarrier in a single time slot, limiting the information content obtained per pilot transition. This constrains the efficiency of narrowband near-field beam training. In contrast, the majority of practical communication systems are wideband systems and it motivates us to develop wideband near-field beam training methods, which utilize information in multiple subcarriers [26] to further reduce the beam training overheads. Unfortunately, to our best knowledge, few works have focused on wideband near-field beam training.

Excavating frequency resources, in our prior work [27], a wideband near-field beam training scheme, namely near-field rainbow, was designed based on near-field beam split. This scheme was built on the true-time-delay (TTD) array. By manipulating the time-delay (TD) parameters, multiple beams at different frequencies are focused on multiple angles in a specific distance ring. In other words, multiple angles can be searched simultaneously in a single time slot, which significantly improves the training efficiency. However, since the beam split in [27] is essentially distance-independent beam split, the optimal distance is obtained by sequentially traversing different distance rings. Thereby, the method in [27] could not get rid of the additional overheads in the distance domain and results in relatively high training overheads.

B. Our Contributions

To tackle this problem, by exploiting *distance-dependent* feature of near-field beam-split, we propose a near-field beam training method where both different angles and distances can be searched simultaneously in one time slot, achieving efficient beam training.[1] Our contributions are summarized as follows.

- Firstly, we reveal the mechanism of near-field distance-dependent beam split. Specifically, [27] has proved that with the elaborate configuration of TD parameters, the beams at different frequencies could be made distributed on multiple angles of *one-distance-ring*. In this paper, we prove that by introducing phase-shift (PS), the collaboration of TD and PS can spread the focused points of multi-frequency beams on different angles of *multi-distance-rings*. In this way, the traditional distance-independent beam split can be transformed into the distance-dependent beam split.
- Secondly, leveraging the mechanism of distance-dependent beam split, an on-grid wideband beam training method is proposed to realize efficient near-field beam training. In the proposed scheme, both the angle and distance domain can be searched with beams of different subcarriers in a single pilot, where the beam with the highest power can be efficiently determined. Moreover, quantitative analysis of the required pilot number to guarantee the full coverage of the whole angle-distance space is presented.
- To further improve the beam training performance for off-grid users, we propose two off-grid beam training algorithms: the auxiliary beam pair-assisted method and match filter-based method. Their principle is to employ the envelope of received powers over multiple subcarriers, rather than relying on the single subcarrier with the highest power, to jointly select the optimal codeword.
- Finally, we present simulation results to validate the efficiency of the proposed method based on distance-dependent beam split. It shows that the proposed beam training scheme is capable of achieving near-optimal rate performance with a lower overhead compared to existing methods. Moreover, the enhanced algorithms tailored for off-grid users surpass the performance of the on-grid algorithm, particularly in regions of low signal-to-noise ratio (SNR).

[1]Simulation codes will be provided to reproduce the results in this paper: http://oa.ee.tsinghua.edu.cn/dailinglong/publications/publications.html.

C. Organization and Notation

1) Organization: The rest of the paper is organized as follows. Section II introduces the near-field XL-MIMO wideband channel model, and formulate the near-field beam training problem. Then the mechanism of existing near-field distance-independent beam split will be reviewed. In Section III, the mechanism of distance-dependent beam split is introduced, and based on it we provide the design of an efficient wideband near-field beam training method. In Section IV, two improved beam training methods for off-grid users are described. Simulation results and conclusions are provided in Section V and Section VI, respectively.

2) Notation: $\mathbf{a}^H$, $\mathbf{A}^H$ denote the conjugate transpose of vector $\mathbf{a}$ and matrix $\mathbf{A}$, respectively; $\|\mathbf{a}\|$ denotes the l_2 norm of vector $\mathbf{a}$; $\odot$ denotes Hadamard product; $\mathcal{CN}(\mu, \mathbf{\Sigma})$ denotes the probability density function of complex multivariate Gaussian distribution with mean μ and variance $\mathbf{\Sigma}$; $\mathrm{round}(a)$, $\lfloor a \rfloor$ and $\lceil a \rceil$ denote rounding to, up to, and down to the nearest integer, respectively.

II. System Model

In this section, we first introduce the near-field wideband channel model in XL-MIMO systems and formulate the near-field beam training problem. Then, we will briefly review the traditional near-field beam split proposed in [27].

A. Near-Field Wideband Channel Model

We examine a millimeter-wave wideband near-field XL-MIMO system. The BS equipped with an N_t-antenna uniform linear array serves a single-antenna user using orthogonal frequency division multiplexing (OFDM) of M subcarriers. The bandwidth, central carrier frequency, speed of light, central

wavelength, and antenna spacing are denoted as B, f_c, c, $\lambda_c = c/f_c$ and $d = \lambda_c/2$, respectively. The lower and upper subcarrier frequencies can be written as $f_L = f_c - B/2$ and $f_H = f_c + B/2$ and the frequency of the m-th subcarrier as $f_m = f_c + \frac{B}{M}(m - 1 - \frac{M-1}{2})$. Besides, the number of BS antennas is assumed to be odd and is denoted by $N_t = 2N+1$ for expression simplicity.

In this paper, we mainly focus on the downlink beam training process. Let $s_m = 1$ be the normalized transmitted symbol at f_m, then the received signal y_m at the user equipment (UE) of the m-th subcarrier is given by

$$y_m = \sqrt{P_t}\mathbf{h}_m^H \boldsymbol{w}_m s_m + n_m, \tag{1}$$

where $P_t > 0$ is the transmit power, $\mathbf{h}_m \in \mathbb{C}^{N_t} \times 1$ the downlink channel, $\boldsymbol{w}_m \in \mathbb{C}^{N_t \times 1}$ the unit-norm beamformer, and $n_m \sim \mathcal{CN}(0, \sigma^2)$ the complex circularly-symmetric Gaussian noise with a variance of σ^2. We adopt the Saleh-Valenzuela channel model in [28] to represent the downlink channel $\mathbf{h}_m \in \mathbb{C}^{N_t \times 1}$ at the m-th subcarrier as

$$\mathbf{h}_m = \sqrt{\frac{N_t}{L}} \sum_{l=1}^{L} \beta_{m,l} e^{-j2\pi f_m \tau_l} \mathbf{a}_m(\theta_l, r_l), \tag{2}$$

where L is the number of paths being either line-of-sight (LoS) or non-line-of-sight (NLoS) path, and $\beta_{m,l}$, τ_l, $\theta_l = \sin\vartheta_l$, r_l denote the path gain, time delay, spatial direction, and distance from the UE/scatter to the center of the antenna array of the l-th path, respectively. Due to the significant path loss experienced by scattered signals, millimeter-wave communications heavily rely on the LoS path [29], which is our primary focus:

$$\mathbf{h}_m = \sqrt{N_t}\beta_m e^{-jk_m r} \mathbf{a}_m(\theta, r), \tag{3}$$

where $k_m = 2\pi f_m/c$ denotes the wavenumber at m-th subcarrier and the path gain of the LoS path is modeled as [30]

$$\beta_m = \frac{\lambda_m}{4\pi r}. \tag{4}$$

Therefore, we can obtain the relationship between the path gain at m-th subcarrier and central frequency f_c as

$$\beta_m = \frac{f_c}{f_m}\beta_c. \tag{5}$$

The near-field array response vector $\mathbf{a}_m(\theta, r)$ in (3) is derived based on spherical wave propagation:

$$\mathbf{a}_m(\theta, r) = \frac{1}{\sqrt{N_t}}[e^{-jk_m(r^{(-N)} - r)}, \ldots, e^{-jk_m(r^{(N)} - r)}]^T \tag{6}$$

where $r^{(n)} = \sqrt{r^2 + n^2 d^2 - 2r\theta nd}$ represents the distance from the n-th BS antenna to the UE. By utilizing the Taylor expansion $\sqrt{1+x} \approx 1 + \frac{1}{2}x - \frac{1}{8}x^2$ [12], the distance difference $r^{(n)} - r$ can be approximated by:

$$r^{(n)} - r \approx -nd\theta + n^2 d^2 \frac{1 - \theta^2}{2r}. \tag{7}$$

For expression simplicity, we denote $\alpha = \frac{1-\theta^2}{2r}$. For fixed α, the curve plotted by $(r = \frac{1-\theta^2}{2\alpha}, \theta)$ can be viewed as a distance ring in the physical space. In this case, $\mathbf{a}_m(\theta, r)$ can be approximated with $\mathbf{b}_m(\theta, \alpha)$ written as

$$\mathbf{a}_m(\theta, r) \approx \mathbf{b}_m(\theta, \alpha) = \frac{1}{\sqrt{N_t}}[e^{jk_m((-N)d\theta - (-N)^2 d^2 \alpha)}, \ldots, e^{jk_m(Nd\theta - N^2 d^2 \alpha)}]. \tag{8}$$

Vector $\mathbf{b}_m(\theta, \alpha)$ effectively captures the spherical wave propagation in wideband scenarios.

B. Formulation of Near-Field Beam Training Problem

The near-field beam training attempts to estimate the location information (θ, α) of the user, thereby aligning the beamforming vector with this location. To adapt to the near-field communication scenario, the optimal beamforming vector is typically selected from the predefined polar-domain codebook [12], which can be represented as

$$\mathcal{A}_m = [\mathbf{b}_m(\theta_1, \alpha_1^1), \cdots, \mathbf{b}_m(\theta_1, \alpha_1^{S_1}), \cdots, \mathbf{b}_m(\theta_{N_t}, \alpha_{N_t}^{S_{N_t}})], \tag{9}$$

where S_n is the number of sampled distance grids at θ_n. Each column of $\mathcal{A}_m$ is a codeword focused on the grid $(\theta_n, \alpha_n^{s_n})$, with $s_n = 1, 2, \cdots, S_n$. When conducting exhaustive beam training, we sequentially employ the codewords to perform beamforming and select the codeword leading to the largest received power to estimate the physical location $(\hat{\theta}, \hat{\alpha})$, i.e.

$$(\hat{\theta}, \hat{\alpha}) = \arg \max_{(\theta_n, \alpha_n^{s_n})} \sum_{m=1}^{M} \|\sqrt{P_t}\mathbf{h}_m^H \mathbf{b}_m(\theta_n, \alpha_n^{s_n}) + n_m\|^2. \tag{10}$$

Finally, we set the beamforming vector as $\boldsymbol{w}_m(\hat{\theta}, \hat{\alpha}) = \mathbf{b}_m(\theta_n, \alpha_n^{s_n})$ to serve the user. Consequently, the size of the polar-domain codebook is determined by the multiplication of the number of sampled angular grids and sampled distance grids. This results in a large codebook size, which in turn leads to prohibitive beam training overheads.

C. Near-Field Distance-Independent Beam Split

In this subsection, we review the mechanism of controllable near-field beam split proposed in [27] and explain how to employ it to achieve efficient beam training. Here we only provide a brief summary of controllable near-field beam split, where we attempt to control the beam split effect (i.e. the focused locations of beams at different frequencies) by elaborately designing the beamformer parameters. For more detailed illustration of conventional beam split, and its effects on beamforming, please refer to [26], [31], [32].

We utilize a TD-PS precoding structure, as presented in Fig. 1, to manipulate the near-field beam split effect. Note that the TD array adopted in [27] is a special case of TD-PS array by forcing the phase shift to zero. Therefore, the TD-PS array can achieve much more flexible control of the beam split effect, particularly in the distance dimension.

