

Triple-Refined Hybrid-Field Beam Training for mmWave Extremely Large-Scale MIMO

Kangjian Chen¹, Graduate Student Member, IEEE, Chenhao Qi¹, Senior Member, IEEE, Octavia A. Dobre², Fellow, IEEE, and Geoffrey Ye Li³, Fellow, IEEE

Abstract—This paper investigates beam training for extremely large-scale multiple-input multiple-output systems. By considering both the near field and far field, a triple-refined hybrid-field beam training scheme is proposed, where high-accuracy estimates of channel parameters are obtained through three steps of progressive beam refinement. First, the hybrid-field beam gain (HFBG)-based first refinement method is developed. Based on the analysis of the HFBG, the first-refinement codebook is designed and the beam training is performed accordingly to narrow down the potential region of the channel path. Then, the maximum likelihood (ML)-based and principle of stationary phase (PSP)-based second refinement methods are developed. By exploiting the measurements of the beam training, the ML is used to estimate the channel parameters. To avoid the high computational complexity of ML, closed-form estimates of the channel parameters are derived according to the PSP. Moreover, the Gaussian approximation (GA)-based third refinement method is developed. The hybrid-field neighboring search is first performed to identify the potential region of the main lobe of the channel steering vector. Afterwards, by applying the GA, a least-squares estimator is developed to obtain the high-accuracy channel parameter estimation. Simulation results verify the effectiveness of the proposed scheme.

Index Terms—Beam training, extremely large-scale multiple-input multiple-output (XL-MIMO), Gaussian approximation, near field, principle of stationary phase.

I. INTRODUCTION

MILLIMETER wave (mmWave) that reserves wide spectrum resources is a promising technology for achieving

Manuscript received 26 July 2023; revised 7 November 2023; accepted 3 January 2024. Date of publication 17 January 2024; date of current version 14 August 2024. This work was supported in part by the National Natural Science Foundation of China under Grant 62071116 and Grant U22B2007, in part by the National Key Research and Development Program of China under Grant 2021YFB2900404, in part by the Natural Sciences and Engineering Research Council of Canada (NSERC) through the Discovery Program, and in part by the Postgraduate Research and Practice Innovation Program of Jiangsu Province under Grant KYCX23_0262. An earlier version of this paper was presented in part at the IEEE Global Communications Conference, Kuala Lumpur, Malaysia, December 2023 [1]. The associate editor coordinating the review of this article and approving it for publication was Z. Zhang. (Corresponding author: Chenhao Qi.)

Kangjian Chen and Chenhao Qi are with the School of Information Science and Engineering, Southeast University, Nanjing 210096, China (e-mail: qch@seu.edu.cn).

Octavia A. Dobre is with the Faculty of Engineering and Applied Science, Memorial University, St. John's, NL A1C 5S7, Canada (e-mail: odobre@mun.ca).

Geoffrey Ye Li is with the Department of Electrical and Electronic Engineering, Imperial College London, SW7 2AZ London, U.K. (e-mail: geoffrey.li@imperial.ac.uk).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TWC.2024.3351712>.

Digital Object Identifier 10.1109/TWC.2024.3351712

high data rates. Thanks to its small wavelength, the space-limited base station (BS) can accommodate a large number of antennas to enhance the spectral efficiency by implementing the massive multiple-input multiple-output (MIMO) [2], [3], [4]. The perfect match between the abundant spectrum resources provided by mmWave and the high spectral efficiency enabled by massive MIMO has led to the boom of mmWave massive MIMO [5].

Recently, extremely large-scale MIMO (XL-MIMO) with far more antennas than existing massive MIMO has been developed to further improve the spectral efficiency via ultrahigh-gain beamforming [6], [7], [8]. Attracted by the tremendous potentials, researchers move on to develop XL-MIMO based on the well-explored massive MIMO. However, extending the works in massive MIMO to XL-MIMO is not straightforward. Due to the much larger array aperture, the channel model of the latter differs greatly from that of the former. Generally, the radiation fields of the electromagnetic waves can be divided into the near field and the far field according to the distance between the BS and the radiation source [9], [10]. In the far field, the channel characteristics conform to the planar-wave model, and the phase differences among antennas can be reasonably approximated as a linear function of the antenna indices [7]. In the near field, the channel characteristics cannot be accurately modeled using the planar-wave model. Instead, they must be characterized by the spherical-wave model, where the phase differences among antennas are expressed as a nonlinear function of the antenna indices [7]. For conventional massive MIMO, where the array aperture is relatively small, the BS coverage predominantly falls within the far field. However, in the emerging XL-MIMO, the substantial expansion of the array aperture results in a significant portion of the BS coverage lying within the near field. In this context, XL-MIMO communications will operate within the coexistence of the far and near fields. In addition, the difference in channel models renders conventional far-field techniques unsuitable for the near field. Therefore, in the realm of XL-MIMO communications, the focus should be shifted towards techniques that can effectively adapt to the hybrid field, including the near and far fields.

One pivotal issue in wireless communications is the channel state information (CSI) acquisition [11], [12], [13]. Due to the large propagation attenuation of mmWave, beam training that can achieve high beamforming gain is preferred [14]. Generally, beam training can be categorized into beam sweeping and multi-stage beam training. Although beam sweeping can

effectively combat noise by using narrow beams, its success is founded on the intensive beam training. In contrast, multi-stage beam training that explores and refines the CSI stage by stage can compare favorably with the beam sweeping while requiring much lower training overhead [15]. A special case of the multi-stage beam training is the hierarchical beam training (HBT) [16], [17], [18], [19], [20]. By comparing the received powers of the codewords in the predefined hierarchical codebook layer by layer, the HBT can gradually narrow down the candidate sets of channel paths, leading to the increasing beamforming gain. Different from the HBT that depends on the hierarchical codebook and mainly exploits the power information of the received signals, the general multi-stage beam training flexibly divides the whole training process into several stages and extracts CSI from the former stage via signal processing techniques to assist the beam training of the latter stage [21], [22], [23], [24]. In [21], an optimized two-stage search algorithm is developed, where the second stage only trains the beam candidates derived from the first stage to accomplish beam alignment. In [22], an adaptive and sequential beam alignment algorithm is proposed, where the codeword for the beam training in the lower layer is determined by the posterior of the channel angle derived in the upper layer. In [23], based on the beam sweeping results, an auxiliary beam pair method is proposed to obtain the high-accuracy estimates of channel angles by comparing the powers of received signals. In [24], to further refine the result of beam sweeping, the far-field array gain is approximated as a Gaussian function and the channel angles are estimated based on additional channel tests.

The fundamental changes in the radiation fields for XL-MIMO communication invalidate the conventional far-field beam training methods in [13], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], and [24]. In addition, the coexistence of the near and far fields appeals for novel hybrid-field beam training methods for XL-MIMO communications. One straightforward method for hybrid field beam training is the hybrid-field beam sweeping (HFBS) [25], which explores the far field via beam scanning and the near field via beam focusing. Nevertheless, the HFBS exhaustively tests every codeword in the predefined hybrid-field codebook and needs an extensive amount of beam training. To reduce the overhead of HFBS, a two-phase beam training (TPBT) method is proposed in [26], where the first phase determines the candidate channel angles via far-field beam sweeping while the second phase finds the channel distance based on shortlisted candidate angles in the first phase. However, the overhead of the TPBT is still high due to the employment of the far-field beam sweeping in the first phase. To further reduce the overhead of TPBT, a hierarchical codebook is designed and a chirp-based HBT (CHBT) method is proposed for the XL-MIMO in [27]. It is worth noting that the imperfection in the beam patterns of codewords in the hierarchical codebook could lead to degradation in beam training performance. For the distance-based HBT (DHBT) method in [28], the codebook is designed by equally sampling the space in distance. However, the neglect of the polar-domain sparsity leads to a deterioration in the training

performance of DHBT. For the two-stage HBT (TSHBT) in [29], the training procedure is divided into two stages, where the first stage only refines the channel angle with wide beams while the second stage simultaneously refines the channel angle and distance with beam focusing. However, a large number of antennas need to be deactivated to form wide beams in the first stage. Owing to the powerful feature extraction ability of neural networks, deep learning has also been exploited for hybrid-field beam training [30], [31], [32]. However, the substantial computational overhead and resource-intensive characteristics of deep learning present significant barriers to its widespread implementation in XL-MIMO systems. Furthermore, the beam training methods in [25], [26], [27], [28], [29], [30], and [31] construct codebooks based on the quantized samples, which not only deteriorates the beamforming gain but also increases errors in near-field localization. While numerous studies have explored high-accuracy near-field sensing and localization [33], [34], [35], they often rely on complicated hardware configurations or necessitate extensive computational resources, which is incompatible with the concise beam training framework. One simple method to improve the accuracy of hybrid-field beam training is the beam refinement [36]. However, the beam refinement method in [36] is specially tailored for the partially-connected hybrid beamforming structure. In the context of XL-MIMO, beam refinement methods for the fully-connected hybrid beamforming structure are missing in the existing works.

In this paper, we investigate hybrid-field beam training for XL-MIMO systems. We aim to achieve high-accuracy beam alignment for XL-MIMO systems with low overhead and low computational complexity, which is beyond the scopes of the existing works in [25], [26], [27], [28], [29], [30], [31], [32], [33], [34], [35], and [36]. Inspired by the idea of the multi-stage beam training, we propose to acquire the CSI of XL-MIMO systems in a progressive refinement way. To obtain the high-accuracy CSI with low overhead, efficient refinement methods and effective progressive strategies are devised. The main contributions of this paper are summarized as follows.

- By considering both the near field and far field, we develop a triple-refined hybrid-field beam training (THBT) scheme, where high-accuracy estimates of channel parameters are obtained through three steps of progressive refinement. Benefiting from the devised refinement methods and progressive strategies, the proposed THBT can accomplish the beam alignment for XL-MIMO systems with substantially reduced training overhead.
- We develop the hybrid-field beam gain (HFBG)-based first refinement method. By analyzing the HFBG of the channel steering vector for the XL-MIMO systems, we design the first-refinement codebook that satisfies the presented two design criteria and perform the beam training accordingly to determine the potential region of the channel path.
- We then exploit the maximum likelihood (ML) and principle of stationary phase (PSP) to develop the second refinement methods. Based on the property of the

HFBG, we first design the second-refinement codebook, where the beam coverage of each codeword contains the potential region from the first refinement. Then, by exploiting the measurements of the beam training, the ML is used to estimate the channel parameters. To avoid the high computational complexity of ML, the phase of the hybrid-field beam gain is analytically expressed according to the PSP, and closed-form estimates of the channel parameters are derived.

- We develop the Gaussian approximation (GA)-based third refinement method. First, we perform the hybrid-field neighboring search to identify the potential region of the main lobe of the channel steering vector. Inspired by the similarities between the Taylor series of the Gaussian function and that of the HFBG, we approximate the main lobe of the HFBG as a two-dimensional Gaussian function. Then, by sampling the surrogate distance and angle uniformly within the potential region, we design the third-refinement codebook and perform the beam training accordingly. Based on the measurements of the beam training, a least-squares (LS) estimator is developed to obtain the high-accuracy channel parameter estimation.

The rest of this paper is organized as follows. The model of the XL-MIMO system is introduced in Section II. The codebook design and beam training for the first, second, and third refinement are presented in Sections III, IV, and V, respectively. The proposed methods are evaluated in Section VI, and the paper is concluded in Section VII.

The notations are defined as follows. Symbols for matrices (upper case) and vectors (lower case) are in boldface. $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^H$ denote the conjugate, transpose, and conjugate transpose (Hermitian), respectively. $[a]_n$, $[A]_{m,\cdot}$, $[A]_{:,n}$, and $[A]_{m,n}$ denote the n th entry of vector \mathbf{a} , the m th row of matrix \mathbf{A} , the n th column of matrix \mathbf{A} , and the entry on the m th row and the n th column of matrix \mathbf{A} , respectively. $\{\mathcal{C}\}_n$ is the n th element of the set \mathcal{C} . In addition, j , $|\cdot|$, \mathbb{C} , and \mathcal{CN} denote the square root of -1 , the absolute value of a scalar, the set of the complex number and the complex Gaussian distribution. Moreover, $f'(\cdot)$ and $f''(\cdot)$ represent the first-order and the second-order derivative of $f(\cdot)$, respectively. $\ln(\cdot)$ denotes the natural logarithm of a number.