As illustrated in Fig. 1, each antenna is sequentially connected to a time-delay unit and a phase-shift unit. A TD unit is capable of tuning *frequency-dependent* phase shift by actively delaying the signal. Accordingly, we can express the n-th element of the TD beamforming vector $\mathbf{w}_m^{\mathrm{TD}}$ at f_m, as

$$[\mathbf{w}_m^{\mathrm{TD}}]_n = \frac{1}{\sqrt{N_t}} e^{-j2\pi f_m \tau^{(n)'}} = \frac{1}{\sqrt{N_t}} e^{-jk_m r_t^{(n)'}} \tag{11}$$

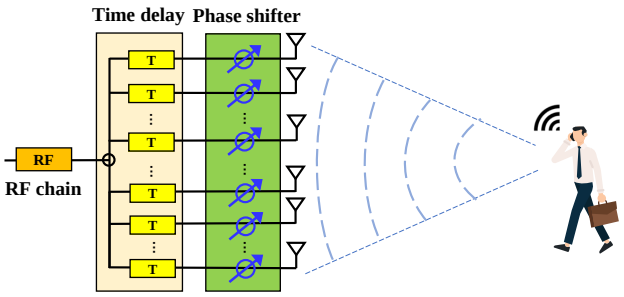


Fig. 1. TD-PS precoding structure.

where $\tau^{(n)'}$ is the adjustable delay of the n-th TD unit, while $r_t^{(n)'} \triangleq c\tau^{(n)'}$. We use the superscript $'$ to indicate the adjustable parameters. Since $\mathbf{w}_m^{\mathrm{TD}}$ has a similar form to the array response vector $\mathbf{b}_m(\theta,\alpha)$, we can set $r_t^{(n)'}$ as $r_t^{(n)'} = nd\theta_t' - n^2d^2\alpha_t'$, where θ_t' and α_t' are defined as the *adjustable TD parameters*. Thereby, the beamforming vector at f_m can be rewritten as

$$[\mathbf{w}_m^{\mathrm{TD}}(\theta_t',\alpha_t')]_n = \frac{1}{\sqrt{N_t}} e^{-jk_m(nd\theta_t' - n^2d^2\alpha_t')}. \tag{12}$$

Furthermore, the n-th element of the beamforming vector, $\mathbf{w}^{\mathrm{PS}}(\theta_p',\alpha_p')$, generated by *frequency-independent* PSs can be similarly set as

$$[\mathbf{w}^{\mathrm{PS}}(\theta_p',\alpha_p')]_n = \frac{1}{\sqrt{N_t}} e^{-jk_c(nd\theta_p' - n^2d^2\alpha_p')}, \tag{13}$$

where θ_p', α_p' are correspondingly *adjustable PS parameters*. By configuring TD-PS beamforming structure with parameters $\theta_t', \alpha_t', \theta_p', \alpha_p'$, the array gain at the m-th subcarrier on an arbitrary location (θ,α) can be presented as

$$\begin{aligned} & |(\mathbf{w}_m^{\mathrm{TD}}(\theta_t',\alpha_t') \odot \mathbf{w}^{\mathrm{PS}}(\theta_p',\alpha_p'))^T \mathbf{b}_m(\theta,\alpha)| \\ = & \frac{1}{N_t} \left| \sum_{n=-N}^{N} e^{jnd(k_m(\theta-\theta_t') - k_c\theta_p') - jn^2d^2(k_m(\alpha-\alpha_t') - k_c\alpha_p')} \right| \\ = & G(k_m\theta - k_m\theta_t' - k_c\theta_p', k_m\alpha - k_m\alpha_t' - k_c\alpha_p'), \end{aligned} \tag{14}$$

where we define $G(x,y) = \frac{1}{N_t}\left|\sum_{n=-N}^{N} e^{jndx - jn^2d^2y}\right|$ and it achieves a maximal value at $(x,y) = (0,0)$. According to [27], the beam at frequency f_m is focused on the location (θ_m,α_m) with the maximum array gain. Thereby we obtain $k_m\theta_m - k_m\theta_t' - k_c\theta_p' = 0$ and $k_m\alpha_m - k_m\alpha_t' - k_c\alpha_p' = 0$, or equivalently:

$$\theta_m = \theta_t' + \frac{f_c}{f_m}\theta_p', \tag{15}$$

$$\alpha_m = \alpha_t' + \frac{f_c}{f_m}\alpha_p'. \tag{16}$$

By setting $\theta_p' \neq 0$, the spatial angle θ_m is related to the frequency f_m, making beams at different frequencies split towards different spatial angles. Moreover, when $\alpha_p' = 0$, α_m becomes independent of f_m, indicating that all beams are focused on the same distance ring $\alpha_m = \alpha_t'$. To perform fast beam training, the authors in [27] propose to distribute multi-frequency beams over multiple angles in the same distance ring simultaneously by setting $\theta_p' \neq 0$ and $\alpha_p' = 0$. Then, it exhaustively searches different distance rings by configuring the value of α_t'.

However, it is clear that the near-field beam split in [27] is essentially distance-independent, where beams can cover multiple angles in only *one* specific distance ring. It fails to fully exploit the degree-of-freedoms provided by the beam split effect, which still requires exhaustive search in the distance domain and limits its efficiency. To deal with this problem, we will introduce the distance-dependent beam split phenomena and how we can search multiple angles as well as distances at the same time with a single pilot by utilizing this phenomenon.

III. Distance-Dependent Beam Split and Proposed Near-Field Beam Training Scheme

In this section, the mechanism of distance-dependent beam split is elaborated. Thanks to the periodicity of the angle dimension, beams can skip periodically between $\theta = 1$ and $\theta = -1$. Therefore, by elaborately designing the TD and PS parameters, we can divide the subcarriers over the entire bandwidth into several groups, and the beams of each group cover the angular range with different distance ranges. Taking advantage of this mechanism, we propose a distance-dependent beam-split-aided near-field beam training method. In the proposed method, multiple distance rings can be covered by the beams at different frequency groups simultaneously in the near field. Based on the discovery, both angle and distance dimensions can be searched simultaneously with different subcarriers in a single pilot during beam training.

A. Distance-Dependent Beam Split

For a better understanding of the mechanism of distance-dependent beam split, in this subsection, we start from the periodical beam split in the far field. Then we will illustrate the distance-dependent beam split in near-field scenarios.

1) Mechanism of Far-field Periodical Beam Split: To illustrate the far-field periodical beam split phenomena, we first analyze the periodicity of the array gain with the TD-PS precoding structure. Specifically, the distance r is assumed to be larger than the Rayleigh distance so that distance-related parameters α reliably approach 0. Then, the array response vector becomes $[\mathbf{b}_m(\theta,\alpha)]_n = [\mathbf{b}_m(\theta,0)]_n = \frac{1}{\sqrt{N_t}} e^{jk_m nd\theta}$. Accordingly, the TD and PS beamforming vector can be set as $[\mathbf{w}_m^{\mathrm{TD}}(\theta_t',0)]_n = \frac{1}{\sqrt{N_t}} e^{-jk_m nd\theta_t'}$ and $[\mathbf{w}^{\mathrm{PS}}(\theta_p',0)]_n = \frac{1}{\sqrt{N_t}} e^{-jk_c nd\theta_p'}$, respectively. Therefore, the array gain in (14) can be simplified into

$$G(k_m\theta - k_m\theta_t' - k_c\theta_p', 0) = \Xi_{N_t}\left(\frac{dk_m}{\pi}(\theta - \theta_t' - \frac{f_c}{f_m}\theta_p')\right), \tag{17}$$

where $\Xi_{N_t}(x) = \sin\frac{N_t\pi}{2}x / N_t \sin\frac{\pi}{2}x$ is the Dirchlet sinc function. The far-field array gain $G(x,0)$ is evidently a periodic function of period $2\pi/d$, i.e.,

$$G(x - \frac{2p\pi}{d}, 0) = G(x,0), \ \forall p \in \mathbb{Z}. \tag{18}$$

This periodicity implies that all the points $x = \frac{2p\pi}{d}, p \in \mathbb{Z}$ can maximize the array gain $G(x, 0)$. In this context, the spatial angle θ_m at frequency f_m should be derived from $k_m\theta_m - k_m\theta'_t - k_c\theta'_p = \frac{2p_m\pi}{d}$, as

$$\theta_m = \theta'_t + \frac{f_c}{f_m}(\theta'_p + 2p_m). \tag{19}$$

The physical significance of p_m in (19) is interpreted as follows.

Remark 1 *(Physical significance of integer p): Since the spatial angle θ_m corresponds to an actual physical angle ϑ, θ_m has to satisfy the angle constraint $\theta_m \in [-1, 1]$. To realize the case of $p_m = 0$ as presented in* (15)*, the variables θ'_t and θ'_p need to be carefully designed to make $\theta_m = \theta'_t + \frac{f_c}{f_m}\theta'_p$ satisfy the angle constraint $\theta_m \in [-1, 1]$. However, as long as the delay circuits allow, the variables θ'_t and θ'_p can be arbitrarily adjusted, making it possible to break the angle constraint: $\theta'_t + \frac{f_c}{f_m}\theta'_p \notin [-1, 1]$. In this context, the integer p_m is activated and it will naturally become non-zero to shift the value of $\theta_m = \theta'_t + \frac{f_c}{f_m}(\theta'_p + 2p_m)$ into the physical angle range, $[-1, 1]$. In other words, by adjusting the delay parameters θ'_t and θ'_p, the integer p_m can be automatically controlled to satisfy the angle constraint.*

(19) allows us to illustrate the far-field periodical beam split. To begin with, we adjust the parameters θ'_t and θ'_p such that the beam of the first subcarrier is steered to $\theta_1 = \theta'_t + \frac{f_c}{f_1}\theta'_p = -1$ with $p_1 = 0$. Without loss of generality, we assume $\theta'_t > 0$ and $\theta'_p < 0$. In this case, as the subcarrier frequency f_m grows, the focused direction θ_m gradually increases from -1 to 1 with p_m remaining to 0. Suppose that the value of θ'_p is large enough such that the focused direction θ_m can approach $\theta_{m_1} = \theta'_t + \frac{f_c}{f_{m_1}}\theta'_p = 1$ at the m_1-th subcarrier where $m_1 < M$. Subsequently, if we continue to increase the frequency, then $\theta_{m_1+1} = \theta'_t + \frac{f_c}{f_{m_1+1}}\theta'_p$ definitely exceeds 1. According to **Remark 1**, the integer p_{m_1+1} will automatically become -1 to shift θ_{m_1+1} from $\theta'_t + \frac{f_c}{f_{m_1+1}}\theta'_p$ to $\theta'_t + \frac{f_c}{f_{m_1+1}}(\theta'_p - 2)$ for guaranteeing the angle constraint $\theta_{m_1+1} \in [-1, 1]$. Besides, the beam at frequency f_{m_1+1} is steered to $\theta'_t + \frac{f_c}{f_{m_1+1}}(\theta'_p - 2) \approx -1$. That is to say, due to the periodic property of the angle domain, the spatial direction of the beam "skips" from 1 to -1 when the frequency arises from f_{m_1} to f_{m_1+1}. In addition, with the continuous increment of the frequency, the focused direction of the beam will gradually rise until the focused direction θ_{m_2}, with $m_1 < m_2 < M$, arrives 1 for the second time. Then, the beam will "skip" again to -1 by automatically tuning the integer p_{m_2+1} to -2. This beam skip process will be repeated until the last subcarrier f_M[2]. This phenomenon is called the *periodical beam split*.

Remark 2 *Utilizing far-field periodical beam split, an important insight is that several beams of different frequencies f may focus in the same physical direction. In other words, the beams over the entire bandwidth can periodically scan the angle range* $[-1, 1]$.

[2]If TD-PS parameters are initialized as $\theta'_p + 2p > 0$, the focus direction of the beam will decrease conversely with respect to frequency f_m. Then the beam will skip from -1 to 1 when automatically increasing p.

2) Mechanism of Near-field Distance-Dependent Beam Split: Consider the periodic pattern of beam split in the polar domain. Similar to the periodic far-field array gain in (18), the near-field array gain $G(x, y)$ exhibits a periodicity as well. As proven in [27], $G(x, y)$ is periodic against the vector variable (x, y) with a period of $(\frac{2\pi}{d}, \frac{2\pi}{d^2})$, i.e., $G(x - \frac{2\pi p}{d}, y - \frac{2\pi q}{d^2}) = G(x, y)$, for $\forall p, q \in \mathbb{Z}$. The periodicity implies that $(x, y) = (\frac{2\pi p}{d}, \frac{2\pi q}{d^2})$ are all optimal solutions to maximize the array gain $G(x, y)$. Then the focused location (θ_m, α_m) of the beam at frequency f_m in (15) and (16) should be rewritten as

$$\theta_m = \theta'_t + \frac{f_c}{f_m}(\theta'_p + 2p_m), \tag{20}$$

$$\alpha_m = \alpha'_t + \frac{f_c}{f_m}(\alpha'_p + \frac{2q_m}{d}), \tag{21}$$

which are derived from $k_m\theta - k_m\theta'_t - k_c\theta'_p = \frac{2\pi p_m}{d}$ and $k_m\alpha - k_m\alpha'_t - k_c\alpha'_p = \frac{2\pi q_m}{d^2}$. Clearly, the periodic beam split pattern in the near field is a generalization of (19) in the far field. When $B \ll f_c$, (20) and (21) can be approximated by the first-order Taylor expansion as

$$\theta_m \approx \theta'_t + 2(\theta'_p + 2p) - \frac{\theta'_p + 2p}{f_c}f_m, \tag{22}$$

$$\alpha_m \approx \alpha'_t + 2(\alpha'_p + \frac{2q_m}{d}) - \frac{\alpha'_p + \frac{2q_m}{d}}{f_c}f_m. \tag{23}$$

(22) and (23) suggest that the focused direction/distance of the beam grows linearly with the frequency f_m.

We summarize the mechanism of near-field distance-dependent beam split as follows. Formulas (20) and (21) reveal that TD-PS arrays can control the coverage range of multi-frequency beams over the *angular* and *distance* dimensions simultaneously. Consider the angular dimension at first. The formula (20) has the same expression as (19) in the far-field case. The beam split pattern over the angular dimension in near-field scenarios is identical to that in the far field. Thereafter, the manipulation of parameters θ'_t and θ'_p can be the same as what we have discussed in Section III-A1. We then consider the focused distance of multi-frequency beams presented in (21). The major difference between the beam split pattern over the angular and distance dimensions is attributed to the physical constraints imposed on θ_m and α_m. The physical direction θ_m is bounded between -1 and 1 while the distance ring, α_m, can grow from 0 to infinity. This difference implies that, as opposed to the periodic pattern of the angle-dimension beam split, the focused distance ring α_m increases/decreases monotonically with the frequency f_m, as long as $\alpha'_p + \frac{2q_m}{d} \neq 0$ and $\alpha'_t + \frac{f_c}{f_m}(\alpha'_p + \frac{2q_m}{d}) \geq 0$. Moreover, we can just ignore the subscript m of q_m because q_m remains the same when changing the frequency f_m. For example, if the multi-frequency beams are desired to cover the entire distance range $[\alpha_p, \alpha_q]$, we can simply set

$$\alpha_L = \alpha'_t + \frac{f_c}{f_L}(\alpha'_p + \frac{2q}{d}) \geqslant \alpha_q, \tag{24}$$

$$\alpha_H = \alpha'_t + \frac{f_c}{f_H}(\alpha'_p + \frac{2q}{d}) \leqslant \alpha_p, \tag{25}$$

where $\alpha_p' + \frac{2q}{d}$ is assumed to be greater than 0 without loss of generality.

In summary, by integrating the periodic pattern of the angle-dimension and the monotonous pattern of the distance-dimension, we arrive at the distance-dependent beam split effect. As presented in Fig. 2(a), the focused directions of the multi-frequency beams fluctuate periodically between $[-1, 1]$ while the focused distance ring increases monotonically with the frequency f_m. In such cases, the focused locations $(\theta_m, \alpha_m), m = 1, 2 \dots M$ are distributed over several inclined strips. Each strip covers the entire angle range $[-1, 1]$ while different strips occupy different distance ranges.

Distance-dependent beam split versus traditional near-field beam split: To create the multi-strip distribution of beam split, it is necessary to make $\theta_p' + 2p$ large enough for triggering the rapid fluctuation in the angular domain and make $\alpha_p' + \frac{2q}{d}$ non-zero for changing the distance ring monotonously. These requirements differ the distance-dependent beam split from the traditional counterpart in [27], which are compared in Fig. 2. In [27], by setting $\theta_p' + 2p$ with a *non-zero but relatively small* value, the focused direction of the beams is linearly related to the frequency f_m in *a single period*; while by setting $\alpha_p' + \frac{2q}{d} = 0$, the focused distance ring remains unchanged. Thus, beams of different frequencies focus in different angles in a single distance ring $\alpha = \alpha_t'$, or equivalently the single horizontal strip in Fig. 2(b). Additionally, in Fig. 2(c), we make $\alpha_p' + \frac{2q}{d}$ non-zero but keep the same $\theta_p' + 2p$ as that of Fig. 2(b). This beam-split pattern suggests that both the focused direction and distance ring increase monotonically with the frequency f_m. However, without exploiting the periodicity of the angle-dimension beam split, the beams can only occupy one inclined strip while the multi-strip distribution is still concealed.

B. Distance-Dependent Beam Split based Near-field Beam Training

Applying the discovered distance-dependent beam split phenomenon, it is promising that during beam training, different angles and distances can be searched simultaneously using one pilot. Thereby, a few pilots are capable of covering the whole two-dimensional space. This observation motivates us to design a new near-field beam training method based on this phenomenon. The design principles, the configuration of TD-PS parameters, and detailed beam training procedures are presented one by one in the sequel.

1) Basic design principles: The basic principles for designing near-field beam training aided by distance-dependent beam split involve the consideration of reliability and efficiency.

- **Reliability.** To support reliable beam training, each location inside the communication region must be covered by one training beam. To quantify this metric, we require that each location within the "angle-distance" region lies in the 3 dB-beam width of a certain beam.
- **Efficiency.** In addition to reliability, we aim to use as few pilots as possible to improve training efficiency.

2) Configuration of TD-PS parameters: Building on the above principles, we elaborate on the manipulation of TD-PS parameters to create multi-strip beam pattern for beam training. We start from the manipulation of a single pilot to guarantee the 3 dB coverage in the angle domain. Then we will explain how to utilize interleaved pilots to guarantee the 3 dB coverage in the distance domain.

For the first pilot, let's consider the angle-related parameters θ_t' and θ_p', as well as the integer p_m first. Based on (20), the focused direction difference of adjacent subcarriers on the same strip with $p_m = p_{m+1}$ is

$$|\theta_{m+1} - \theta_m| = |f_c(\theta_p' + 2p_m)(\frac{1}{f_{m+1}} - \frac{1}{f_m})|. \tag{26}$$

For ease of discussion, we assume $\theta_p' + 2p_m$ is always greater than 0 in the sequel, while all our conclusions can be straightforwardly generalized to negative $\theta_p' + 2p_m$ situations. Formula (26) suggests that a larger $\theta_p' + 2p_m$ leads to a faster change in the focused direction versus frequency f_m. Thus, a relatively large $\theta_p' + 2p_m$ can broaden the coverage area of beams, thereby improving beam training efficiency. However, since the direction difference in (26) grows linearly with $\theta_p' + 2p_m$, excessively large $\theta_p' + 2p_m$ will lead to uncovered angular range between adjacent subcarriers. Taking these two factors into consideration, **Lemma 1** provides an upper limit of $\theta_p' + 2p_m$ that can avoid angular coverage hole.

Lemma 1 *Consider an arbitrary physical location $(\bar{\theta}, \bar{\alpha})$. If $\theta_p' + 2p_m > 0$, to make the direction $\bar{\theta}$ lie in the angle-domain 3 dB-beamwidth of a certain beam, θ_p' should satisfy*

$$\theta_p' \in \left\{ \theta_p' : \theta_p' + 2p_m \leqslant \frac{0.88}{N_t}(f_m + f_{m+1})\frac{M}{B}, \forall p_m \in \mathbb{Z} \right\}. \tag{27}$$

Proof: (See Appendix A). ∎

It is easy to prove that when $\theta_p' + 2p_m > 0$, p_m is a non-decreasing function with frequency f_m (i.e. $\theta_p' + 2p_m \leqslant \theta_p' + 2p_M$). Besides, $\frac{0.88}{N_t}(f_m + f_{m+1})\frac{M}{B}$ has minimum value at $f_m = f_L$. Thus, by setting

$$\theta_p' \in \left\{ \theta_p' : \theta_p' + 2p_M \leqslant \frac{1.76 f_L M}{N_t B}, \forall p_M \in \mathbb{Z} \right\}, \tag{28}$$

(27) holds for all subcarriers. Therefore, by setting $p_M = \lfloor \frac{0.88\gamma f_L M}{N_t B} \rfloor$, θ_p' can be determined by

$$\theta_p' = \frac{1.76\gamma f_L M}{N_t B} - 2p_M, \tag{29}$$

where $\gamma \in [0, 1]$ is a hyperparameter. When $\gamma = 1$, the equality in (28) holds, while otherwise, it denotes an oversampling strategy in the angle domain.