II. SYSTEM MODEL

As shown in Fig. 1, we consider the downlink beam training between the BS and the user, where a half-wavelength-interval array with N_t antennas is equipped at the BS while a single-antenna transceiver is adopted at the user. To ease the notation, we assume that N_t is an odd number, and $N_t = 2N + 1$. With the hybrid beamforming structure, the N_t antennas are connected to N_{RF} radio frequency (RF) chains through a phase shifter network. In this work, the beam training is performed based on the analog beamforming, and hence, we only focus on one of the multiple RF chains for simplicity. Then the s th received signal, for $s = 1, 2, \dots, S$, can be expressed as

$$y_s = \mathbf{h}^H \mathbf{f}_s x_s + \eta \quad (1)$$

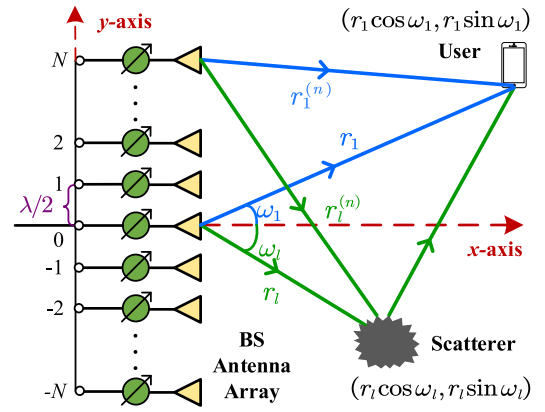


Fig. 1. Illustration of the system model.

where $\mathbf{h} \in \mathbb{C}^{N_t}$, $\mathbf{f}_s \in \mathbb{C}^{N_t}$, and x_s denote the channel between the BS and the user, the analog beamformer at the BS, and the transmit signal, respectively. In addition, $\eta \sim \mathcal{CN}(0, \sigma^2)$ denotes the additive white Gaussian noise.

To depict the channel, we first establish a Cartesian coordinate, which sets the center, the normal direction, and the tangent direction of the antenna array as the origin, the x-axis, and the y-axis, respectively. Then the coordinate of the n th antenna can be expressed as $(0, n\lambda/2)$ for $n \in \mathcal{I}$, where $\mathcal{I} \triangleq \{-N, \dots, 0, \dots, N\}$ and λ denotes the wavelength. From Fig. 1, the coordinate of the radiation source at the l th path is $(r_l \cos \omega_l, r_l \sin \omega_l)$, for $l = 1, 2, \dots, L$, where L denotes the number of paths, r_l is the distance between the origin and the l th radiation source, and $\omega_l \in [-\pi/2, \pi/2]$ represents the angle of the l th radiation source relative to the x-axis. The distance between the l th radiation source and the n th antenna is calculated as

$$r_l^{(n)} = \sqrt{r_l^2 + n^2 \lambda^2 / 4 - nr_l \Omega_l \lambda} \quad (2)$$

where $\Omega_l \triangleq \sin \omega_l \in [-1, 1]$. Then, the channel steering vector between the BS and the user can be expressed as

$$\mathbf{h} = \sum_{l=1}^L g_l \boldsymbol{\alpha}(\Omega_l, r_l) \quad (3)$$

where g_l denotes the channel gain of the l th path and the channel steering vector $\boldsymbol{\alpha}(\cdot)$ is defined as

$$\boldsymbol{\alpha}(\Omega_l, r_l) = \frac{1}{\sqrt{N_t}} \left[e^{-j \frac{2\pi}{\lambda} (r_l^{(-N)} - r_l)}, \dots, e^{-j \frac{2\pi}{\lambda} (r_l^{(N)} - r_l)} \right]^T \quad (4)$$

Note that the channel steering vector in (4) adapts to both the near and far fields. To ease the notation, we omit the subscript “ l ” and focus on processing the path of interest. Then, $r^{(n)}$ can be simplified as

$$r^{(n)} \approx r - n\Omega\lambda/2 + \frac{n^2\lambda^2(1 - \Omega^2)}{8r} \quad (5)$$

according to $\sqrt{1 + \epsilon} \approx 1 + \frac{1}{2}\epsilon - \frac{1}{8}\epsilon^2$, which is verified to be accurate if $r^{(n)} \geq 0.5\sqrt{N^3\lambda^2}$ [7], [9]. In fact, $0.5\sqrt{N^3\lambda^2}$ is quite small compared to the coverage of the BS. For example, if $N = 128$ and $\lambda = 0.005$ m, we have $0.5\sqrt{N^3\lambda^2} \approx 3.6$ m,

which is much smaller than the typical coverage of the BS. Therefore, in this work, we focus on the radiation field with $r^{(n)} \geq 0.5\sqrt{N^3\lambda^2}$. Substituting (5) into (4), we have

$$[\alpha(\Omega, r)]_n \approx \frac{1}{\sqrt{N_t}} e^{j\pi\left(\Omega n - \frac{\lambda(1-\Omega^2)}{4r} n^2\right)} \quad (6)$$

for $n \in \mathcal{I}$. We define

$$b \triangleq \frac{\lambda(1-\Omega^2)}{4r} \quad (7)$$

which is referred to as the ‘‘surrogate distance’’ in this paper. Then (4) can be approximated as

$$\begin{aligned} \alpha(\Omega, r) &\approx \frac{1}{\sqrt{N_t}} \left[e^{j\pi(-\Omega N - bN^2)}, \dots, e^{j\pi(\Omega N - bN^2)} \right]^T \\ &\triangleq \gamma(\Omega, b). \end{aligned} \quad (8)$$

We then propose the THBT scheme to obtain high-accuracy channel parameter estimation through three steps of progressive refinement.

III. HFBG-BASED FIRST REFINEMENT METHOD

In this section, we first analyze the HFBG of the channel steering vector for XL-MIMO systems. Based on the analysis, we develop the HFBG-based first refinement method, where the first-refinement codebook is designed and the beam training is performed accordingly to determine the potential region of the channel path.

A. Hybrid-Field Beam Gain

For an arbitrary steering vector, $\mathbf{u} \triangleq \gamma(\Theta, k)$, we define its HFBG as

$$\begin{aligned} G(\mathbf{u}, \Omega, b) &\triangleq N_t \gamma(\Omega, b)^H \mathbf{u} \\ &= \sum_{n=-N}^N e^{j\pi((\Theta-\Omega)n + (b-k)n^2)} \\ &\stackrel{(a)}{\approx} \int_{-N-1/2}^{N+1/2} e^{j\pi((\Theta-\Omega)z + (b-k)z^2)} dz \\ &= \int_{-\infty}^{\infty} U(z) e^{j\pi((\Theta-\Omega)z + (b-k)z^2)} dz \end{aligned} \quad (9)$$

where

$$U(z) = \begin{cases} 1, & -N - 1/2 \leq z \leq N + 1/2 \\ 0, & \text{others.} \end{cases} \quad (10)$$

In (9), we approximate the summation as the integral in (a). Due to the quadratic phase structure of the integrand, it is hard to obtain the closed-form solution of (9). Alternatively, a simple but effective approximation based on the PSP is widely adopted [37], [38], [39]. We define

$$J(z, \Omega, b) \triangleq \pi((\Theta - \Omega)z + (b - k)z^2) \quad (11)$$

and denote z_0 as the stationary phase. According to the PSP, the differential of $J(z, \Omega, b)$ at the stationary phase satisfies the condition that $J'(z_0, \Omega, b) = 0$. Therefore, we can obtain

$$z_0 = \frac{\Omega - \Theta}{2(b - k)}. \quad (12)$$

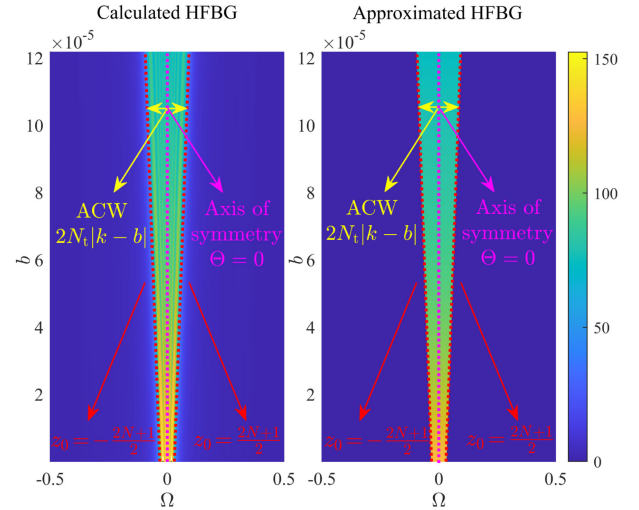


Fig. 2. Comparison of the calculated and the approximated HFBGs.

Then, based on the PSP, (9) can be approximated as

$$\begin{aligned} G(\mathbf{u}, \Omega, b) &\approx \sqrt{\frac{-2\pi}{J''(z_0, \Omega, b)}} e^{-j\pi/4} U(z_0) e^{jJ(z_0, \Omega, b)} \\ &= \frac{e^{-j\pi/4}}{\sqrt{k-b}} U(z_0) e^{j\pi \frac{(\Omega-\Theta)^2}{4(k-b)}} \\ &= \begin{cases} \frac{e^{-j\pi/4}}{\sqrt{k-b}} e^{j\pi \frac{(\Omega-\Theta)^2}{4(k-b)}}, & -\frac{N_t}{2} \leq z_0 \leq \frac{N_t}{2} \\ 0, & \text{others.} \end{cases} \end{aligned} \quad (13)$$

By taking the absolute value of $G(\mathbf{u}, \Omega, b)$, we have

$$|G(\mathbf{u}, \Omega, b)| \approx \begin{cases} \frac{1}{\sqrt{|b-k|}}, & -\frac{N_t}{2} \leq z_0 \leq \frac{N_t}{2} \\ 0, & \text{others.} \end{cases} \quad (14)$$

Combining (12) and (14), we can express the beam coverage of \mathbf{u} as

$$\mathcal{B}(\mathbf{u}) = \left\{ (\Omega, b) \mid \left| \frac{\Omega - \Theta}{b - k} \right| \leq N_t \right\}. \quad (15)$$

From (15), the angle coverage of \mathbf{u} for a fixed b is

$$\Theta - N_t |b - k| \leq \Omega \leq \Theta + N_t |b - k| \quad (16)$$

which indicates that the beam coverage is symmetric about angle Θ and the angle coverage width (ACW) for a fixed b is $2N_t |b - k|$.

In Fig. 2, we illustrate the calculated and the approximated HFBGs of \mathbf{u} , where we set $N_t = 513$, $N = 256$, $\Theta = 0$ and $k = -6.09 \times 10^{-5}$. From the figure, the calculated HFBG and the approximated HFBG share similarities in several aspects, including the boundary, the ACW and the axis of the symmetry, which indicates that the PSP can provide a good approximation for the HFBG.

B. First-Refinement Codebook Design and Beam Training

Suppose the initial potential region of the interested channel path is

$$\Phi_1 = \left\{ (\Omega, b) \mid -\bar{\Omega} \leq \Omega \leq \bar{\Omega}, 0 \leq b \leq \bar{b} \right\} \quad (17)$$

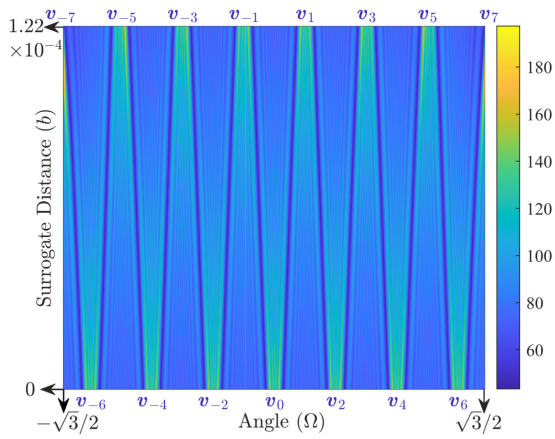


Fig. 3. Illustration of the beam gain of \mathbf{v}_m for $m = -M_1, \dots, 0, \dots, M_1$.

where $\bar{\Omega}$ and \bar{b} denote boundaries of Ω and b , respectively. Note that in (17) we assume the angle of the potential region is symmetric about zero for simplicity. However, more complicated cases can also be developed by exploiting the phase-shift invariance property of the channel steering vector [27]. In addition, b is lower bounded by zero due to its definition in (7).

Generally, the beam training focuses on narrowing down the potential region by testing the codewords in the predefined codebook. Then, directional beams with high beamforming gains can be formed to combat the noise. To narrow down the potential region effectively and efficiently, the codeword design commonly considers the following two criteria [40]: 1) The beam coverages of the codewords do not overlap with each other. 2) The potential region should be fully covered by the union of the beam coverage of the codewords. In the following, by considering the aforementioned two criteria, the first-refinement codebook is designed accordingly.