While the PS parameter θ_p' influences the direction difference, the TD parameter θ_t' determines the intercept of the strip according to (22). In other words, it affects the starting/ending direction (i.e. θ_1 or θ_M) of this pilot. Without loss of generality, we expect that the ending direction of the beam at frequency f_M is $\theta_M = 1$. Accordingly, the parameter θ_t' can be calculated by

$$\theta_t' = 1 - \frac{f_c}{f_M}(\theta_p' + 2p_M). \tag{30}$$

For future usage, here we also calculate the integer p_1 of the first subcarrier. p_1 should be automatically adjusted to satisfy

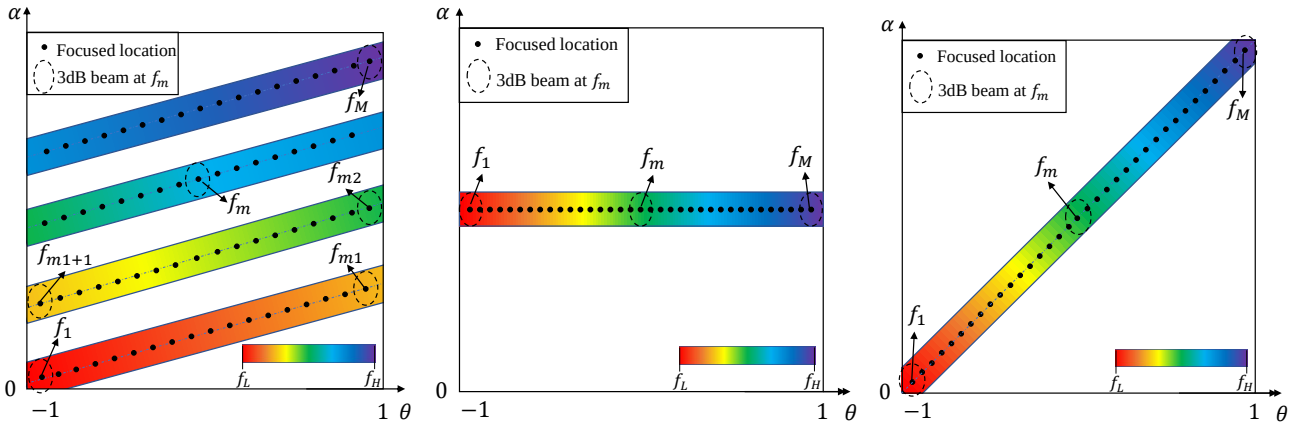


Fig. 2. Comparison of distance-dependent beam split with traditional beam split with different settings: (a) Beams of different subcarriers in proposed distance-dependent beam split; (b) Beams of different subcarriers in traditional beam split with $\alpha_p' + \frac{2q}{d} = 0$; (c) Beams of different subcarriers in traditional beam split with $\alpha_p' + \frac{2q}{d} \neq 0$.

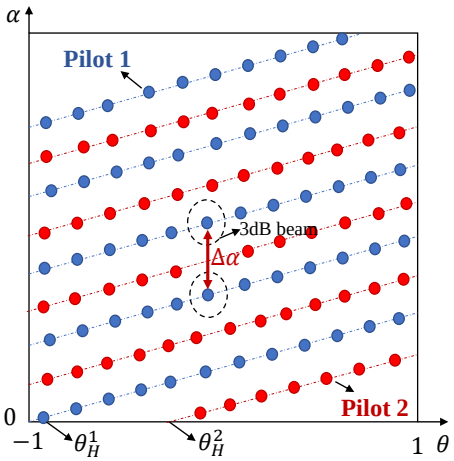


Fig. 3. Add another pilot to guarantee coverage of possible regions.

the angle constraint $\theta_t' + \frac{f_c}{f_L}(\theta_p' + 2p_1) \in [-1, 1]$, so it can be presented by

$$p_1 = \text{round}\left(\frac{-\theta_t'\frac{f_L}{f_c} - \theta_p'}{2}\right). \tag{31}$$

Next, we continue to discuss the manipulation of distance-related parameters α_t' and α_p', and the corresponding integer q. To guarantee that beams can cover the desired distance range, α_t' and α_p' should meet the requirements in (24) and (25). We assume that the possible user distribution distance is $[\alpha_{\min}, \alpha_{\max}]$. By setting $(\alpha_p, \alpha_q) = (\alpha_{\min}, \alpha_{\max})$ and subtracting (25) from (24), $\alpha_p' + \frac{2q}{d}$ is supposed to satisfy

$$\alpha_p' \in \left\{\alpha_p' : \alpha_p' + \frac{2q}{d} \geq \frac{\alpha_{\max} - \alpha_{\min}}{\frac{f_c}{f_L} - \frac{f_c}{f_H}}, \forall q \in \mathbb{Z}\right\}. \tag{32}$$

Therefore, by setting $q = \lfloor \frac{\alpha_{\max} - \alpha_{\min}}{\frac{f_c}{f_L} - \frac{f_c}{f_H}} \frac{d}{2} \rfloor$, α_p' can be set as

$$\alpha_p' = \frac{\alpha_{\max} - \alpha_{\min}}{\frac{f_c}{f_L} - \frac{f_c}{f_H}} - \frac{2q}{d}, \tag{33}$$

After obtaining α_p', the parameter α_t' can be arbitrarily selected from the range

$$\alpha_t' \in [\alpha_{\max} - \frac{f_c}{f_L}(\alpha_p' + \frac{2q}{d}), \alpha_{\min} - \frac{f_c}{f_H}(\alpha_p' + \frac{2q}{d})], \tag{34}$$

to meet the constraints (24) and (25).

Remark 3 *In summary, the parameters designed in* (29), (30), (33), *and* (34) *allow us to create the multi-strip beam pattern in Fig. 2(a) using one single pilot. These beams can guarantee the 3 dB coverage of the entire angular range in each strip.*

However, one remaining issue is that the 3 dB coverage in the distance domain may not be satisfied, since there exists coverage hole between adjacent strips as shown in Fig. 2(a). Specifically, consider two searched locations at the same angle but adjacent strips as illustrated by the circled blue points in Fig. 3. Based on (22) and (23), the angle and distance differences, $\Delta\theta$ and $\Delta\alpha$, of these two locations can be calculated as

$$\Delta\theta = \frac{\theta_p' + 2p_m}{f_c}\Delta f = 2, \tag{35}$$

$$\Delta\alpha = \frac{\alpha_p' + \frac{2q}{d}}{f_c}\Delta f, \tag{36}$$

where Δf denotes the frequency difference. From equations (35) and (36), the distance difference $\Delta\alpha$ can be expressed as

$$\Delta\alpha = \frac{2(\alpha_p' + \frac{2q}{d})}{\theta_p' + 2p_m} \leqslant \frac{2(\alpha_p' + \frac{2q}{d})}{\theta_p' + 2p_1} \triangleq \Delta\alpha^{\text{upper}}. \tag{37}$$

To realize the coverage over the entire distance domain, we expect that the distance difference $\Delta\alpha$ is smaller than the sum of the distance-domain 3 dB-beam width of the two beams, which can be presented as

$$\Delta\alpha \leqslant \eta_\alpha^{f_a} + \eta_\alpha^{f_b}. \tag{38}$$

Here, f_a and f_b refer to the frequency of the corresponding beams and η_α^f is the distance-domain 3 dB-beam width at frequency f. As proven in Appendix B, the beam width η_α^f

has an analytical expression: $\eta_\alpha^f = \frac{4\beta^2 f_c^2}{N_t^2 cf}$, where $\beta = 1.318$.

In practice, however, the constraint (38) is hard to satisfy, indicating that a single pilot is not sufficient to cover all distances. To fill up the coverage hole of a single pilot, we plug in more pilots to create *interleaved multi-strip beam pattern*, as illustrated in Fig. 3. To this aim, recall that the intercept of a strip is dependent on the TD parameter θ_t' according to (22). By changing the value of θ_t', the ending direction θ_M is modified and thus the strips of one pilot are shifted. This observation motivates us to retain the parameters θ_p', α_t', and α_p' but set θ_t' as different values for different pilots to create the interleaved multi-strip beam pattern. We disperse different pilots to create uniformly spaced strips (the circled red points in Fig. 3) to efficiently fill in the uncovered ranges.

To elaborate, we denote the total pilot number as K and introduce the superscript k to index the TD parameter $\theta_t'^k$ and the ending direction θ_M^k of the k-th pilot. The minimal pilot number required for distance coverage is given as follows.

Remark 4 *The K pilots need to fill up the distance-domain coverage hole with an upper length of $\Delta\alpha^{\text{upper}}$. Given that the distance-domain beam width has a lower bound $\eta_\alpha^f \geq \eta_\alpha^{f_H}$, the minimal pilot number is determined by*

$$K \geq \frac{\Delta\alpha^{\text{upper}}}{2\eta_\alpha^{f_H}} = \frac{(\alpha_p' + \frac{2q}{d})N_t^2 c f_H}{4(\theta_p' + 2p_1)\beta^2 f_c^2}. \tag{39}$$

After obtaining the pilot number K, we can uniformly sample the ending directions from the range $[-1, 1]$ such that $\theta_M^k = 1 - \frac{2}{K}(k-1)$ for $k \in \{1, 2, \cdots, K\}$ to ensure uniformly spaced strips. Therefore, $\{\theta_t'^k\}$ for each pilot are given as

$$\theta_t'^k = \theta_M^k - \frac{f_c}{f_M}(\theta_p' + 2p_M), \quad k \in \{1, 2 \cdots, K\}. \tag{40}$$

It is notable that the parameter θ_t' designed in (30) is exactly the first parameter $\theta_t'^1$ designed in (40) since $\theta_M^1 = 1$.

Based on the discussions above, we summarize the TD-PS parameter manipulation procedures in **Algorithm 1**. For Step 1-3, by setting the angle-related parameters of the first pilot, we manipulate the beams to periodically cover the angular range while ensuring the 3-dB coverage in the angle domain. Then based on Step 4, the beams are set to cover the desired distance range. Finally, we utilize interleaved pilots to ensure 3-dB coverage in distance domain, as shown in Step 5-6.

Algorithm 1 TD-PS parameter manipulation for distance-dependent beam-split-based beam training.

Input: N_t, f_c, B, M, $\alpha_{\max}$, $\alpha_{\min}$, and γ.

1: Set $p_M = \lfloor \frac{0.88\gamma f_L M}{N_t B} \rfloor$ and obtain θ_p' by (29);
2: Set $\theta_M^1 = 1$ and obtain $\theta_t'^1$ for the first pilot by (30);
3: Set p_1 as (31);
4: Set $q = \lfloor \frac{\alpha_{\max} - \alpha_{\min}}{\frac{f_c}{f_L} - \frac{f_c}{f_H}} \frac{d}{2} \rfloor$ and obtain α_p' and α_t' based on (33) and (34);
5: Calculate the required pilot number K based on (39);
6: Obtain $\{\theta_t'^k\}$ for the remaining $K-1$ pilots based on (40).

Output: Designed parameters $\{\theta_t'^k\}$, θ_p', α_t' and α_p'.

3) Beam training process: In the beam training process, the BS sequentially transmits K pilots to the user. The received array gains over all subcarriers and pilots are written as

$$\begin{aligned}\boldsymbol{g} =& [g(1,1), g(2,1), \cdots, g(M,1), \\ & \cdots, g(1,K), g(2,K), \cdots, g(M,K)],\end{aligned} \tag{41}$$

where $g(m,k)$ is the array gain at the m-th subcarrier of the k-th pilot. The user selects the subcarrier with the highest received power (i.e. the strongest element of $\boldsymbol{g}(\hat{m}, \hat{k})$) and the estimated physical location can be calculated based on (20) and (21). Moreover, we emphasize that the proposed method can be directly extended to multi-user scenarios. Specifically, since the proposed method only involves K predefined codewords that remain unchanged during beam training, the proposed method can simultaneously perform beam training in multi-user cases.

C. Theoretical Analysis of the Complexity Proposed Method

In this subsection, we analyze the training overhead of the proposed method with respect to antenna number N_t, and compare it with other methods.