Denote the first-refinement codebook as

$$\mathcal{C}_1 = \{\mathbf{v}_{-M_1}, \dots, \mathbf{v}_0, \dots, \mathbf{v}_{M_1}\} \quad (18)$$

where $2M_1 + 1$ is the number of codewords. We design the m th codeword in \mathcal{C}_1 as

$$\mathbf{v}_m = \gamma(m\bar{\Theta}_1, \bar{k}_1) \quad (19)$$

for $m = -M_1, -M_1 + 1, \dots, M_1$, where

$$\bar{k}_1 = \begin{cases} \tilde{k}_1, & m \text{ is even,} \\ \bar{b} - \tilde{k}_1, & m \text{ is odd,} \end{cases} \quad (20)$$

$\tilde{k}_1 < 0$ is the introduced parameter to control the beam coverage of the designed codewords, and $\bar{\Theta}_1 \triangleq (\bar{b} - 2\tilde{k}_1)N_t$ denotes the angle deviation between adjacent codewords. By substituting (19) and (20) into (15), it is convenient to verify that the beam coverages of the designed codewords fit closely without overlapping, which indicates that the first criterion is satisfied. In addition, we set only two values for \bar{k}_1 in (20) so that the designed codewords have similar beam coverages and can provide comparable beamforming gains for every position in the space. To satisfy the second criterion, the potential region in (17) should be fully covered by the

Algorithm 1 HFBG-Based First Refinement Methods

- 1: **Input:** N , N_t , λ , \mathbf{h} , $\bar{\Omega}$, \bar{b} , and \tilde{k}_1
- 2: Obtain M_1 via (22).
- 3: Obtain \mathcal{C}_1 via (18).
- 4: Obtain y_m for $m = -M_1, \dots, 0, \dots, M_1$ via (23).
- 5: Obtain \bar{m} via (24).
- 6: Obtain $\tilde{\Phi}_1$ via (25).
- 7: **Output:** \bar{m} and $\tilde{\Phi}_1$.

$2M_1 + 1$ codewords in (18). Since both the potential region and the beam coverage of the designed codewords are symmetric, we only focus on the region with $\Omega \geq 0$. From (16), the right angle boundary of \mathbf{v}_{M_1} is $M_1\bar{\Theta}_1 + N_t|b - \bar{k}_1|$, where $0 \leq b \leq \bar{b}$. Then according to (17) and the second criterion, we have

$$\min_{0 \leq b \leq \bar{b}} M_1\bar{\Theta}_1 + N_t|b - \bar{k}_1| > \bar{\Omega}. \quad (21)$$

From (21), we have

$$M_1 > (\bar{\Omega} + N_t\tilde{k}_1)/\bar{\Theta}_1 \quad (22)$$

which provides guidance for the design of M_1 .

Then, we perform the beam training based on \mathcal{C}_1 to narrow down the potential region. When testing the m th codeword in \mathcal{C}_1 , the signal received by the user can be expressed as

$$y_m = \mathbf{h}^H \mathbf{v}_m + \eta \quad (23)$$

for $m = -M_1, \dots, 0, \dots, M_1$, where we set the transmit signal as “1”. The result of the beam training can be expressed as

$$\bar{m} = \arg \max_{m=-M_1, \dots, 0, \dots, M_1} |y_m|. \quad (24)$$

From (24), the channel path locates in the beam coverage of $\mathbf{v}_{\bar{m}}$. Therefore, we can update the potential region in (17) as

$$\tilde{\Phi}_1 = \left\{ (\Omega, b) \left| 0 \leq b \leq \bar{b}, \frac{|\Omega - \bar{m}\bar{\Theta}_1|}{|b - \bar{k}_1|} \leq N_t \right. \right\}, \quad (25)$$

where \bar{k}_1 can be determined by substituting \bar{m} into (20).

Now, we provide an example for the first-refinement codebook in **Design Example 1**.

Design Example 1: Consider an XL-MIMO system with $N = 256$, $N_t = 513$, $\lambda = 0.005$ m, $\tilde{k}_1 = -6.09 \times 10^{-5}$, $\bar{b} = 1.22 \times 10^{-4}$ and $\bar{\Omega} = \sqrt{3}/2$. From (22), we can obtain $M_1 \geq 6.68$. Since M_1 must be an integer, we set $M_1 = 7$, leading to totally 15 codewords in \mathcal{C}_1 . According to (19), \mathbf{v}_m for $m = -M_1, \dots, 0, \dots, M_1$ can be designed. In Fig. 3, we illustrate the calculated HFBG of \mathbf{v}_m within the potential region $\tilde{\Phi}_1$. From the figure, the potential region can be fully covered by the codewords in \mathcal{C}_1 and there is little overlap between the beam coverages of two different codewords in \mathcal{C}_1 , which indicates that the two aforementioned design criteria are well satisfied.

Finally, we summarize the HFBG-based first refinement method in **Algorithm 1**.

IV. ML-BASED AND PSP-BASED SECOND REFINEMENT METHODS

In this section, by exploiting the ML and PSP, we develop the second refinement methods to further improve the accuracy of channel parameter estimation.

Note that the beam training is more likely to fail when channel paths locate in the transition zone of the codeword due to the similar beam gains of adjacent codewords [41]. Therefore, we extend the potential region in (25) as

$$\Phi_2 = \left\{ (\Omega, b) \mid 0 \leq b \leq \bar{b}, \frac{|\Omega - \bar{m}\bar{\Theta}_1|}{|b - \bar{k}_1| + 1/N_t^2} \leq N_t \right\} \quad (26)$$

to include the transition zone, where the width of the transition zone is typically set as $1/N_t$ [42]. Then we perform the beam training within Φ_2 .

We denote the second-refinement codebook as

$$\mathcal{C}_2 = \{\mathbf{w}_{-M_2}, \dots, \mathbf{w}_0, \dots, \mathbf{w}_{M_2}\} \quad (27)$$

where $2M_2 + 1$ is the number of codewords. We design the m th codeword in (27) as

$$\mathbf{w}_m = \gamma(\tilde{\Theta}_m, \bar{k}_2), \quad \tilde{\Theta}_m = \bar{m}\bar{\Theta}_1 + m\bar{\Theta}_2 \quad (28)$$

for $m = -M_2, \dots, 0, \dots, M_2$, where \bar{k}_2 and $\bar{\Theta}_2 > 0$ are introduced parameters to control the beam coverage and the angle deviation of the codewords, respectively. According to (15), the beam coverage of the m th codeword can be expressed as

$$\Psi_m = \left\{ (\Omega, b) \mid 0 \leq b \leq \bar{b}, \frac{|\Omega - \tilde{\Theta}_m|}{|b - \bar{k}_2|} \leq N_t \right\}. \quad (29)$$

To provide high beam gain during beam training, the beam coverage of each codeword in \mathcal{C}_2 should contain Φ_2 . Based on (26) and (29), \bar{k}_2 should satisfy

$$\frac{1}{N_t} + N_t|b - \bar{k}_1| \leq N_t|b - \bar{k}_2| + m\bar{\Theta}_2. \quad (30)$$

From (30), we have

$$\begin{cases} \bar{k}_2 \leq \bar{k}_1 - M_2\bar{\Theta}_2/N_t - 1/N_t^2, \bar{m} \text{ is even} \\ \bar{k}_2 \geq \bar{k}_1 + M_2\bar{\Theta}_2/N_t + 1/N_t^2, \bar{m} \text{ is odd.} \end{cases} \quad (31)$$

Then, we perform the beam training based on \mathcal{C}_2 . When testing the m th codeword in \mathcal{C}_2 , the signal received by the user can be expressed as

$$\begin{aligned} \bar{y}_m &= \mathbf{h}^H \mathbf{w}_m + \eta \\ &\stackrel{(a)}{\approx} \mathbf{h}^H \gamma(\tilde{\Theta}_m, \bar{k}_2) \\ &\stackrel{(b)}{\approx} g^* \gamma(\Omega, b)^H \gamma(\tilde{\Theta}_m, \bar{k}_2) \end{aligned} \quad (32)$$

where we omit the noise term in (a) and the effects of the non-line-of-sight (NLoS) paths in (b).

1) *ML-Based Second Refinement Method*: Based on the measurements in (32), the ML estimation of channel parameters can be expressed as

$$(\hat{\Omega}_2, \hat{b}_2, \hat{g}) = \arg \min_{\Omega, b, g} \sum_{m=-M_2}^{M_2} |\bar{y}_m^* - g\gamma(\tilde{\Theta}_m, \bar{k}_2)^H \gamma(\Omega, b)|^2 \quad (33)$$

We express (33) in a vector form as

$$(\hat{\Omega}_2, \hat{b}_2, \hat{g}) = \arg \min_{\Omega, b, g} \|\bar{\mathbf{y}} - g\mathbf{\Gamma}^H \gamma(\Omega, b)\|_2^2 \quad (34)$$

where $\bar{\mathbf{y}}$ and $\mathbf{\Gamma}$ are the stack of \bar{y}_m^* and $\gamma(\tilde{\Theta}_m, \bar{k}_2)$, respectively. Given $\gamma(\Omega, b)$, the optimal solution of g is $\hat{g} = \gamma(\Omega, b)^H \mathbf{\Gamma} \bar{\mathbf{y}} / \|\mathbf{\Gamma}^H \gamma(\Omega, b)\|_2^2$. Then, (34) is converted to

$$(\hat{\Omega}_2, \hat{b}_2) = \arg \max_{\Omega, b} \|\bar{\mathbf{y}}^H \mathbf{\Gamma}^H \gamma(\Omega, b)\|_2 / \|\mathbf{\Gamma}^H \gamma(\Omega, b)\|_2 \quad (35)$$

which can be solved by the two-dimensional search. We term the THBT with this method as the THBT-ML.

2) *PSP-Based Second Refinement Method*: From (31), the channel path locates in the beam coverage of $\gamma(\tilde{\Theta}_m, \bar{k}_2)$, which indicates that $|\gamma(\Omega, b)^H \gamma(\tilde{\Theta}_m, \bar{k}_2)|$ has similar values for $m = -M_2, \dots, 0, \dots, M_2$. Therefore, we may normalize amplitudes of the measurements to simplify the analysis. Then we convert (33) to

$$(\hat{\Omega}_2, \hat{b}_2, \hat{g}) = \arg \min_{\Omega, b, g} \sum_{m=-M_2}^{M_2} (\mathcal{P}(\bar{y}_m^*) - \psi_m)^2 \quad (36)$$

where $\psi_m \triangleq \mathcal{P}(g\gamma(\tilde{\Theta}_m, \bar{k}_2)^H \gamma(\Omega, b))$ and $\mathcal{P}(\cdot)$ denotes the phase of a complex number. One method to solve (36) is performing the multi-dimensional search similar to (35). However, a vast number of calculations are needed especially for intensive samples of Ω and b . To avoid the high computational complexity, we revisit the PSP in (13). By extracting the phase of $G(\mathbf{u}, \Omega, b)$, we have

$$\mathcal{U}(G(\mathbf{u}, \Omega, b)) \approx \begin{cases} \frac{(\Omega - \Theta)^2}{4(k-b)^2} \pi - \frac{\pi}{4}, & -\frac{N_t}{2} \leq z_0 \leq \frac{N_t}{2} \\ 0, & \text{others} \end{cases} \quad (37)$$

where $\mathcal{U}(\cdot)$ denotes the unwrapped phase of a complex number. By applying the PSP approximation, we can convert (36) to

$$(\hat{\Omega}_2, \hat{b}_2, \hat{\phi}_2) = \arg \min_{\Omega, b, \phi} \sum_{m=-M_2}^{M_2} \left(\mathcal{U}(\bar{y}_m^*) - \pi \frac{(\Omega - \tilde{\Theta}_m)^2}{4(b - \bar{k}_2)} - \phi \right)^2 \quad (38)$$

where ϕ is a parameter introduced to eliminate the constant phases. We define

$$\mathcal{L}(\alpha, \beta, \gamma) \triangleq \sum_{m=-M_2}^{M_2} (\mathcal{U}(\bar{y}_m^*) + \alpha \tilde{\Theta}_m^2 + \beta \tilde{\Theta}_m + \gamma)^2 \quad (39)$$

where $\alpha \triangleq -\frac{\pi}{4(b - \bar{k}_2)}$, $\beta \triangleq \frac{\pi\Omega}{2(b - \bar{k}_2)}$ and $\gamma \triangleq -\frac{\pi\Omega^2}{4(b - \bar{k}_2)} - \phi$. The optimal solution of (38) is achieved when

$$\frac{\partial \mathcal{L}(\alpha, \beta, \gamma)}{\partial (2\alpha)} = \sum_{m=-M_2}^{M_2} (\mathcal{U}(\bar{y}_m^*) + \alpha \tilde{\Theta}_m^2 + \beta \tilde{\Theta}_m + \gamma) \tilde{\Theta}_m^2 = 0$$