Combining (39) with (28) and (32), the required pilot overhead of the proposed method can be calculated as

$$K \geq \frac{N_t^3 c f_H^2 (\alpha_{\max} - \alpha_{\min})}{7.04\gamma\beta^2 f_c^3 M} = \mathcal{O}(\frac{N_t}{M} N_t^2), \tag{42}$$

where N_t, M, f_c, f_H, c and $[\alpha_{\min}, \alpha_{\max}]$ are the antenna number, the subcarrier number, the central carrier frequency, the upper subcarrier frequency, the speed of light, and the distance range of users. Besides, we have also provided the training overhead of near-field exhaustive beam sweeping and near-field rainbow-based method for comparison. For near-field exhaustive beam sweeping, the number of sampled angles is exactly N_t. Moreover, according to [12], the number of sampled distances is calculated as

$$S = \lceil \frac{N_t^2 d^2}{2\beta^2 \lambda_c \rho_{\min}} \rceil. \tag{43}$$

Hence, the training overhead of near-field exhaustive beam sweeping is the product of the sampled direction number and sampled distance number, i.e.

$$K_{ex} \geq \frac{N_t^3 d^2}{2\beta^2 \lambda_c \rho_{\min}} = \mathcal{O}(N_t^3), \tag{44}$$

where $\rho_{\min}$ denotes the minimal users' distance. Besides, the training overhead of near-field rainbow-based method equals the number of sampled distances:

$$K_{rainbow} = S \geq \frac{N_t^2 d^2}{2\beta^2 \lambda_c \rho_{\min}} = \mathcal{O}(N_t^2). \tag{45}$$

Based on the analysis above, we conclude that, as the increase of antenna number, the training overheads of all the methods increase. The main reason is that, with larger antennas, the 3-dB beamwidth can become narrower and the near-field regions become larger. In this case, more pilots are required to search the entire possible region. Specifically, the training overhead of the proposed method, the exhaustive, and the near-field rainbow methods are $\mathcal{O}(\frac{N_t}{M} N_t^2)$, $\mathcal{O}(N_t^3)$, and

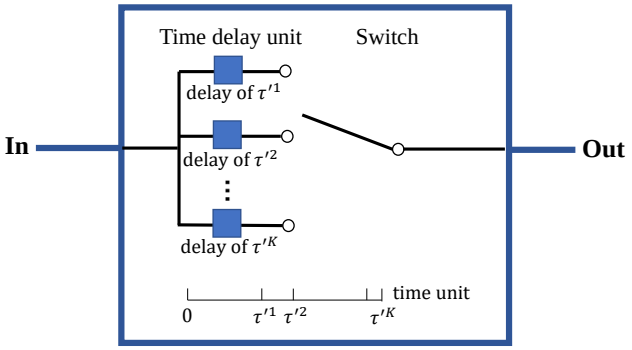


Fig. 4. Hardware implementations of TD by fixed time delay network.

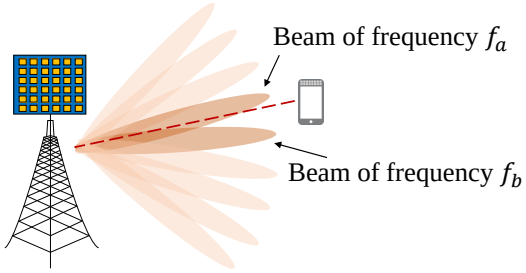


Fig. 5. Auxiliary beam pair-assisted method.

$\mathcal{O}(N_t^2)$, respectively. Here, we emphasize that in practical wideband communication, the subcarrier number is significantly larger than the antenna number. For example, consider a wideband system with bandwidth $B = 3$ GHz, the subcarrier number may be up to $M = 25000$ if the subcarrier spacing is 120 kHz (as in 5G standard). In contrast, the antenna number N_t for XL-MIMO systems is usually several hundred or thousand. Thus, we obtain $M \gg N_t$, which indicates that the proposed methods can significantly reduce the training overhead compared to existing methods.

D. Hardware Implementations of Proposed Method

This subsection introduces the hardware implementation of the TD-PS structure to realize the proposed beam training algorithm. In particular, we mainly focus on the expensive TD circuits. The function of *time delay* can be implemented mainly by two kinds of approaches, namely the adjustable true-time-delay-based methods [33], [34] and fixed true-time-delay-based methods [35], [36]. The former one can provide continuously adjustable time delay over the entire range with significantly high power consumption and hardware complexity [36]. In contrast, the available time-delays of fixed true-time-delay-based methods are restricted to several fixed values. Thus, this architecture has much lower power consumption and hardware complexity, with the sacrifice of TD resolution.

Fortunately, the low-cost fixed true-time-delay-based methods could adequately satisfy the requirements for realizing the proposed beam training algorithm. To be specific, the scanned positions of beams are fixed and the corresponding time delay values are elaborately designed and can be determined and calculated in advance. Besides, only K (usually a small number such as 2,3) time delay values need to be realized by TD. Therefore, as shown in Fig. 4, we propose to deploy a nonuniform-quantization TD with fixed time delay unit and a switch network, which can be implemented by an extensively-utilized microstrip line or waveguide of fixed length [35]. For the n-th TD, we calculate the K desired time delay value as $\tau^{(n)'k}, k = 1, 2, \cdots, K$:

$$\tau^{(n)'} = \frac{r_1^{(n)'}}{c} = \frac{nd\theta_t' - n^2d^2\alpha_t'}{c}. \tag{46}$$

Then we construct the TD with the fixed TD comprising K fixed time delay units and a switch, indicating the selected fixed TD in different time slots. In this context, only $\log_2 K$ bits are required to represent the selected time delay. Thus, the required resolution of fixed TD is quite low, which highlights the advantage of the low hardware complexity of our proposed method.

Note that the incorporation of TDs may introduce additional Insertion Loss (IL) into the link budget, indicating the signal attenuation experienced as it traverses the hardware device. For the proposed implementation, the IL associated with the TD is the cumulative effect of the insertion losses from one fixed time delay unit and one switch, i.e., $IL_{FTD} + IL_{switch}$, which is significantly lower than adjustable TDs [35]. Besides, as discussed in [37], additional IL introduced by fixed TD may not have significant influence on beamforming gain and it can be compensated by improving the amplification factor of the power amplifiers.

IV. Improved Beam Training for Off-grid Users

Compared to the method in [27], the sampled density of direction within $[-1, 1]$ of the proposed method in Section III is relatively low, which limits its performance for off-grid users (i.e. users not on the searched locations). In this section, we propose two algorithms, namely the auxiliary beam pair-assisted method and the match filter-based method, to solve this problem. The core idea of the proposed methods is to employ the received power of multiple subcarriers, rather than only the one with the highest received power to jointly estimate the users' locations.

A. Auxiliary beam pair-assisted method

As shown in Fig. 5, when the user is located between two beams of frequency f_a and f_b, the scheme in Section III, which chooses the codeword having the highest received power at frequency f_a, is suboptimal. Inspired by [38], [39], we design an auxiliary beam pair-assisted beam training method based on near-field distance-dependent beam split. The method utilizes the amplitude distribution of the auxiliary beam pair, f_a and f_b, to accurately estimate the angle and distance of off-grid users. To this aim, we will first analyze the amplitude distribution of near-field beams, and then the beam training method is elaborated.

1) Analysis of amplitude distribution of near-field beams: The equal-power contour of near-field beams can be clarified by the following **Proposition 1**.

Proposition 1 *The equal-power contour of near-field beams can be approximated as an ellipse equation:*

$$G(k(\theta-\bar{\theta}),k(\alpha-\bar{\alpha}))=1-\sigma_1(\theta-\bar{\theta})^2-\sigma_2(\alpha-\bar{\alpha})^2, \quad (47)$$

where (θ,α) *is the focused location of the beam,* $(\bar{\theta},\bar{\alpha})$ *is an arbitrary physical location within the curve of* 3 dB*-beam width. Parameters* σ_1,σ_2 *are presented as*

$$\sigma_1=\frac{1}{24}N_t^2\pi^2\frac{f^2}{f_c} \quad \text{and} \quad \sigma_2=\frac{\pi^2}{90}\frac{N_t^4d^4}{\lambda^2}. \quad (48)$$

Proof: (See Appendix C). ∎

2) Procedures of auxiliary beam pair-assisted method: According to the analysis above, the received array gain of the beam at frequency f_a and f_b each defines an ellipse in which the user is located. The angle and distance to be estimated can be obtained through the intersection of the ellipses. This is the basic idea of the auxiliary beam pair-assisted method. Then we will illustrate the procedures step by step as follows.

- Firstly, we perform the beam training proposed in Section III to obtain subcarrier f_a with the highest received power. Suppose we can obtain the value of the highest power P_a, we calculate the array gain as G_a. Besides, we denote the focus location of the beam as (θ_a,α_a).
- Then we compare the received power at subcarriers f_{a-1} and f_{a+1}, and select the one with higher power, denoted as f_b. Similarly, record the focus location (θ_b,α_b) and array gain as G_b.
- The corresponding equal-amplitude ellipse defined by subcarrier f_a and f_b can be written as

$$\begin{cases}\sigma_1(\theta_a-\theta)^2+\sigma_2(\alpha_a-\alpha)^2 &= 1-G_a\\ \sigma_1(\theta_b-\theta)^2+\sigma_2(\alpha_b-\alpha)^2 &= 1-G_b\end{cases}, \quad (49)$$

  based on which the estimated $\hat{\theta}$ and $\hat{\alpha}$ can be obtained using the Newton method.

B. Match filter-based method

The auxiliary beam pair-assisted beam training utilizes a pair of beams to improve the training performance for off-grid users. In this subsection, we generalize it to a match filter-based beam training method, which advocates the use of array gains over all subcarriers and pilots for accurate beam training.

In this method, a substantial number of grids with different angles and distances are sampled offline. For each grid, the distribution of array gain over different subcarriers and pilots can be calculated by the location of the grid. Then, in the online stage, we select the grid whose array gain distribution matches the best with the received array gain distribution of the user. Since the sampled grids can be much denser than the searched locations, the performance for off-grid users is improved. Besides, the features of match filter also protect training from noise. The framework is elaborated below.

- **Offline stage**: Sample the possible user locations by L grids on the angle domain and S grids on the distance domain. The sampled grids are supposed to be denser than the searched locations by the proposed beam training method. For the sampled location (θ_l,α_s), the array gain at the m-th subcarrier of the k-th pilot is presented as

$$g_{l,s}(m,k) = |(\mathbf{w}_m^{\mathrm{TD}}(\theta_t'^k,\alpha_t')\odot\mathbf{w}^{\mathrm{PS}}(\theta_p',\alpha_p'))^T\mathbf{b}_m(\theta_l,\alpha_s)|. \quad (50)$$

  Then, the corresponding array gain vector, $\boldsymbol{g}_{l,s}$, is

$$\boldsymbol{g}_{l,s}=[g_{l,s}(1,1),g_{l,s}(2,1),\cdots,g_{l,s}(M,1), \cdots,g_{l,s}(1,K),g_{l,s}(2,K),\cdots,g_{l,s}(M,K)]. \quad (51)$$

- **Online stage**: For each of the $L\times S$ grids, the similarity between the array gain vector of location (θ_l,α_s) and the received array gain vector of the user is calculated as

$$r_{l,s}=\boldsymbol{g}^T\boldsymbol{g}_{l,s}, s=1,2,\cdots,S, l=1,2,\cdots,L. \quad (52)$$

  Finally, we select the grid with largest similarity $r_{l,s}$ and the estimated location of the user is $\hat{\theta}=\theta_l$ and $\hat{\alpha}=\alpha_s$.

To reduce feedback overhead, the users are supposed to estimate $\hat{\theta}$ and $\hat{\alpha}$ locally and feed back the results to the BS, rather than feeding back all the received power to the BS.