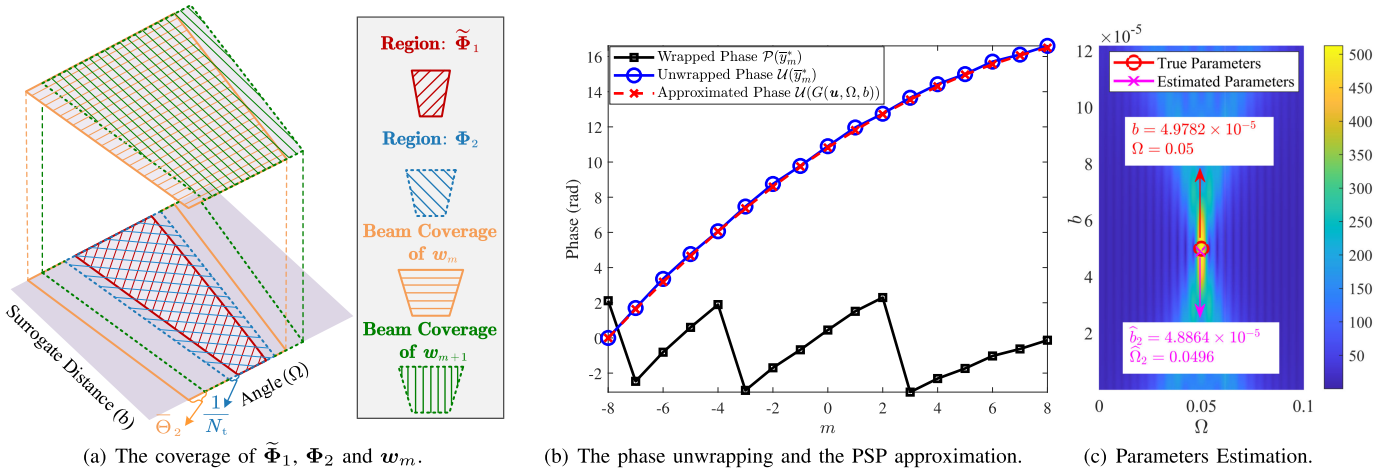


Fig. 4. Illustration of the second-refinement codebook and beam training.

$$\begin{aligned} \frac{\partial \mathcal{L}(\alpha, \beta, \gamma)}{\partial (2\beta)} &= \sum_{m=-M_2}^{M_2} (\mathcal{U}(\bar{y}_m^*) + \alpha \tilde{\Theta}_m^2 + \beta \tilde{\Theta}_m + \gamma) \tilde{\Theta}_m = 0 \\ \frac{\partial \mathcal{L}(\alpha, \beta, \gamma)}{\partial (2\gamma)} &= \sum_{m=-M_2}^{M_2} (\mathcal{U}(\bar{y}_m^*) + \alpha \tilde{\Theta}_m^2 + \beta \tilde{\Theta}_m + \gamma) = 0. \end{aligned} \quad (40)$$

Note that (40) is a system of linear equations and the closed-form solutions can be obtained via the Gaussian elimination. We omit the details and denote its solutions as $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\gamma}$. Then we can express the results of the second refinement as

$$\hat{b}_2 = -\frac{\pi}{4\hat{\alpha}} + \bar{k}_2, \text{ and } \hat{\Omega}_2 = -\frac{\hat{\beta}}{2\hat{\alpha}}. \quad (41)$$

We term the THBT with this method as the THBT-PSP.

Remark 1: One remaining problem is how to obtain the unwrapped phase of \bar{y}_m^* in (38). The classic method unwraps the phase by comparing the phase difference between the adjacent sample points with the threshold π [43]. To guarantee the successful phase unwrap, we have

$$\frac{1}{\pi} |\mathcal{U}(\bar{y}_{m+1}^*) - \mathcal{U}(\bar{y}_m^*)| \leq 1. \quad (42)$$

Note that

$$\begin{aligned} & \frac{1}{\pi} |\mathcal{U}(\bar{y}_{m+1}^*) - \mathcal{U}(\bar{y}_m^*)| \\ & \stackrel{(a)}{\approx} \left| \frac{(\Omega - \tilde{\Theta}_{m+1})^2}{4(b - \bar{k}_2)} - \frac{(\Omega - \tilde{\Theta}_m)^2}{4(b - \bar{k}_2)} \right| \\ & = \frac{\bar{\Theta}_2 |2\Omega - \tilde{\Theta}_{m+1} - \tilde{\Theta}_m|}{4|b - \bar{k}_2|} \\ & \stackrel{(b)}{\leq} \frac{\bar{\Theta}_2 (1/N_t + |b - \bar{k}_1| N_t + M_2 \bar{\Theta}_2)}{2|b - \bar{k}_2|} \\ & \stackrel{(c)}{\leq} \frac{\bar{\Theta}_2 (1/N_t + (\bar{b} - \tilde{k}_1) N_t + M_2 \bar{\Theta}_2)}{2|b - \bar{k}_2|} \end{aligned} \quad (43)$$

where we obtain (a) by applying the PSP approximation in (37), obtain (b) by considering the range of Ω in (26) and the expression of $\tilde{\Theta}_m$ in (28), and obtain (c) by considering

the range of b in (26) and the expression of \bar{k}_1 in (20). Combining (42) and (43), we have

$$\begin{cases} \bar{k}_2 \leq -B, & \bar{m} \text{ is even} \\ \bar{k}_2 \geq \bar{b} + B, & \bar{m} \text{ is odd} \end{cases} \quad (44)$$

where

$$B \triangleq \frac{\bar{\Theta}_2 (1/N_t + (\bar{b} - \tilde{k}_1) N_t + M_2 \bar{\Theta}_2)}{2}. \quad (45)$$

With the setting of \bar{k}_2 in (44), the constraints in (42) are satisfied. Then, based on $\mathcal{P}(\bar{y}_m^*)$, the unwrapped phase of \bar{y}_m^* can be expressed as

$$\mathcal{U}(\bar{y}_{m+1}^*) = \mathcal{U}(\bar{y}_m^*) + \text{mod}(\mathcal{P}(\bar{y}_{m+1}^*) - \mathcal{P}(\bar{y}_m^*) + \pi, 2\pi) - \pi \quad (46)$$

for $m = -M_2, \dots, 0, \dots, M_2 - 1$, where $\mathcal{U}(\bar{y}_{-M_2}^*)$ is initialized to be zero.

Remark 2: Note that two constraints on \bar{k}_2 are derived in (31) and (44), respectively. To ensure the successful performing of the beam training, both (31) and (44) should be satisfied. In addition, according to (14), a larger deviation between the surrogate distance of the channel and that of the codeword will result in a smaller beam gain. Therefore, we need to satisfy both (31) and (44) while reducing the deviation between the surrogate distance of the channel and that of the codeword. Based on the above discussions, we set \bar{k}_2 as

$$\bar{k}_2 = \begin{cases} \min\{-B, \bar{k}_1 - M_2 \bar{\Theta}_2 / N_t - 1/N_t^2\}, & \bar{m} \text{ is even} \\ \max\{\bar{b} + B, \bar{k}_1 + M_2 \bar{\Theta}_2 / N_t + 1/N_t^2\}, & \bar{m} \text{ is odd.} \end{cases} \quad (47)$$

Now, we provide an example for the second-refinement codebook design and the beam training based on the PSP in **Design Example 2**.

Design Example 2: Consider an XL-MIMO system with $N = 256$, $N_t = 513$, $\lambda = 0.005$ m, $\bar{k}_1 = -6.09 \times 10^{-5}$, $\bar{b} = 1.22 \times 10^{-4}$, $\bar{\Theta}_2 = 2/N_t$, and $M_2 = 8$. The number of channel paths are set to $L = 1$ and the corresponding channel steering vector is $\gamma(\Omega, b)$, where $b = 4.9782 \times 10^{-5}$ and $\Omega = 0.05$. Suppose the beam training result of the first

Algorithm 2 ML-Based and PSP-Based Second Refinement Method

- 1: **Input:** N , N_t , λ , \mathbf{h} , \bar{m} , \bar{b} , \bar{k}_1 , \tilde{k}_1 , M_2 , $\bar{\Theta}_2$.
 - 2: Obtain \bar{k}_2 via (47).
 - 3: Obtain \mathcal{C}_2 via (27).
 - 4: Obtain \bar{y}_m for $m = -M_2, \dots, 0, \dots, M_2$ via (32).
 - 5: **if** ML is adopted **then**
 - 6: Obtain $\hat{\Omega}_2$ and \hat{b}_2 via (35).
 - 7: **else if** PSP is adopted **then**
 - 8: Obtain $\mathcal{U}(\bar{y}_m^*)$ via (46).
 - 9: Obtain $\hat{\Omega}_2$ and \hat{b}_2 via (41).
 - 10: **end if**
 - 11: **Output:** $\hat{\Omega}_2$ and \hat{b}_2 .
-

refinement is $\bar{m} = 0$. According to (25) and (26), we can obtain potential region $\tilde{\Phi}_1$ and extended potential region Φ_2 , which are illustrated as the red and the blue trapezoidal regions in Fig. 4(a), respectively. According to (47), we can obtain \bar{k}_2 . Then, we can design \mathbf{w}_m based on (28). As shown in Fig. 4(a), the beam coverage of \mathbf{w}_{m+1} can be obtained through translating the beam coverage of \mathbf{w}_m by $\bar{\Theta}_2$. In Fig. 4(b), we illustrate the wrapped phase, the unwrapped phase, and the approximated phase of \bar{y}_m^* via the PSP. It is shown that the phase of \bar{y}_m^* is well unwrapped via (46) and the unwrapped phase can be well approximated by the PSP. In Fig. 4(c), we illustrate the beam gain of $\gamma(\Omega, b)$ and the values of the real parameters as well as the estimated parameters in (41). It is shown that the estimated parameters are close to the real parameters, which verifies the effectiveness of the second refinement.

Finally, we summarize the ML-based and PSP-based second refinement methods in **Algorithm 2**.

V. GA-BASED THIRD REFINEMENT METHOD

In this section, we develop the GA-based third refinement method to further improve the beam training accuracy. First, we perform the hybrid-field neighboring search to identify the potential region of the main lobe of the channel steering vector. Then, by applying the GA, an LS estimator is developed to obtain the high-accuracy channel parameter estimation.

A. Hybrid-Field Neighboring Search

In this part, we present the details of the hybrid-field neighboring search, which includes initialization, beam training, and stop conditions.

1) *Initialization:* We divide the hybrid-field neighboring search into M_n groups, each consisting of five times of beam training. The parameters of the central codeword in the neighboring search are initialized to be $\tilde{\Theta}_n^{(1)} = \hat{\Omega}_2$ and $\tilde{k}_n^{(1)} = \hat{b}_2$. The surrogate distance deviation and the angle deviation between adjacent codewords are set to \bar{k}_n and $\bar{\Theta}_n$, respectively. The index of the current group is initialized to be $m = 0$.

2) *Beam Training:* We update $m \leftarrow m + 1$ and design the codebook for the beam training of the m th group as

$$\mathcal{C}_n^{(m)} = \{ \gamma(\tilde{\Theta}_n^{(m)} - \bar{\Theta}_n, \tilde{k}_n^{(m)}), \gamma(\tilde{\Theta}_n^{(m)} + \bar{\Theta}_n, \tilde{k}_n^{(m)}) \},$$

$$\gamma(\tilde{\Theta}_n^{(m)}, \tilde{k}_n^{(m)} - \bar{k}_n), \gamma(\tilde{\Theta}_n^{(m)}, \tilde{k}_n^{(m)} + \bar{k}_n), \gamma(\tilde{\Theta}_n^{(m)}, \tilde{k}_n^{(m)}) \}. \quad (48)$$

After that, we perform beam training with $\mathcal{C}_n^{(m)}$ and obtain

$$\tilde{y}_s^{(m)} = \mathbf{h}^H \{ \mathcal{C}_n^{(m)} \}_s + \eta \quad (49)$$

for $s = 1, 2, 3, 4, 5$. The index of the codeword in $\mathcal{C}_n^{(m)}$ with the biggest received power can be expressed as

$$\tilde{s}^{(m)} = \arg \max_{s=1,2,3,4,5} |\tilde{y}_s^{(m)}|. \quad (50)$$

Then we update the parameters of the central codeword as

$$(\tilde{\Theta}_n^{(m+1)}, \tilde{k}_n^{(m+1)}) = \Xi(\{ \mathcal{C}_n^{(m)} \}_{\tilde{s}^{(m)}}) \quad (51)$$

where $\Xi(\cdot)$ denotes the channel parameters of a steering vector, e.g., $(\Omega, b) = \Xi(\gamma(\Omega, b))$.

3) *Stop Conditions:* We repeat the beam training procedure in Section V-A.2 until $\tilde{s}^{(m)} = 5$ or $m = M_n$, where the index of the training group at this point is expressed as \tilde{m} . Note that $\tilde{s}^{(m)} = 5$ implies that the neighboring search has converged due to $\tilde{\Theta}_n^{(\tilde{m}+1)} = \tilde{\Theta}_n^{(\tilde{m}+2)}$ and $\tilde{k}_n^{(\tilde{m}+1)} = \tilde{k}_n^{(\tilde{m}+2)}$. We express the neighboring search results as

$$\hat{\Omega}_n = \tilde{\Theta}_n^{(\tilde{m}+1)}, \text{ and } \hat{b}_n = \tilde{k}_n^{(\tilde{m}+1)}. \quad (52)$$

Remark 3: Usually, \bar{k}_n and $\bar{\Theta}_n$ can be designed according to the coherence of the adjacent codewords, which is equivalent to the main lobe region determination of the channel steering vector. Note that the coherence of the two codewords that are adjacent in angle equals that of the two adjacent discrete Fourier transform (DFT) codewords [7]. Therefore, we can set $\bar{\Theta}_n = 2/N_t$ emulating the merits of the DFT codebook [18]. The coherence of the two codewords that are adjacent in distance can be calculated as

$$\begin{aligned} \rho &= \frac{1}{N_t} |G(\mathbf{u}, \Theta, k + \bar{k}_n)| \\ &\stackrel{(a)}{\approx} \frac{1}{N_t} \left| \int_{-N}^N e^{j\pi \bar{k}_n z^2} dz \right| \\ &= \frac{1}{N_t} \sqrt{\frac{2\mathcal{C}(\sqrt{2\bar{k}_n}N)^2 + 2\mathcal{S}(\sqrt{2\bar{k}_n}N)^2}{\bar{k}_n}} \end{aligned} \quad (53)$$

where we approximate the summation as integral in (a), $\mathcal{C}(\sqrt{2\bar{k}_n}N) = \int_0^{\sqrt{2\bar{k}_n}N} \cos(\pi z^2/2) dz$ and $\mathcal{S}(\sqrt{2\bar{k}_n}N) = \int_0^{\sqrt{2\bar{k}_n}N} \sin(\pi z^2/2) dz$ are the Fresnel functions. Given ρ , the value of \bar{k}_n can be obtained via (53). For example, if $N_t = 513$ and $\rho = 0.35$, we can obtain $\bar{k}_n = 2.28 \times 10^{-5} \approx 6/N_t^2$.