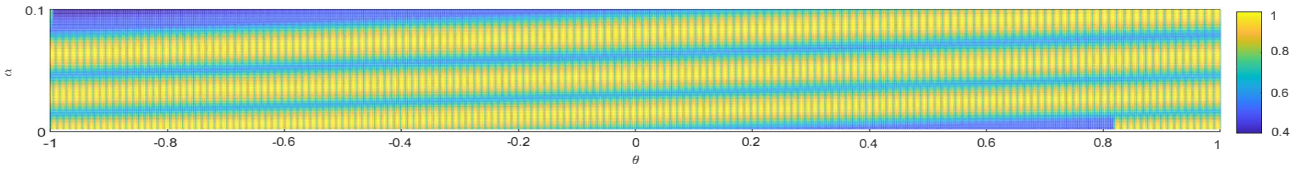
V. Simulation Results

In this section, we present simulations to validate the proposed distance-dependent beam split phenomenon. Subsequently, the performance of the proposed distance-dependent beam split-based beam training schemes, encompassing both on-grid algorithms and off-grid algorithms, is assessed.

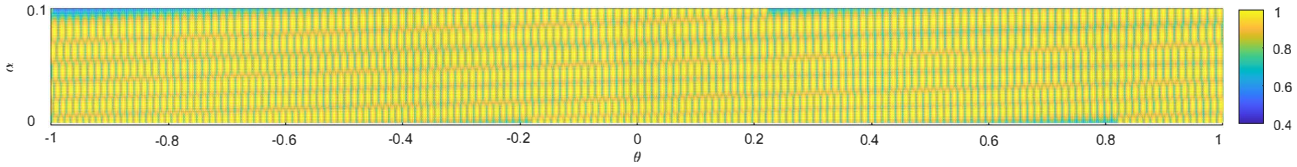
A. The Demonstration of Distance-dependent Beam Split

We first verify that the designed TD-PS parameters in **Algorithm 1** are capable of producing distance-dependent near-field beam split. As an example, we suppose the BS is equipped with $N_t=128$ antennas. The carrier frequency is $f_c=10$ GHz, the number of subcarriers is $M=512$, and the bandwidth is $B=2$ GHz. The users are uniformly distributed within angle range $[-\pi/3,\pi/3]$, and distance range $[\rho_{\min},\rho_{\max}]=[5\text{ m},200\text{ m}]$, corresponding to a distance ring range $[\alpha_{\min},\alpha_{\max}]=[1/400,1/10]$. Note that if the users only locate within part of the angular range, $[-\pi/3,\pi/3]$ in this paper for instance, we only inject transmit power to $\sqrt{3}/2$ subcarriers whose focused direction is within $[-\pi/3,\pi/3]$ and shut off the wasted subcarriers. In this way, we can fully exploit the transmit power and improve the efficiency of the proposed method. According to **Algorithm 1**, the TD-PS parameters are defined as follows. Based on (28), $\theta_p'+2p_M$ should satisfies $\theta_p'+2p_M\leqslant\frac{1.76f_L}{N_t}\frac{M}{B}=31.68$ and we set $\theta_p'=1.68$, $p_M=15$ with $\gamma=1$. Then (30) and (31) determine that $\theta_t'^1-27.8$ and $p_1=12$. Then based on (32), $\alpha_p'+\frac{2q}{d}\geqslant\frac{\alpha_{\max}-\alpha_{\min}}{\frac{f_c}{f_L}-\frac{f_c}{f_H}}=0.483$, and thus α_p' is set as 0.5 with $q=0$ here. Accordingly, α_t' is manipulated as $\alpha_t'=-0.454\in[\alpha_{\max}-\frac{f_c}{f_L}(\alpha_p'+\frac{2q}{d}),\alpha_{\min}-\frac{f_c}{f_H}(\alpha_p'+\frac{2q}{d})]=[-0.456,-0.452]$.

We respectively depict the beam pattern of individual subcarrier generated by the designed beamformer in the same figure. Thus, for an arbitrary location in the (θ,α) domain, the quantity in the figure denotes the highest beam gain of all the subcarriers at this location. Then as discussed in Section. III-B,



(a) Distance-dependent beam split with a single pilot.



(b) Distance-dependent beam split with two pilots.

Fig. 6. Illustration of near-field distance-dependent beam split.

if the normalized beam gain $G > \frac{1}{\sqrt{2}}$ at a specific location, it means that the location lies within the 3 dB-beam width of a certain beam. In other words, the location has been effectively searched by the pilot.

Fig. 6 (a) presents the distribution of beams over M subcarriers of a single pilot generated by the TD-PS beamformer. As we expect, the focused points of beams create a multi-strip beam pattern, validating the effectiveness of distance-dependent beam split. It is also observable that the beams of a single pilot in Fig. 6 (a) cannot cover the entire user region, leaving a coverage hole between strips (i.e. there are locations whose highest beam gain is lower than $\frac{1}{\sqrt{2}}$). According to (39), the minimum pilot number required is calculated as $K = 2$. Thereby, we can add another pilot with the parameter $\theta_t'^2 = -28.8$ to make the second ending direction be $\theta_M^2 = 0$ based on (40). The beam patterns of the two pilots are plotted together in Fig. 6 (b). In this context, the entire possible region is thoroughly searched, guaranteeing the reliability of beam training.

B. Beam Training performance

We then evaluate the performance of the proposed distance-dependent beam-split-based near-field beam training. The default simulation settings are as follows. The BS equips an $N_t = 256$-element ULA while the carrier frequency is $f_c = 30$ GHz. The bandwidth is $B = 5$ GHz with subcarrier number $M = 1024$. The users are uniformly distributed within angle range $[-\pi/3, \pi/3]$, and distance range $[\rho_{\min}, \rho_{\max}] = [5\text{ m}, 200\text{ m}]$, corresponding to a distance ring range $[\alpha_{\min}, \alpha_{\max}] = [1/400, 1/10]$. According to **Algorithm 1**, we employ $K = 3$ pilots for beam training. The TD-PS parameters are set as $\theta_p' = 0.784$, $\{\theta_t'^k\} = \{-32.95, -33.62, -34.29\}$, $\alpha_p' = 0.58$, and $\alpha_t' = -0.53$, while the corresponding $\gamma = 0.95$, $p_1 = 15$, $p_M = 18$ and $q = 0$. To evaluate the effectiveness of beam training, we employ the average rate performance as a metric:

$$\begin{aligned} R &= \frac{1}{M}\sum_{m=1}^{M}\log_2\left(1+\frac{P_t}{\sigma^2}\left\|\mathbf{h}_m^T\mathbf{w}_m\right\|^2\right) \\ &= \frac{1}{M}\sum_{m=1}^{M}\log_2\left(1+\frac{P_tN_t\beta_m^2}{\sigma^2}\left\|\mathbf{b}_m^T\left(\theta_0,r_0\right)\mathbf{w}_m\right\|^2\right) \end{aligned} \tag{53}$$

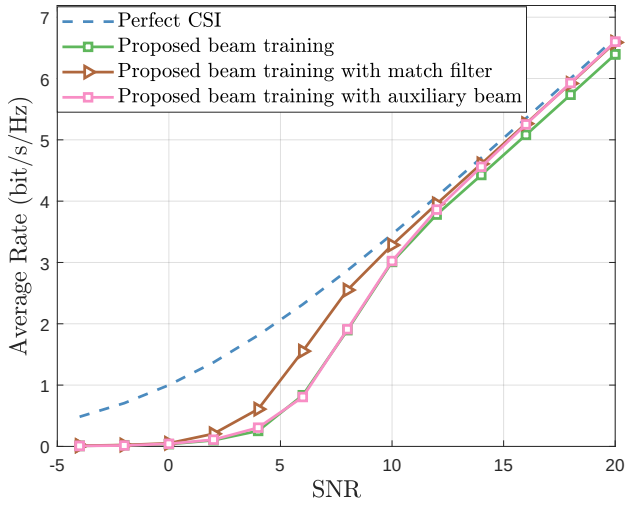


Fig. 7. Average rate performance vs. SNR.

where $\text{SNR} = \frac{P_tN_t\beta_m^2}{\sigma^2}$. The simulation results below are obtained by 1000 realizations of Monte Carlo trails.

To begin with, Fig. 7 presents the average rate performance against SNR ranging from 5 dB to 20 dB. The proposed beam training algorithms for both on-grid and off-grid cases are included. The training overhead is $K = 3$. The sampled grids in the angle and distance domain are $L = 1024$ and $S = 10$. When the SNR is relatively high (e.g. SNR > 10 dB), both the proposed off-grid algorithms can improve the performance of distance-dependent beam split-based beam training performance. While at low SNR scenarios, the performance of the auxiliary beam pair-assisted method is almost the same as the on-grid method. This may result from the fact that an inaccurate beam gains G_a and G_b with noise poses the "error accumulation" phenomena when calculating the estimated $\hat{\theta}$ and $\hat{\alpha}$. Therefore, the auxiliary beam pair-assisted method might be vulnerable to noise. In contrast, thanks to the capability of restraining noise interference, match-filter-based method can ensure significant performance improvement for off-grid users even in low SNR regimes.

Then we compare the proposed distance-dependent beam-

split-based near-field beam training enhanced by match filter with the following benchmarks (BT denotes beam training in the following figures):

- **Perfect CSI:** BS has access to the perfect channel vector $\mathbf{h}_m$, which establishes the upper bound for performance.
- **Near-field exhaustive beam training:** This method conducts an exhaustive search across all potential directions and distances to identify the optimal codeword.
- **Near-field rainbow based beam training:** As mentioned above, the near-field rainbow-based method [27] searches multiple angles in a single distance ring with different subcarriers concurrently and sequentially searches through various distance rings exhaustively.
- **Far-field rainbow-based beam training:** Proposed in [40], this method employs TD beamforming to concurrently search for the optimal angle across different subcarriers while disregarding the distance information.
- **Near-field hierarchical beam training:** The method [14] searches from low-resolution codebooks to high-resolution ones, which gradually reduces the angle and distance estimation granularity layer-by-layer.
- **Inference from far-field beam training:** The method [18], [23] first employs far-field beam sweeping to estimate the angles. Then a near-field codebook is utilized to search for distance within the candidate angles.

Table I summarizes the required pilot overheads for different techniques. We sample $N_t = 256$ grids for the angle domain and $S = 10$ grids for the distance ring. Since for the near-field exhaustive beam training all codewords in the polar-domain codebook should be sequentially tested, the training overhead is up to $N_t S = 2560$. Besides, the pilot overhead of the near-field rainbow [27] equals the number of sampled distance $S = 10$. The training overhead of near-field hierarchical beam training can be written as $\sum U_l S_l$, where U_l, S_l denote the sampled directions and distances to be tested in the l-th layer, respectively. Take [14] for instance, the required codewords in three layers are 32×4, 2×4, and 4×1. Thus, the total overhead is 140. As for the two-stage near-field beam training method in [18], [23], the far-field beam sweeping for angle estimation takes up N_t time slots while the near-field beam sweeping for distance estimation takes up S time slots. The accumulative overhead is reduced to $N_t + S = 266$. Moreover, the pilot number is $K = 3$ for proposed distance-dependent beam split-based beam training. As for the far-field rainbow technique, the optimal angle can be obtained with only a single pilot through TD beamforming, while the distance information is neglected.