Remark 4: From Section V-A.3, the procedure of the hybrid-field neighboring search stops when $\tilde{s}^{(\tilde{m})} = 5$ or $\tilde{m} = M_n$. If $\tilde{s}^{(\tilde{m})} = 5$ (\tilde{m} equals M_n or not), which implies that the central codeword has the largest received power, we say that the neighboring search is successful. Then we can implement the GA to further improve the training performance. Otherwise, we say that the neighboring search fails. Then the THBT stops at this step and its results are expressed as

$$\hat{b}_f = \hat{b}_n, \text{ and } \hat{\Omega}_f = \hat{\Omega}_n. \quad (54)$$

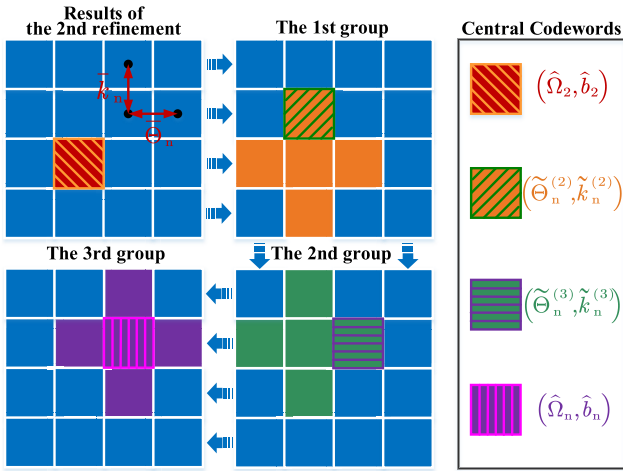


Fig. 5. Illustration of the hybrid-field neighboring search.

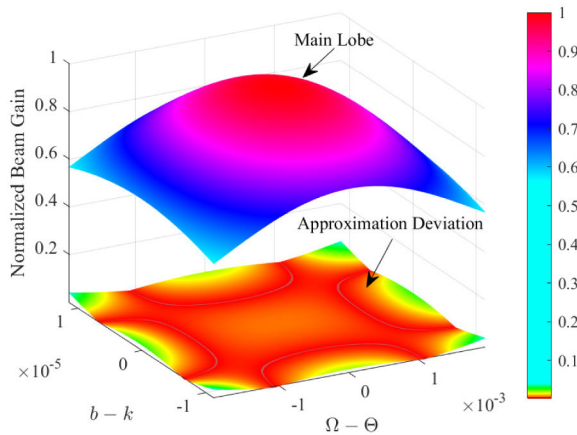


Fig. 6. Illustration of the Gaussian approximation.

Now, we provide an example for the hybrid-field neighboring search in **Design Example 3**.

Design Example 3: In Fig. 5, we consider the neighboring search consisting of $M_n = 3$ groups, where each group tests the adjacent four codewords of the central codeword. If the best codeword in the third group is the one with pink borders, the neighboring search is successful. Otherwise, it fails. In addition, it is noteworthy that there are two overlapping codewords between the adjacent groups. Therefore, each group only needs to test three additional codewords except the first one. In total, the maximum training overhead of the hybrid-field neighboring search is $3M_n + 2$.

B. Gaussian Approximation

From (9), we have

$$\begin{aligned}
 |G(\mathbf{u}, \Omega, b)|^2 &= \left(\sum_{n=-N}^N e^{j(\tilde{b}n^2 - \tilde{\Omega}n)} \right) \left(\sum_{n=-N}^N e^{j(\tilde{b}n^2 - \tilde{\Omega}n)} \right)^* \\
 &= \sum_{n=-N}^N \sum_{i=-N}^N e^{j(\tilde{\Omega}(i-n) + \tilde{b}(n^2 - i^2))} \\
 &= \sum_{n=-N}^N \sum_{i=-N}^N \cos(\tilde{\Omega}(i-n) + \tilde{b}(n^2 - i^2))
 \end{aligned} \tag{55}$$

where $\tilde{\Omega} \triangleq (\Omega - \Theta)\pi$ and $\tilde{b} \triangleq (b - k)\pi$. In fact, it is hard to obtain a closed-form expression of (55) due to the quadratic phase term $\tilde{b}(n^2 - i^2)$. To obtain a deeper scope of $|G(\mathbf{u}, \Omega, b)|^2$, we resort to the widely-used Taylor series. At the bottom of the next page, we provide the Taylor series of $|G(\mathbf{u}, \Omega, b)|^2$ and $e^{-\tilde{\Omega}^2/\sigma_1^2 - \tilde{b}^2/\sigma_2^2}$ about the point $(0, 0)$, where $C_p, \tilde{C}_p, \bar{C}_{p,q}, D_p, \tilde{D}_p, \bar{D}_{p,q}$ are all constant coefficients. In addition, we omit lower-order terms in (a) of (56) for simplicity. From (56) and (57), the Taylor series of $|G(\mathbf{u}, \Omega, b)|^2$ and $e^{-\tilde{\Omega}^2/\sigma_1^2 - \tilde{b}^2/\sigma_2^2}$ share similarities in several aspects including the orders and coefficients of the series. Inspired by their similarities, we use the two-dimensional Gaussian function

$$f(\Omega, b) \triangleq ae^{-\frac{(\Omega - \Theta)^2}{2\sigma_1^2} - \frac{(b - k)^2}{2\sigma_2^2}} \tag{58}$$

to approximate the main lobe of $|G(\mathbf{u}, \Omega, b)|$, which is termed as the Gaussian approximation. Note that the Gaussian approximation has been investigated in [24], [44], and [45]. However, they only focus on the far-field scenarios. In this work, we demonstrate that the main lobe of the HFBG can also be approximated by the Gaussian functions. According to the **Remark 3** of Section V, the main lobe of $|G(\mathbf{u}, \Omega, b)|$ is restricted to

$$\{(\Omega, b) | \Omega \in [\Theta - \bar{\Theta}_3, \Theta + \bar{\Theta}_3], b \in [k - \bar{k}_3, k + \bar{k}_3]\} \tag{59}$$

where $\bar{\Theta}_3 = \bar{\Theta}_n/2$ and $\bar{k}_3 = \bar{k}_n/2$. Then, the Gaussian approximation is formulated as

$$\min_{a, \sigma_1, \sigma_2} \int_{k - \bar{k}_3}^{k + \bar{k}_3} \int_{\Theta - \bar{\Theta}_3}^{\Theta + \bar{\Theta}_3} |f(\Omega, b) - |G(\mathbf{u}, \Omega, b)||^2 d\Omega db \tag{60}$$

which is a nonlinear least-square problem and can be solved by the trust-region optimization algorithm [46]. We omit the details and denote the solutions of (60) as $\hat{a}, \hat{\sigma}_1$ and $\hat{\sigma}_2$. Then, the optimized Gaussian function can be expressed as

$$\hat{f}(\Omega, b) = \hat{a}e^{-\frac{(\Omega - \Theta)^2}{2\hat{\sigma}_1^2} - \frac{(b - k)^2}{2\hat{\sigma}_2^2}}. \tag{61}$$

Note that for another channel steering vector, $\vec{\mathbf{u}} = \gamma(\vec{\Theta}, \vec{k})$, we have

$$\begin{aligned}
 G(\vec{\mathbf{u}}, \Omega, b) &= \sum_{n=-N}^N e^{j\pi((\vec{\Theta} - \Omega)n - (\vec{k} - b)n^2)} \\
 &= \sum_{n=-N}^N e^{j\pi((\Theta - (\Theta - \vec{\Theta} + \Omega))n - (k - (k - \vec{k} + b))n^2)} \\
 &= G(\mathbf{u}, \Omega + (\Theta - \vec{\Theta}), b + (k - \vec{k}))
 \end{aligned} \tag{62}$$

which indicates that the beam gain of $\vec{\mathbf{u}}$ is the translation of that of \mathbf{u} . Therefore, we only need to solve (60) for \mathbf{u} , and the Gaussian approximation for other channel steering vectors can be obtained via the translation in (62).

In Fig. 6, we illustrate the main lobe of $|G(\mathbf{u}, \Omega, b)|$ and the deviation of the Gaussian approximation, where we set $N = 256$, $N_t = 513$, $\bar{k}_3 = 3/N_t$, and $\bar{\Theta}_3 = 1/N_t$. From the figure, the approximation deviation is quite small compared to the beam gain of the main lobe. For example, the maximum and the averaged approximation deviations are only 4% and 0.5% of the maximum beam gain, respectively.

C. GA-Based Channel Parameter Estimation

In this part, we implement the channel parameter estimation based on Gaussian approximation to improve the estimation accuracy with low overhead.

First, we determine the potential region of Ω and b based on the results of the neighboring search. From (52), the potential intervals of Ω and b can be expressed as

$$\begin{aligned}\Omega &= [\hat{\Omega}_n - \bar{\Theta}_3, \hat{\Omega}_n + \bar{\Theta}_3], \\ \mathcal{B} &= [\hat{b}_n - \bar{k}_3, \hat{b}_n + \bar{k}_3].\end{aligned}\quad (63)$$

Then, we further narrow down the regions of Ω and b by comparing the powers of the received signals adjacent to $\tilde{y}_5^{(\tilde{m})}$. The narrowed potential region of Ω is expressed as

$$\tilde{\Omega} = \begin{cases} [\hat{\Omega}_n, \hat{\Omega}_n + \bar{\Theta}_3], & |\tilde{y}_2^{(\tilde{m})}| \geq |\tilde{y}_1^{(\tilde{m})}| \\ [\hat{\Omega}_n - \bar{\Theta}_3, \hat{\Omega}_n], & |\tilde{y}_2^{(\tilde{m})}| < |\tilde{y}_1^{(\tilde{m})}| \end{cases}\quad (64)$$

where the conditions can be obtained conveniently according to (48) and (49). Similarly, the narrowed potential region of b can be expressed as

$$\tilde{\mathcal{B}} = \begin{cases} [\hat{b}_n, \hat{b}_n + \bar{k}_3], & |\tilde{y}_4^{(\tilde{m})}| \geq |\tilde{y}_3^{(\tilde{m})}| \\ [\hat{b}_n - \bar{k}_3, \hat{b}_n], & |\tilde{y}_4^{(\tilde{m})}| < |\tilde{y}_3^{(\tilde{m})}|. \end{cases}\quad (65)$$

After that, we quantize the narrowed potential regions of Ω and b by M_3 samples. We denote the left and right boundaries of $\tilde{\Omega}$ as Ψ_L and Ψ_R , respectively. Similarly, we have $\tilde{\mathcal{B}} = [d_L, d_R]$. Then the m th sample of the angle and the surrogate distance can be expressed as

$$\begin{aligned}\tilde{\Theta}_3^{(m)} &= \Psi_L + \frac{(m-1)(\Psi_R - \Psi_L)}{M_3 - 1}, \\ \tilde{k}_3^{(m)} &= d_L + \frac{(m-1)(d_R - d_L)}{M_3 - 1}.\end{aligned}\quad (66)$$

We stack the codewords directing to the quantized samples in (66) as \mathcal{C}_3 , where

$$\begin{aligned}\mathcal{C}_3 &= \left\{ \tilde{\mathcal{C}}_3^{(1)}, \tilde{\mathcal{C}}_3^{(2)}, \dots, \tilde{\mathcal{C}}_3^{(M_3)} \right\}, \\ [\tilde{\mathcal{C}}_3^{(m)}]_t &= \gamma(\tilde{\Theta}_3^{(m)}, \tilde{k}_3^{(t)})\end{aligned}\quad (67)$$

Algorithm 3 GA-Based Third Refinement Method

- 1: **Input:** $N, N_t, \lambda, \mathbf{h}, \hat{\Omega}_2, \hat{b}_2, M_n, M_3$.
- 2: Obtain \hat{k}_n and $\hat{\Theta}_n$ via **Remark 3**.
- 3: Obtain \tilde{m} and $\tilde{s}^{(\tilde{m})}$ via Section V-A.3.
- 4: Obtain \hat{b}_n and $\hat{\Omega}_n$ via (52).
- 5: **if** $\tilde{m} = M_n$ and $\tilde{s}^{(\tilde{m})} \neq 5$ **then**
- 6: $\hat{b}_f \leftarrow \hat{b}_n, \hat{\Omega}_f \leftarrow \hat{\Omega}_n$.
- 7: **else**
- 8: Obtain \mathcal{C}_3 via (67).
- 9: Obtain $\hat{y}_{m,t}$ via (68).
- 10: Obtain \hat{b}_3 and $\hat{\Omega}_3$ via (76).
- 11: $\hat{b}_f \leftarrow \hat{b}_3, \hat{\Omega}_f \leftarrow \hat{\Omega}_3$.
- 12: **end if**
- 13: Obtain \hat{r}_f via (78).
- 14: **Output:** $\hat{b}_f, \hat{\Omega}_f, \hat{r}_f$.

for $m = 1, 2, \dots, M_3$ and $t = 1, 2, \dots, M_3$. Then we perform the beam training with \mathcal{C}_3 . Similar to (32), the received signal can be expressed as

$$\hat{y}_{m,t} \approx g^* \gamma(\Omega, b)^H \gamma(\tilde{\Theta}_3^{(m)}, \tilde{k}_3^{(t)}) + \eta. \quad (68)$$

We specify that all the variants of η denote the noise terms in the following texts.