Fig. 8 presents the average rate performance against the training overhead. The SNR is set as 15 dB. The training overhead ranges from 0 to 2048. In each time slot, we employ the optimal beamforming vector identified in prior time slots to serve the user. The far-field rainbow-based technique requires only one pilot. However, since the far-field approach focuses solely on angle information and disregards distance information, its average rate performance is suboptimal. Furthermore, near-field exhaustive scheme requires more than 2000 pilots to attain a satisfactory performance, which is impractical for

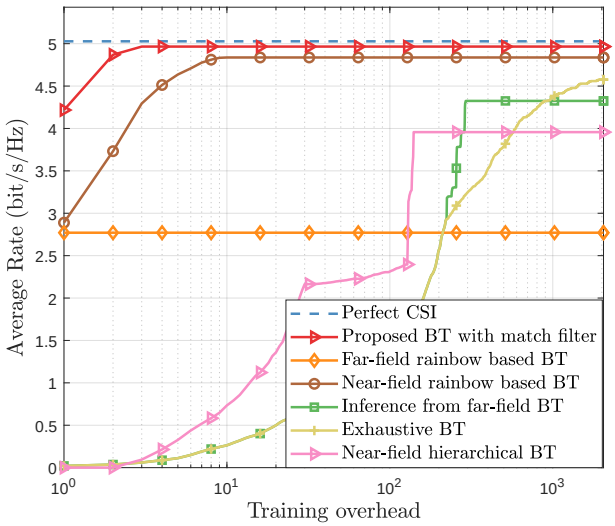


Fig. 8. Average rate performance vs. training overhead.

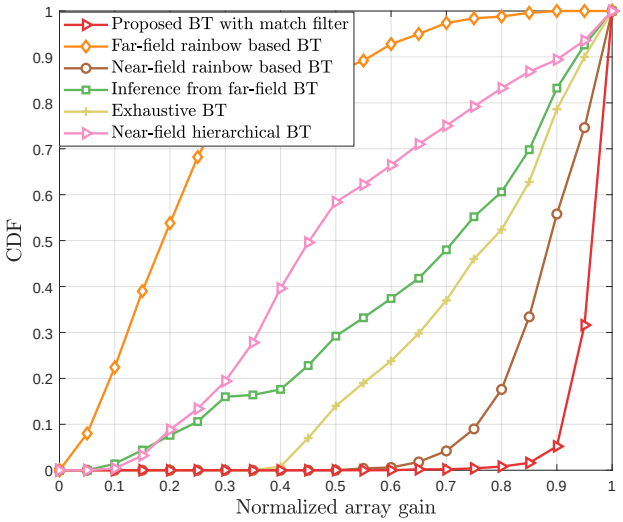


Fig. 9. CDF vs. Normalized array gain.

real-world systems. Narrow-band methods, both hierarchical beam training and inference with far-field beam training, can to some extent reduce the training overhead. Nevertheless, the methods still consume 140 and 266 time slots to acquire accurate beam training, since they do not fully exploit the frequency resources. Then we focus on the comparison of near-field rainbow-based beam training [27] and the proposed method. For fair comparison, the number of sampled directions for the match filter is set as $L = 1024$ to guarantee the same sample rate in the angle domain with near field rainbow [27] (the number of sampled directions is $M = 1024$) while the number of sampled distances is both $S = 10$. As depicted in Fig. 8, the proposed method employs only 3 pilots to achieve comparable average rate performance while near-field rainbow-based method [27] requires 10. Therefore, since both the angle and distance domain can be searched with different subcarriers simultaneously, the proposed method further improves the near-field beam training efficiency.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2024.3507795

TABLE I
COMPARISONS OF OVERHEADS FOR DIFFERENT SCHEMES

Schemes	Training Overheads	Values
Near-field exhaustive beam training	N_tS	2560
Near-field rainbow based beam training	S	10
Far-field rainbow based beam training	1	1
Near-field hierarchical beam training	$\sum U_lS_l$	140
Inference from far-field beam training	N_t+S	266
Proposed distance-dependent beam split based beam training	K	3

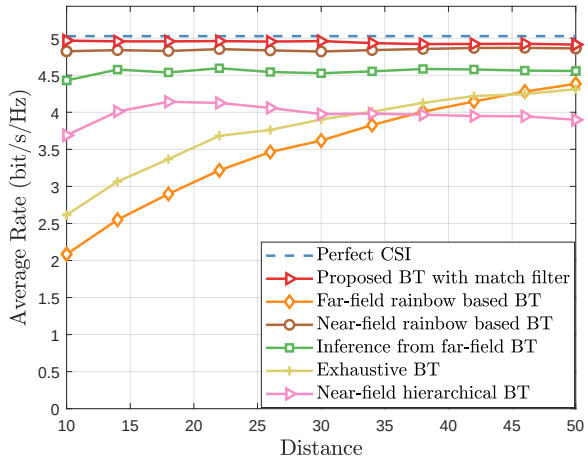


Fig. 10. Average rate performance vs. distance.

In Fig. 9, cumulative probability functions (CDFs) of normalized array gains are plotted. Fig. 9 reveals the improvements in array gains with our proposed method. Specifically, the harvested array gain of far-field method is relatively low without considering near-field effect. The near-field narrowband methods can significantly enhance the performance but at a higher complexity. Wideband methods, both [27] and proposed method, enable more reliable beam training with low overhead. For our proposed method, more than 95% users can achieve the normalized array gain of more than 0.9, which verifies the effectiveness of our proposed method.

Fig. 10 illustrates the average rate performance versus the distance. We have capped the maximum training overhead for all the methods under consideration as 300 and set SNR as 15 dB. Additionally, the distance between the user and BS increases from 10 to 50 meters. As illustrated in Fig. 10, the proposed method with match filter achieves near-optimal rate performance with lower training overheads compared to near-field rainbow-based method and near-field narrowband methods. Regarding the far-field scheme, as the distance decreases, near-field propagation becomes more prominent, leading to a rapid degradation in the average rate for the far-field scheme. Moreover, due to limited maximum training overhead, the exhaustive scheme faces difficulties in searching for locations within the near field.

VI. CONCLUSIONS

In this paper, we demonstrate a phenomenon called distance-dependent beam split where we can divide the subcarriers into different groups which periodically cover the angular range with different distance ranges by elaborately manipulation of the TD-PS parameters. Then a wideband beam training method has been proposed based on the new discovery and we further improve the performance for off-grid users. The simulation results have confirmed the efficacy of the proposed method, which emerges as a promising approach for achieving efficient beam training in XL-MIMO systems. The paper also presents the first attempt to fully unleash the capability of near-field beam split. Future research could concentrate on extending the distance-dependent beam-split to uniform planar array configurations. In addition, investigation of the low-hardware-complexity TD-PS structure with a small number of TDs [41] is also left for future research.

APPENDIX

A. *Proof of* ***Lemma 1***

Consider a beam focused on the location (θ,α) at frequency f. To evaluate the beam coverage in the angle domain, we can assume that $\alpha=\bar{\alpha}$. Then, the achieved beamforming gain on the location $(\bar{\theta},\bar{\alpha})$ is expressed as

$$G(k(\theta-\bar{\theta}),0)=\left|\frac{\sin\frac{N_t\pi(\theta-\bar{\theta})f}{2f_c}}{N_t\sin\frac{\pi(\theta-\bar{\theta})f}{2f_c}}\right|\overset{(a)}{\approx}\left|\frac{\sin\frac{N_t\pi(\theta-\bar{\theta})f}{2f_c}}{\frac{N_t\pi(\theta-\bar{\theta})f}{2f_c}}\right|, \quad (54)$$

where (a) holds because N_t is very large. In this case, 3 dB-beam width in the angle domain can be numerically solved from $|G(k(\theta-\bar{\theta}),0)|=\frac{1}{\sqrt{2}}$:

$$|\theta-\bar{\theta}|\approx\frac{0.88f_c}{N_tf}. \quad (55)$$

To guarantee the entire coverage in the angle domain according to the reliability criterion, the angle difference of θ_m and θ_{m+1} should be smaller than the sum of the 3 dB-beam width of the beam at the m-th and $m+1$-th subcarrier, i.e.,

$$|\theta_{m+1}-\theta_m|\leqslant\eta_\theta^{f_m}+\eta_\theta^{f_{m+1}}=\frac{0.88}{N_t}(\frac{f_c}{f_{m+1}}+\frac{f_c}{f_m}), \quad (56)$$

where η_θ^f is the angle-domain 3 dB-beam width of the beam at frequency f. Combining the above equation and (26) we obtain

$$|f_c(\theta_p'+2p_m)(\frac{1}{f_{m+1}}-\frac{1}{f_m})|\leq\frac{0.88}{N_t}(\frac{f_c}{f_{m+1}}+\frac{f_c}{f_m}), \quad (57)$$

which gives rise to the results of (27) in **Lemma 1**.

B. Derivation of the distance-domain 3 dB-beam width

In (14), when $\theta = \bar{\theta}$, the array gain on an arbitrary physical distance $\bar{\alpha}$ is

$$G(0, k(\alpha - \bar{\alpha})) = \frac{1}{N_t} \left| \sum_{n=-N}^{N} e^{-jn^2 d^2 k(\alpha - \bar{\alpha})} \right| = F(\beta) \quad (58)$$

where $F(\beta) \triangleq |C(\beta) + jS(\beta)|/\beta$, and $C(\beta) = \int_0^\beta \cos(\frac{\pi}{2}t^2)\mathrm{d}t$ and $S(\beta) = \int_0^\beta \sin(\frac{\pi}{2}t^2)\mathrm{d}t$ are Fresnel functions. The variable β is written as $\beta = \sqrt{\frac{N_t^2 d^2}{2\lambda}(\alpha - \bar{\alpha})}$ [12]. For 3 dB-beam width, it can be numerically solved from $|F(\beta)| = \frac{1}{\sqrt{2}}$ that $\beta \approx 1.318$. Thus, the distance-domain 3 dB-beam width at frequency f can be denoted as

$$\eta_\alpha^f = |\alpha - \bar{\alpha}| = \frac{\beta^2 \lambda}{N_t^2 d^2} = \frac{4\beta^2 f_c^2}{N_t^2 c f}. \quad (59)$$

C. Proof of **Proposition 1**

We use the product the angle-domain array gain in (54) and distance-domain array gain in (58) to approximate the array gain (14) on an arbitrary physical location $(\bar{\theta}, \bar{\alpha})$ as

$$G(k(\theta - \bar{\theta}), k(\alpha - \bar{\alpha})) \approx \Xi_{N_t}\left(\frac{(\theta - \bar{\theta})f}{f_c}\right) F(\beta). \quad (60)$$

Based on Taylor expansion, if $\bar{\theta}, \bar{\alpha}$ is within 3 dB-beam width of mainlobe (i.e. $\theta - \bar{\theta} \to 0, \alpha - \bar{\alpha} \to 0$), two parts of G can be written as

$$\Xi_{N_t}\left(\frac{(\theta - \bar{\theta})f}{f_c}\right) \approx 1 - \frac{1}{24} N_t^2 \pi^2 \frac{f^2}{f_c}(\theta - \bar{\theta})^2 \quad (61)$$

$$F(\beta) \approx 1 - \frac{\pi^2}{90} \frac{N_t^4 d^4}{\lambda^2}(\alpha - \bar{\alpha})^2. \quad (62)$$

Combining (60)-(62) and neglecting the high-order small numbers, we arrive at (47).

REFERENCES

[1] P. P. Ray, N. Kumar, and M. Guizani, "A vision on 6G-enabled NIB: Requirements, technologies, deployments, and prospects," *IEEE Wireless Commun.*, vol. 28, no. 4, pp. 120–127, May 2021.