By applying the Gaussian approximation to (68), we have

$$\begin{aligned}|\hat{y}_{m,t}| &\approx |\tilde{g}| \hat{f}(\tilde{\Theta}_3^{(m)}, \tilde{k}_3^{(t)}) + \tilde{\eta} \\ &= |\tilde{g}| \hat{a} e^{-\frac{(\Omega - \tilde{\Theta}_3^{(m)})^2}{2\hat{\sigma}_1^2} - \frac{(b - \tilde{k}_3^{(t)})^2}{2\hat{\sigma}_2^2}} + \tilde{\eta}\end{aligned}\quad (69)$$

where $\tilde{g} = g^*/N_t$. Note that the Gaussian function is the exponential of the quadratic function. A straightforward method to simplify the analysis is to take the natural logarithm of $|\hat{y}_{m,t}|$. Then, we have

$$\ln |\hat{y}_{m,t}| \approx \ln(|\tilde{g}| \hat{a}) - \frac{(\Omega - \tilde{\Theta}_3^{(m)})^2}{2\hat{\sigma}_1^2} - \frac{(b - \tilde{k}_3^{(t)})^2}{2\hat{\sigma}_2^2} + \tilde{\eta}, \quad (70)$$

which is a quadratic function about the channel parameters Ω and b . Note that

$$\ln |\hat{y}_{m,t}| \approx \ln(|\tilde{g}| \hat{f}(\tilde{\Theta}_3^{(m)}, \tilde{k}_3^{(t)})) + \ln \left(1 + \frac{\tilde{\eta}}{|\tilde{g}| \hat{f}(\tilde{\Theta}_3^{(m)}, \tilde{k}_3^{(t)})} \right)$$

$$\begin{aligned}|G(\mathbf{u}, \Omega, b)|^2 &= \sum_{n=-N}^N \sum_{i=-N}^N \sum_{p=0}^{\infty} \sum_{q=1}^{2p-1} \frac{(-1)^p (i-n)^{2p} \tilde{\Omega}^{2p}}{(2p)!} + \frac{(-1)^p (n^2 - i^2)^{2p} \tilde{b}^{2q}}{(2p)!} + 2 \frac{(-1)^p (i-n)^q (n^2 - i^2)^{2p-q} \tilde{\Omega}^q \tilde{b}^{2p-q}}{(q)!(2p-q)!} \\ &\stackrel{(a)}{\approx} \sum_{p=0}^{\infty} \sum_{q=1}^{2p-1} \frac{(-1)^p C_p N^{2p} \tilde{\Omega}^{2p}}{(2p)!} + \frac{(-1)^p \tilde{C}_p (N^2)^{2p} \tilde{b}^{2q}}{(2p)!} + 2 \frac{(-1)^p \tilde{C}_{p,q} N^{4p-q+2} \tilde{\Omega}^q \tilde{b}^{2p-q}}{(q)!(2p-q)!}\end{aligned}\quad (56)$$

$$\begin{aligned}e^{-\frac{\hat{\Omega}_2^2}{\hat{\sigma}_1^2} - \frac{\hat{b}_2^2}{\hat{\sigma}_2^2}} &= \sum_{p=0}^{\infty} \sum_{q=1}^{p-1} \frac{(-1)^p \left(\frac{2}{\hat{\sigma}_1^2}\right)^p \tilde{\Omega}^{2p} \prod_{i=1}^p (2i-1)}{(2p)!} + \frac{(-1)^p \left(\frac{2}{\hat{\sigma}_2^2}\right)^p \tilde{b}^{2q} \prod_{i=1}^p (2i-1)}{(2p)!} \\ &+ 2 \frac{(-1)^p \left(\frac{2}{\hat{\sigma}_1^2}\right)^q \left(\frac{2}{\hat{\sigma}_2^2}\right)^{p-q} \left(\prod_{i=1}^q (2i-1)\right) \left(\prod_{i=1}^{p-q} (2i-1)\right) \tilde{\Omega}^{2q} \tilde{b}^{2p-2q}}{(2q)!(2p-2q)!} \\ &= \sum_{p=0}^{\infty} \sum_{q=1}^{p-1} \frac{(-1)^p D_p (\hat{\sigma}_1^{-1})^{2p} \tilde{\Omega}^{2p}}{(2p)!} + \frac{(-1)^p \tilde{D}_p (\hat{\sigma}_2^{-1})^{2p} \tilde{b}^{2q}}{(2p)!} + 2 \frac{(-1)^p \tilde{D}_{p,q} (\hat{\sigma}_1^{-1})^{2q} (\hat{\sigma}_2^{-1})^{2p-2q} \tilde{\Omega}^{2q} \tilde{b}^{2p-2q}}{(2q)!(2p-2q)!}\end{aligned}\quad (57)$$

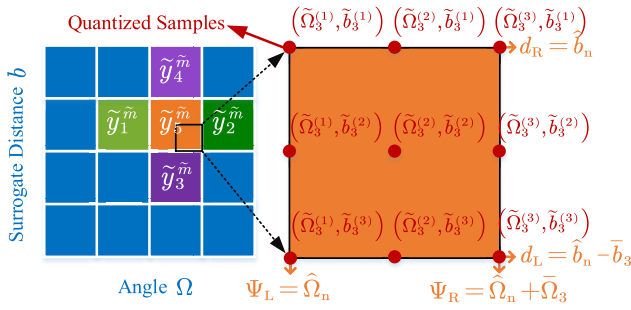


Fig. 7. Illustration of the GA-based channel parameter estimation.

$$\stackrel{(a)}{\approx} \ln(|\tilde{g}| \hat{f}(\tilde{\Theta}_3^{(m)}, \tilde{k}_3^{(t)})) + \frac{\tilde{\eta}}{|\tilde{g}| \hat{f}(\tilde{\Theta}_3^{(m)}, \tilde{k}_3^{(t)})} \quad (71)$$

where (a) holds because $\ln(1 + \epsilon) \approx \epsilon$. The relations in (71) indicate that the noise term will be magnified by $\frac{1}{|\tilde{g}| \hat{f}(\tilde{\Theta}_3^{(m)}, \tilde{k}_3^{(t)})}$ times in the process of taking the logarithm. Therefore, large errors will be introduced for small values of $\hat{f}(\tilde{\Theta}_3^{(m)}, \tilde{k}_3^{(t)})$. To avoid the noise amplification effects, we multiply (70) by $|\hat{y}_{m,t}|$ [47], [48], and obtain

$$\Upsilon_{m,t} = \frac{|\hat{y}_{m,t}| \tilde{\Theta}_3^{(m)}}{\hat{\sigma}_1^2} \Omega + \frac{|\hat{y}_{m,t}| \tilde{k}_3^{(t)}}{\hat{\sigma}_2^2} b + |\hat{y}_{m,t}| \chi + \hat{\eta} \quad (72)$$

where

$$\begin{aligned} \Upsilon_{m,t} &= |\hat{y}_{m,t}| \ln |\hat{y}_{m,t}| + \frac{|\hat{y}_{m,t}| (\tilde{\Theta}_3^{(m)})^2}{2\hat{\sigma}_1^2} + \frac{|\hat{y}_{m,t}| (\tilde{k}_3^{(t)})^2}{2\hat{\sigma}_2^2} \\ \chi &= \ln(|\tilde{g}| \hat{a}) - \frac{\Omega^2}{2\hat{\sigma}_1^2} - \frac{b^2}{2\hat{\sigma}_2^2}. \end{aligned} \quad (73)$$

We rewrite (72) in the vector form as

$$\mathbf{A}z + \hat{\boldsymbol{\eta}} = \mathbf{y}, \quad (74)$$

where $z \triangleq [\Omega, b, \chi]^T$ and $\hat{\boldsymbol{\eta}}$ is the stack of noise terms. The u th row of \mathbf{A} and the u th entry of \mathbf{y} are expressed as

$$[\mathbf{A}]_{u,:} = \left[\frac{|\hat{y}_{m,t}| \tilde{\Theta}_3^{(m)}}{\hat{\sigma}_1^2}, \frac{|\hat{y}_{m,t}| \tilde{k}_3^{(t)}}{\hat{\sigma}_2^2}, |\hat{y}_{m,t}| \right], \quad [\mathbf{y}]_u = \Upsilon_{m,t}, \quad (75)$$

respectively, for $u = (m-1)M_3 + t$, $m = 1, 2, \dots, M_3$, and $t = 1, 2, \dots, M_3$. The LS solution of (74) is $\hat{z} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{y}$. Then the estimates of Ω and b can be expressed as

$$\hat{\Omega}_3 = [\hat{z}]_1, \quad \text{and} \quad \hat{b}_3 = [\hat{z}]_2. \quad (76)$$

Now we provide an example for the GA-based channel parameter estimation in **Design Example 4**.

Design Example 4: In Fig. 7, we illustrate the process of the GA-based channel parameter estimation, where $M_3 = 3$, $|\tilde{y}_2^{(m)}| > |\tilde{y}_1^{(m)}|$, and $|\tilde{y}_3^{(m)}| > |\tilde{y}_4^{(m)}|$. From (64) and (65), $\Psi_L = \hat{\Omega}_n$, $\Psi_R = \hat{\Omega}_n + \hat{\Theta}_3$, $d_L = \hat{b}_n - \hat{k}_3$, and $d_R = \hat{b}_n$. Then the quantized samples can be obtained via (66). Note that one of the quantized samples is exactly the central codeword in the hybrid-field neighboring search, whose received signal can be reused for the GA. Therefore, the training overhead of the GA-based channel parameter estimation is $M_3^2 - 1$.

The final results of the THBT are expressed as

$$\hat{b}_f = \hat{b}_3, \quad \text{and} \quad \hat{\Omega}_f = \hat{\Omega}_3. \quad (77)$$

According to (7), the estimation of the distance can be expressed as

$$\hat{r}_f = \frac{\lambda(1 - \hat{\Omega}_f^2)}{4\hat{b}_f}. \quad (78)$$

Finally, we summarize the details of the GA-based third refinement method in **Algorithm 3**.

Now we evaluate the computational complexity of the proposed THBT-ML and THBT-PSP. For the THBT-PSP, the first refinement method only involves a straightforward comparison of the received signal power, resulting in a computational complexity of $\mathcal{O}(M_1)$. The second and third refinement methods provide closed-form expressions for channel parameter estimation, as shown in (41) and (76), respectively. Their computational complexities are $\mathcal{O}(M_2)$ and $\mathcal{O}(M_3^2)$, respectively. Therefore, the computational complexity of the THBT-PSP is $\mathcal{O}(\max\{M_1, M_2, M_3^2\})$. On the other hand, the THBT-ML shares the same first and third refinement methods as the THBT-PSP. However, during the second refinement step, it employs a two-dimensional search to solve (35). Denote the number of the searches as V . Then, the computational complexity of the THBT-ML is $\mathcal{O}(V(2M_2 + 1))$.

Remark 5: The proposed scheme is also adaptable to the uniform planar array (UPA) configuration. According to Lemma 3 of [49], the hybrid-field steering vector of a UPA can be approximated as the Kronecker product of channel steering vectors for the horizontal uniform linear array (ULA) and the vertical ULA, which indicates that beam training of a UPA can be decoupled into the beam training of two ULAs. By performing the beam training for the horizontal ULA and the vertical ULA separately, the proposed scheme can be extended to the UPA configuration.

VI. SIMULATION RESULTS

Now we evaluate the performance of the proposed THBT scheme. We consider an XL-MIMO system including a BS equipped with $N_t = 513$ antennas and a single-antenna user. We set the wavelength as $\lambda = 0.005$ m corresponding to the carrier frequency of 60 GHz. The channel between the BS and the user is composed of one line-of-sight (LoS) path and two NLoS paths, where the channel gain of the LoS path obeys $g_1 \sim \mathcal{CN}(0, \delta_1^2)$ and that of the NLoS paths obeys $g_l \sim \mathcal{CN}(0, \delta_l^2)$ for $l \in \{2, 3\}$. The channel angles distribute uniformly within $[-\sqrt{3}/2, \sqrt{3}/2]$. The HFBS [25], TPBT [26], CHBT [27], DHBT [28], and TSHBT [29] are adopted as benchmarks. The parameters of different methods are set in the **Parameter Settings** of Table I, where P , Q , K , O , T , R , W_1 , and W_2 denote the angle samples, the distance samples, the number of candidate angles, the number of codewords in the first layer, the number of layers in the hierarchical codebook, and the number of space samples in each layer, the number of layers in the first stage of TSHBT, and the number of layers in the second stage of TSHBT, respectively. In addition, for the introduced parameters in the THBT, we set $\tilde{k}_1 = -6.09 \times 10^{-5}$, $\bar{\Theta}_2 = 2/N_t$, $\bar{k}_n = 6/N_t^2$ and $\bar{\Theta}_n = 2/N_t$.

In Fig. 8, we compare the beamforming gains of different methods. The beamforming gain after beam alignment is

TABLE I
COMPARISON OF TRAINING OVERHEAD FOR DIFFERENT METHODS

| Methods | Training Overhead | Parameter Settings | Calculated Training Overhead |
|---------------|-----------------------------------|--------------------------------------|------------------------------|
| HFBS [25] | PQ | $P = 513, Q = 9$ | 4617 |
| TPBT [26] | $N_t + KQ$ | $N_t = 513, K = 3, Q = 9$ | 540 |
| CHBT [27] | $O + 4(T - 1)$ | $O = 33, T = 5$ | 49 |
| DHBT [28] | TR | $R = 100, T = 2$ | 200 |
| TSHBT [29] | $2W_1 + 4W_2$ | $W_1 = 7, W_2 = 3$ | 26 |
| Proposed THBT | $2(M_1 + M_2) + 3M_n + M_3^2 + 3$ | $M_1 = 7, M_2 = 8, M_n = 3, M_3 = 2$ | 46 |

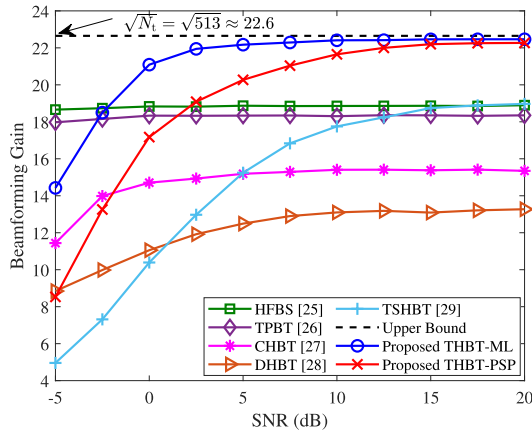


Fig. 8. Comparisons of the beamforming gains for different methods.

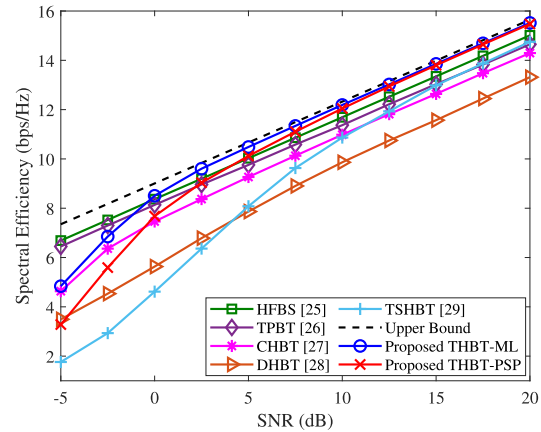


Fig. 9. Comparisons of the spectral efficiency for different methods.

defined as

$$\xi = \max_l \frac{|g_l|}{g_m} |\alpha(\Omega_l, r_l)^H \alpha(\Omega_f, r_f)|, \quad (79)$$

where $g_m = \max_l |g_l|$. The distances between the BS and the user or scatterers obey the uniform distribution within $[10, 200]$ m. We set $\delta_1 = 1$ and $\delta_l = 0.1$ for $l \in \{2, 3\}$. From Fig. 8, at low signal-to-noise-ratios (SNRs), the HFBS achieves the best performance among all the methods; the justification is that the HFBS exhaustively tests the codewords in the hybrid-field codebook [25] and needs far more times of beam training than other methods. Then the TPBT achieves the second-best performance due to the accurate identification of candidate angles by far-field beam sweeping. With the increase of the SNR, the performance of the proposed THBT-ML and THBT-PSP improves significantly and exceeds that of other methods when the SNR is larger than 5 dB. The advantage of the proposed methods mainly comes from the three steps of progressive refinement. In addition, the CHBT performs worse than other methods because of the imperfect hierarchical codebook [27] while the poor performance of the DHBT is attributed to the neglect of the polar-domain sparsity [7]. The TSHBT performs worse than other methods at low SNRs due to low beamforming gains of the upper-layer codewords. Furthermore, at high SNRs, the proposed THBT-ML and THBT-PSP outperform other methods and can approach the upper bound.

In Fig. 9, we evaluate the spectral efficiency performance of different methods, where the simulation settings are the same as those in Fig. 8. From the figure, we can notice

that the spectral efficiency performance is consistent with the beamforming gain performance in Fig. 8. In addition, the proposed THBT-ML and THBT-PSP can approach the upper bound with a negligible gap, e.g., 0.2 bps/Hz at 15 dB.

In Fig. 10, we compare the beamforming gains of different methods with varying distances. The SNR is set to 10 dB and the distances between the BS and the user or scatterers obey the uniform distribution between $[10, r]$ m, where r ranges from 50 m to 600 m. From the figure, the THBT-ML achieves the best performance among all the methods, followed by the THBT-PSP, and then the HFBS, TPBT, TSHBT, CHBT and DHBT in descending order. In addition, the THBT-ML, THBT-PSP, HFBS, and CHBT are robust to distance changes due to the consideration of both the near-field and far-field effects. The performance of the TSHBT deteriorates at short distances because approximating the near-field channels with far-field ones by deactivating part of antennas may be not accurate enough for short distances. The performance of the DHBT deteriorates for short distances due to the neglect of the polar-domain sparsity.

In Fig. 11, we compare the positioning performance of different methods, where the SNR is set to 20 dB and the distances between the BS and the user or scatterers obey the uniform distribution between $[10, 30]$ m. The deviation between the true position and the estimated position is denoted as E and 10^5 times of Monte Carlo simulation are used to calculate the cumulative distribution function (CDF). From the figure, the THBT-ML and THBT-PSP achieve much better performance than the other five methods. For example, when

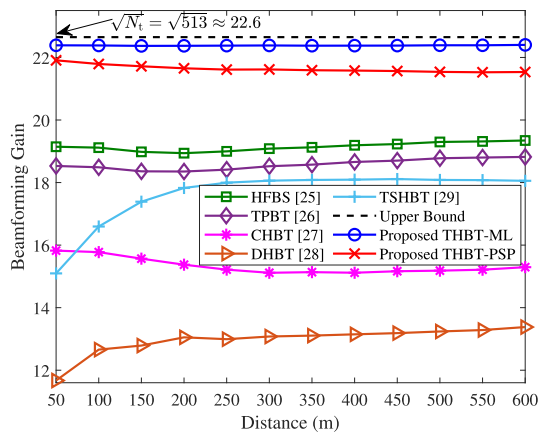


Fig. 10. Comparisons of the beamforming gains for different methods with varying distances.

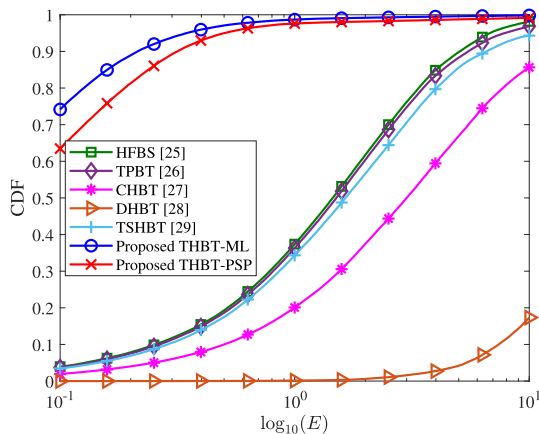


Fig. 11. Comparisons of the positioning performance for different methods.

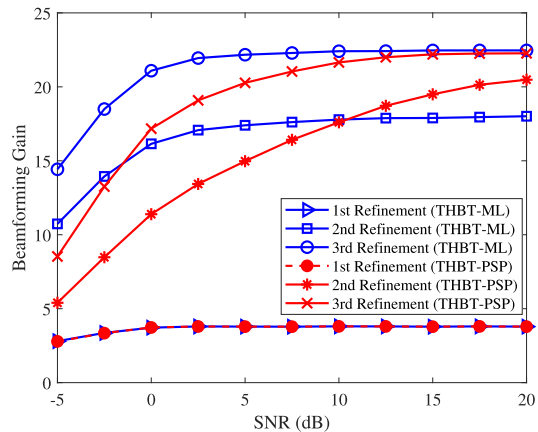


Fig. 12. Comparisons of the beamforming gains for different stages of refinement.

$E = 1$ m, the values of CDF for the THBT-ML, THBT-PSP, HFBS, TPBT, TSHBT, CHBT, and DHBT are 98.7%, 97.6%, 37.4%, 36.3%, 34.3%, 20.1%, and 0.13%, respectively, which indicates that the proposed schemes outperform existing ones and can attain high-accuracy positioning in most cases.

In Fig. 12, we compare the beamforming gains of the THBT-ML and THBT-PSP for different stages of refinement. From the figure, the THBT-ML and THBT-PSP have similar performance for the first refinement, which is due to

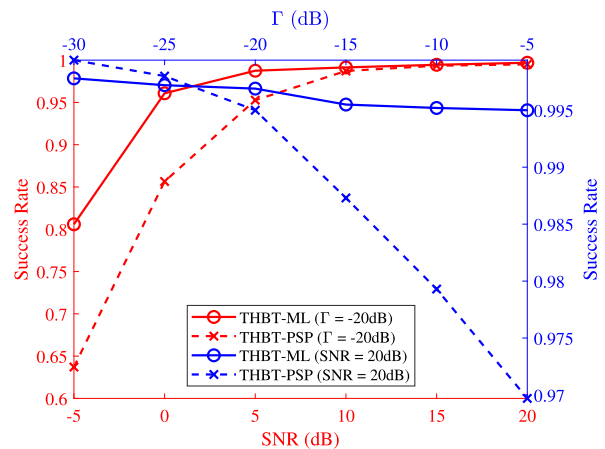


Fig. 13. Illustration of the success rate of the neighboring search.

the fact that the THBT-ML and THBT-PSP share the same first-refinement beam training. In addition, the beamforming gains of both the THBT-ML and the THBT-PSP improve with the progress of the THBT. Moreover, the THBT-ML outperforms the THBT-PSP at low SNRs due to the robustness of ML to noise while the performance of the THBT-PSP can approach that of the THBT-ML at high SNRs thanks to the exploitation of the phase property of the hybrid-field beam gain.

In Fig. 13, we evaluate the performance of the neighboring search in terms of success rate, where Γ denotes the ratio of the LoS path power to the NLoS path power, i.e., $\Gamma \triangleq 20 \log_{10}(\delta_1/\delta_l)$ for $l \in \{2, 3\}$. From the figure, the success rate of the neighboring search increases with SNR and Γ , which indicates that the performance of the neighboring search is affected by the effects of noise and NLoS paths. In addition, when $\Gamma = -20$ dB and the SNR is larger than 5 dB, the success rate of the neighboring search is considerably high, e.g. 98.8% for the THBT-ML. Therefore, the neighboring search can successfully find the mainlobe of the channel path in most cases.

In Table I, we compare the training overheads of different methods. The training overheads of the HFBS, TPBT, CHBT, DHBT, TSHBT, and the proposed THBT are PQ , $N_t + KQ$, $O + 4(T - 1)$, TR , $2W_1 + 4W_2$, and $2(M_1 + M_2) + 3M_n + M_3^3 + 3$, respectively. Under the simulation settings, these six methods require 4617, 540, 49, 200, 26, and 46 times of beam training, respectively. Specifically, when the SNR is larger than 5 dB, the proposed THBT outperforms the HFBS with a 99% reduction in training overhead.

VII. CONCLUSION

In this paper, beam training for XL-MIMO systems has been investigated. By considering both the near field and far field, a triple-refined hybrid-field beam training scheme has been proposed, where the HFBG-based first refinement method, the ML-based and PSP-based second refinement methods, and the GA-based third refinement method have been developed. Simulation results have shown that the proposed scheme outperforms the existing methods. In our future work, we will try to extend this work to the THz band by considering the

molecular absorption loss, THz spectrum windows, and beam split effects, following the works in [50], [51], and [52]. In addition, we will also focus on reducing the attainable latency for beam alignment by jointly considering the training overhead and feedback.