[2] Z. Wang, J. Zhang, H. Du, D. Niyato, S. Cui, B. Ai, M. Debbah, K. B. Letaief, and H. V. Poor, "A tutorial on extremely large-scale MIMO for 6G: Fundamentals, signal processing, and applications," *IEEE Commun. Surv. Tutor.*, 2024.

[3] 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.

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

[5] H. Lu and Y. Zeng, "Communicating with extremely large-scale array/surface: Unified modeling and performance analysis," *IEEE Trans. Wireless Commun.*, vol. 21, no. 6, pp. 4039–4053, Jun. 2022.

[6] H. Tataria, M. Shafi, A. F. Molisch, M. Dohler, H. Sjöland, and F. Tufvesson, "6G wireless systems: Vision, requirements, challenges, insights, and opportunities," *Proc. IEEE*, vol. 109, no. 7, pp. 1166–1199, Jul. 2021.

[7] W. Liu, C. Pan, H. Ren, F. Shu, S. Jin, and J. Wang, "Low-overhead beam training scheme for extremely large-scale RIS in near field," *IEEE Trans. Commun.*, vol. 71, no. 8, pp. 4924–4940, Aug. 2023.

[8] X. Gao, L. Dai, Z. Chen, Z. Wang, and Z. Zhang, "Near-optimal beam selection for beamspace mmwave massive MIMO systems," *IEEE Commun. Lett.*, vol. 20, no. 5, pp. 1054–1057, May 2016.

[9] A. Alkhateeb, G. Leus, and R. W. Heath, "Limited feedback hybrid precoding for multi-user millimeter wave systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6481–6494, Nov. 2015.

[10] M. Ke, Z. Gao, Y. Wu, X. Gao, and R. Schober, "Compressive sensing-based adaptive active user detection and channel estimation: Massive access meets massive MIMO," *IEEE Trans. Signal Process.*, vol. 68, pp. 764–779, Jan. 2020.

[11] S. H. Lim, S. Kim, B. Shim, and J. W. Choi, "Efficient beam training and sparse channel estimation for millimeter wave communications under mobility," *IEEE Trans. Commun.*, vol. 68, no. 10, pp. 6583–6596, Oct. 2020.

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

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

[14] Y. Lu, Z. Zhang, and L. Dai, "Hierarchical beam training for extremely large-scale MIMO: From far-field to near-field," *IEEE Trans. Commun,*, 2023.

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

[16] X. Shi, J. Wang, Z. Sun, and J. Song, "Spatial-chirp codebook-based hierarchical beam training for extremely large-scale massive MIMO," *IEEE Trans. Wireless Commun.*, 2023.

[17] X. Zhang, H. Zhang, C. Li, Y. Huang, and L. Yang, "Environment-specific beam training for extremely large-scale MIMO systems via contrastive learning," *IEEE Commun. Lett.*, vol. 27, no. 10, pp. 2638–2642, Oct. 2023.

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

[19] K. Chen, C. Qi, and C.-X. Wang, "Two-stage hybrid-field beam training for ultra-massive mimo systems," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, 2022, pp. 1074–1079.

[20] W. Liu, H. Ren, C. Pan, and J. Wang, "Deep learning based beam training for extremely large-scale massive MIMO in near-field domain," *IEEE Commun. Lett.*, vol. 27, no. 1, pp. 170–174, Jan. 2023.

[21] W. Liu, C. Pan, H. Ren, F. Shu, S. Jin, and J. Wang, "Low-overhead beam training scheme for extremely large-scale RIS in near field," *IEEE Trans. Commun.*, vol. 71, no. 8, pp. 4924–4940, Aug. 2023.

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

[23] S. Hu, H. Wang, and M. C. Ilter, "Design of near-field beamforming for large intelligent surfaces," *IEEE Trans. Wireless Commun.*, vol. 23, no. 1, pp. 762–774, 2024.

[24] K. Chen, C. Qi, C.-X. Wang, and G. Y. Li, "Beam training and tracking for extremely large-scale mimo communications," *IEEE Trans. Wireless Commun.*, vol. 23, no. 5, pp. 5048–5062, May. 2024.

[25] K. Chen, C. Qi, O. A. Dobre, and G. Ye Li, "Triple-refined hybrid-field beam training for mmwave extremely large-scale mimo," *IEEE Trans. Wireless Commun.*, vol. 23, no. 8, pp. 8556–8570, Aug. 2024.

[26] M. Cui and L. Dai, "Near-field wideband beamforming for extremely large antenna arrays," *IEEE Trans. Wireless Commun.*, pp. 1–1, 2024.

[27] M. Cui, L. Dai, Z. Wang, S. Zhou, and N. Ge, "Near-field rainbow: Wideband beam training for XL-MIMO," *IEEE Trans. Wireless Commun.*, vol. 22, no. 6, pp. 3899–3912, Jun. 2023.

[28] 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.

[29] H. Elayan, O. Amin, B. Shihada, R. M. Shubair, and M.-S. Alouini, "Terahertz band: The last piece of rf spectrum puzzle for communication systems," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 1–32, 2020.

[30] T. S. Rappaport, Y. Xing, O. Kanhere, S. Ju, A. Madanayake, S. Mandal, A. Alkhateeb, and G. C. Trichopoulos, "Wireless communications and applications above 100 GHz: Opportunities and challenges for 6G and beyond," *IEEE Access*, vol. 7, pp. 78 729–78 757, 2019.

[31] B. Wang, F. Gao, S. Jin, H. Lin, and G. Y. Li, "Spatial- and frequency-wideband effects in millimeter-wave massive mimo systems," *IEEE Trans. Signal Process.*, vol. 66, no. 13, pp. 3393–3406, Jul. 2018.

[32] L. Dai, J. Tan, Z. Chen, and H. V. Poor, "Delay-phase precoding for wideband thz massive mimo," *IEEE Transactions on Wireless Communications*, vol. 21, no. 9, pp. 7271–7286, Sep. 2022.

[33] X. Zhao, U. Shah, O. Glubokov, and J. Oberhammer, "Micromachined subterahertz waveguide-integrated phase shifter utilizing supermode propagation," *IEEE Trans. Microw. Theory Techn.*, vol. 69, no. 7, pp. 3219–3227, Jul. 2021.

[34] J.-C. Jeong, I.-B. Yom, J.-D. Kim, W.-Y. Lee, and C.-H. Lee, "A 6–18-GHz GaAs multifunction chip with 8-bit true time delay and 7-bit amplitude control," *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 5, pp. 2220–2230, May 2018.
[35] N. Llombart, K. B. Cooper, R. J. Dengler, T. Bryllert, G. Chattopadhyay, and P. H. Siegel, "Time-delay multiplexing of two beams in a terahertz imaging radar," *IEEE Trans. Microw. Theory Techn.*, vol. 58, no. 7, pp. 1999–2007, Jul. 2010.
[36] L. Yan, C. Han, and J. Yuan, "Energy-efficient dynamic-subarray with fixed true-time-delay design for terahertz wideband hybrid beamforming," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 10, pp. 2840–2854, Oct. 2022.
[37] Z. Wu and L. Dai, "The manifestation of spatial wideband effect in circular array: From beam split to beam defocus," *IEEE Trans. Commun.*, 2024.
[38] D. Zhu, J. Choi, and R. W. Heath, "Auxiliary beam pair enabled aod and aoa estimation in closed-loop large-scale millimeter-wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4770–4785, Jul. 2017.
[39] D. Zhu, J. Choi, Q. Cheng, W. Xiao, and R. W. Heath, "High-resolution angle tracking for mobile wideband millimeter-wave systems with antenna array calibration," *IEEE Trans. Wireless Commun.*, vol. 17, no. 11, pp. 7173–7189, Nov. 2018.
[40] J. Tan and L. Dai, "Wideband beam tracking in THz massive MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 6, pp. 1693–1710, Jun. 2021.
[41] S. Kim, J. Wu, J. Moon, and B. Shim, "Frequency-dependent precoding for wideband terahertz communication systems," in *Proc. 2023 IEEE Global Commun. Conf. (IEEE GLOBECOM'23)*, 2023, pp. 5787–5792.

Tianyue Zheng received the B.E. degree in information engineering from Southeast University, Nanjing, China, in 2022. She is currently pursuing the Ph.D. degree in the Department of Electronic Engineering, Tsinghua University, Beijing, China. Her research interests include extremely large scale MIMO (XL-MIMO), CSI acquisition and AI for communications. She has received the National Scholarship in 2019 and the Excellent Student of Jiangsu Province in 2021.

Mingyao Cui received the B.E. and M.S. degrees in the Department of electronic engineering from Tsinghua University, Beijing, China, in 2020 and 2023, respectively. He is currently pursuing the Ph.D. degree in the University of HongKong. His research interests include Rydberg atomic receiver, massive MIMO, and near-field communications. He received the IEEE ICC Outstanding Demo Award and the National Scholarship in 2022, and the HKPF Scholarship in 2023.

Zidong Wu received the B.E. and Ph.D. degree in Department of Electronic Engineering from Tsinghua University, Beijing, China, in 2019 and 2024, respectively. His research interests include massive MIMO, mmWave communications, and near-field communications. He has received the Honorary Mention of IEEE ComSoc Student Competition in 2019, IEEE ICC Oustanding Demo Award in 2022, and IEEE Best Paper Award at IEEE GLOBECOM 2023.

Linglong Dai (Fellow, IEEE) received the B.S. degree from Zhejiang University, Hangzhou, China, in 2003, the M.S. degree from the China Academy of Telecommunications Technology, Beijing, China, in 2006, and the Ph.D. degree from Tsinghua University, Beijing, in 2011. From 2011 to 2013, he was a Post-Doctoral Researcher with the Department of Electronic Engineering, Tsinghua University, where he was an Assistant Professor from 2013 to 2016, an Associate Professor from 2016 to 2022, and has been a Professor since 2022. His current research interests include massive MIMO, reconfigurable intelligent surface (RIS), millimeter-wave and Terahertz communications, near-field communications, machine learning for wireless communications, and electromagnetic information theory.

He has coauthored the book *MmWave Massive MIMO: A Paradigm for 5G* (Academic Press, 2016). He has authored or coauthored over 100 IEEE journal papers and over 60 IEEE conference papers. He also holds over 20 granted patents. He has received five IEEE Best Paper Awards at the IEEE ICC 2013, the IEEE ICC 2014, the IEEE ICC 2017, the IEEE VTC 2017-Fall, the IEEE ICC 2018, and the IEEE GLOBECOM 2023. He has also received the Tsinghua University Outstanding Ph.D. Graduate Award in 2011, the Beijing Excellent Doctoral Dissertation Award in 2012, the China National Excellent Doctoral Dissertation Nomination Award in 2013, the URSI Young Scientist Award in 2014, the IEEE Transactions on Broadcasting Best Paper Award in 2015, the Electronics Letters Best Paper Award in 2016, the National Natural Science Foundation of China for Outstanding Young Scholars in 2017, the IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award in 2017, the IEEE ComSoc Asia-Pacific Outstanding Paper Award in 2018, the China Communications Best Paper Award in 2019, the IEEE Access Best Multimedia Award in 2020, the IEEE Communications Society Leonard G. Abraham Prize in 2020, the IEEE ComSoc Stephen O. Rice Prize in 2022, the IEEE ICC Best Demo Award in 2022, and the National Science Foundation for Distinguished Young Scholars in 2023. He was listed as a Highly Cited Researcher by Clarivate Analytics from 2020 to 2023. He was elevated as an IEEE Fellow in 2022.