REFERENCES

- [1] K. Chen, C. Qi, and O. A. Dobre, "Beam refinement for THz extremely large-scale MIMO systems based on Gaussian approximation," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Kuala Lumpur, Malaysia, Dec. 2023, pp. 1–6.
- [2] E. Bjornson, L. Van der Perre, S. Buzzi, and E. G. Larsson, "Massive MIMO in sub-6 GHz and mmWave: Physical, practical, and use-case differences," *IEEE Wireless Commun.*, vol. 26, no. 2, pp. 100–108, Apr. 2019.
- [3] M. Chen, J. Guo, C.-K. Wen, S. Jin, G. Y. Li, and A. Yang, "Deep learning-based implicit CSI feedback in massive MIMO," *IEEE Trans. Commun.*, vol. 70, no. 2, pp. 935–950, Feb. 2022.
- [4] Z. Qiu, S. Zhou, M. Zhao, and W. Zhou, "Low-complexity precoding by exploiting spatial sparsity in massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 7, pp. 4740–4753, Jul. 2022.
- [5] C. Qi, P. Dong, W. Ma, H. Zhang, Z. Zhang, and G. Y. Li, "Acquisition of channel state information for mmWave massive MIMO: Traditional and machine learning-based approaches," *Sci. China Inf. Sci.*, vol. 64, no. 8, Aug. 2021, Art. no. 181301.
- [6] 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.
- [7] 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.
- [8] Y. Lu and L. Dai, "Near-field channel estimation in mixed LoS/NLoS environments for extremely large-scale MIMO systems," *IEEE Trans. Commun.*, vol. 71, no. 6, pp. 3694–3707, Jun. 2023.
- [9] K. T. Selvan and R. Janaswamy, "Fraunhofer and Fresnel distances: Unified derivation for aperture antennas," *IEEE Antennas Propag. Mag.*, vol. 59, no. 4, pp. 12–15, Aug. 2017.
- [10] 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.
- [11] Z. Chen, B. Ning, C. Han, Z. Tian, and S. Li, "Intelligent reflecting surface assisted terahertz communications toward 6G," *IEEE Wireless Commun.*, vol. 28, no. 6, pp. 110–117, Dec. 2021.
- [12] B. Ning et al., "Beamforming technologies for ultra-massive MIMO in terahertz communications," *IEEE Open J. Commun. Soc.*, vol. 4, pp. 614–658, 2023.
- [13] B. Ning, Z. Chen, Z. Tian, C. Han, and S. Li, "A unified 3D beam training and tracking procedure for terahertz communication," *IEEE Trans. Wireless Commun.*, vol. 21, no. 4, pp. 2445–2461, Apr. 2022.
- [14] X. Sun, C. Qi, and G. Y. Li, "Beam training and allocation for multiuser millimeter wave massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 18, no. 2, pp. 1041–1053, Feb. 2019.
- [15] C. Liu, M. Li, S. V. Hanly, I. B. Collings, and P. Whiting, "Millimeter wave beam alignment: Large deviations analysis and design insights," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 7, pp. 1619–1631, Jul. 2017.
- [16] J. Wang et al., "Beam codebook based beamforming protocol for multi-gbps millimeter-wave WPAN systems," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 8, pp. 1390–1399, Oct. 2009.
- [17] Z. Xiao, H. Dong, L. Bai, P. Xia, and X.-G. Xia, "Enhanced channel estimation and codebook design for millimeter-wave communication," *IEEE Trans. Veh. Technol.*, vol. 67, no. 10, pp. 9393–9405, Oct. 2018.
- [18] C. Qi, K. Chen, O. A. Dobre, and G. Y. Li, "Hierarchical codebook-based multiuser beam training for millimeter wave massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 8142–8152, Dec. 2020.
- [19] B. Ning, Z. Chen, W. Chen, Y. Du, and J. Fang, "Terahertz multi-user massive MIMO with intelligent reflecting surface: Beam training and hybrid beamforming," *IEEE Trans. Veh. Technol.*, vol. 70, no. 2, pp. 1376–1393, Feb. 2021.
- [20] B. Ning, T. Wang, C. Huang, Y. Zhang, and Z. Chen, "Wide-beam designs for Terahertz massive MIMO: SCA-ATP and S-SARV," *IEEE Internet Things J.*, vol. 10, no. 12, pp. 10857–10869, Jun. 2023.
- [21] M. Li, C. Liu, S. V. Hanly, I. B. Collings, and P. Whiting, "Explore and eliminate: Optimized two-stage search for millimeter-wave beam alignment," *IEEE Trans. Wireless Commun.*, vol. 18, no. 9, pp. 4379–4393, Sep. 2019.
- [22] S.-E. Chiu, N. Ronquillo, and T. Javidi, "Active learning and CSI acquisition for mmWave initial alignment," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 11, pp. 2474–2489, Nov. 2019.
- [23] 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.
- [24] J. A. Zhang, K. Wu, X. Huang, and Y. J. Guo, "Beam alignment for analog arrays based on Gaussian approximation," *IEEE Trans. Veh. Technol.*, vol. 72, no. 6, pp. 8152–8157, Jun. 2023.
- [25] 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)*, Foshan, China, Aug. 2022, pp. 1074–1079.
- [26] 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.
- [27] 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.*, early access, Aug. 14, 2023, doi: [10.1109/TWC.2023.3303229](https://doi.org/10.1109/TWC.2023.3303229).
- [28] 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.
- [29] C. Wu, C. You, Y. Liu, L. Chen, and S. Shi, "Two-stage hierarchical beam training for near-field communications," *IEEE Trans. Veh. Technol.*, early access, Sep. 5, 2023, doi: [10.1109/TVT.2023.3311868](https://doi.org/10.1109/TVT.2023.3311868).
- [30] G. Jiang and C. Qi, "Near-field beam training based on deep learning for extremely large-scale MIMO," *IEEE Commun. Lett.*, vol. 27, no. 8, pp. 2063–2067, Aug. 2023.
- [31] 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.
- [32] 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.
- [33] J. Liang and D. Liu, "Passive localization of mixed near-field and far-field sources using two-stage MUSIC algorithm," *IEEE Trans. Signal Process.*, vol. 58, no. 1, pp. 108–120, Jan. 2010.
- [34] Z. Wang, X. Mu, and Y. Liu, "Near-field integrated sensing and communications," *IEEE Commun. Lett.*, vol. 27, no. 8, pp. 2048–2052, Aug. 2023.
- [35] B. Friedlander, "Localization of signals in the near-field of an antenna array," *IEEE Trans. Signal Process.*, vol. 67, no. 15, pp. 3885–3893, Aug. 2019.
- [36] 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.*, early access, Oct. 19, 2023, doi: [10.1109/TWC.2023.3324176](https://doi.org/10.1109/TWC.2023.3324176).
- [37] N. M. Temme, *Asymptotic Methods for Integrals*, vol. 11. Singapore: World Scientific, 2014.
- [38] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers*. New York, NY, USA: McGraw-Hill, 1978.
- [39] F. W. J. Olver, *Asymptotics and Special Functions*. Boca Raton, FL, USA: CRC Press, 1997.
- [40] K. Chen, C. Qi, and G. Y. Li, "Two-step codeword design for millimeter wave massive MIMO systems with quantized phase shifters," *IEEE Trans. Signal Process.*, vol. 68, pp. 170–180, 2020.
- [41] J. Zhang, Y. Huang, Q. Shi, J. Wang, and L. Yang, "Codebook design for beam alignment in millimeter wave communication systems," *IEEE Trans. Commun.*, vol. 65, no. 11, pp. 4980–4995, Nov. 2017.
- [42] K. Chen and C. Qi, "Beam training based on dynamic hierarchical codebook for millimeter wave massive MIMO," *IEEE Commun. Lett.*, vol. 23, no. 1, pp. 132–135, Jan. 2019.
- [43] K. Itoh, "Analysis of the phase unwrapping algorithm," *Appl. Opt.*, vol. 21, no. 14, p. 2470, Jul. 1982.
- [44] G. Buttazzoni and R. Vescovo, "Density tapering of linear arrays radiating pencil beams: A new extremely fast Gaussian approach," *IEEE Trans. Antennas Propag.*, vol. 65, no. 12, pp. 7372–7377, Dec. 2017.
- [45] K. W. Lo, "Theoretical analysis of the sequential lobing technique," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 35, no. 1, pp. 282–293, Jan. 1999.

- [46] T. F. Coleman and Y. Li, "An interior trust region approach for nonlinear minimization subject to bounds," *SIAM J. Optim.*, vol. 6, no. 2, pp. 418–445, May 1996.
- [47] H. Guo, "A simple algorithm for fitting a Gaussian function [DSP tips and tricks]," *IEEE Signal Process. Mag.*, vol. 28, no. 5, pp. 134–137, Sep. 2011.
- [48] K. Wu, J. A. Zhang, and Y. J. Guo, "Fast and accurate linear fitting for an incompletely sampled Gaussian function with a long tail [tips & tricks]," *IEEE Signal Process. Mag.*, vol. 39, no. 6, pp. 76–84, Nov. 2022.
- [49] Z. Wu and L. Dai, "Multiple access for near-field communications: SDMA or LDMA?" *IEEE J. Sel. Areas Commun.*, vol. 41, no. 6, pp. 1918–1935, Jun. 2023.
- [50] I. F. Akyildiz, C. Han, Z. Hu, S. Nie, and J. M. Jornet, "Terahertz band communication: An old problem revisited and research directions for the next decade," *IEEE Trans. Commun.*, vol. 70, no. 6, pp. 4250–4285, Jun. 2022.
- [51] A. Shafie, N. Yang, C. Han, J. M. Jornet, M. Juntti, and T. Kürner, "Terahertz communications for 6G and beyond wireless networks: Challenges, key advancements, and opportunities," *IEEE Netw.*, vol. 37, no. 3, pp. 162–169, May 2023.
- [52] J. Tan and L. Dai, "THz precoding for 6G: Challenges, solutions, and opportunities," *IEEE Wireless Commun.*, vol. 30, no. 4, pp. 132–138, Aug. 2023.



the IEEE/CIC International Conference on Communications in China (ICCC) in 2022.

Kangjian Chen (Graduate Student Member, IEEE) received the B.S. degree from the Nanjing University of Science and Technology, Nanjing, China, in 2017, and the M.S. degree in signal processing from Southeast University, Nanjing, in 2020, where he is currently pursuing the Ph.D. degree. His current research interests include integrated sensing and communication (ISAC), massive MIMO, and reconfigurable intelligent surface (RIS). He received the Best Paper Awards from the IEEE Global Communications Conference (GLOBECOM) in 2019 and



currently a Professor and the Head of the Jiangsu Multimedia Communication and Sensing Technology Research Center. He received the Best Paper Awards from IEEE GLOBECOM in 2019, IEEE/CIC ICC in 2022, and the 11th International Conference on Wireless Communications and Signal Processing (WCSP) in 2019. He has served as an Associate Editor for IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE COMMUNICATIONS LETTERS, IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY, IEEE OPEN JOURNAL OF VEHICULAR TECHNOLOGY, and *China Communications*.

Chenhao Qi (Senior Member, IEEE) received the B.S. degree (Hons.) in information engineering from the Chien-Shiung Wu Honored College, Southeast University, China, in 2004, and the Ph.D. degree in signal and information processing from Southeast University in 2010.

From 2008 to 2010, he visited the Department of Electrical Engineering, Columbia University, New York, NY, USA. Since 2010, he has been a Faculty Member with the School of Information Science and Engineering, Southeast University, where he is



Octavia A. Dobre (Fellow, IEEE) was a Visiting Professor with the Massachusetts Institute of Technology, USA, and Université de Bretagne Occidentale, France. She is currently a Professor and the Canada Research Chair Tier 1. Her research interests encompass wireless communication and networking technologies, and optical and underwater communications. She has (co)authored more than 450 refereed articles in these areas.

She is an Elected Member of the European Academy of Sciences and Arts and a fellow of the Engineering Institute of Canada and the Canadian Academy of Engineering. She received the Best Paper Awards from various conferences, including IEEE ICC, IEEE GLOBECOM, IEEE WCNC, and IEEE PIMRC. She serves as the VP of Publications of the Communications Society. She was the Director of Journals, the inaugural Editor-in-Chief (EiC) of IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY, and the EiC of IEEE COMMUNICATIONS LETTERS. She was a Fulbright Scholar, a Royal Society Scholar, and a Distinguished Lecturer of the IEEE Communications Society.



Geoffrey Ye Li (Fellow, IEEE) is currently a Chair Professor with the Imperial College London, U.K. Before joining Imperial College London in 2020, he was a Professor with the Georgia Institute of Technology, USA, for 20 years, and a Principal Technical Staff Member with AT&T Laboratories–Research (previous Bell Laboratories), NJ, USA, for five years. He made fundamental contributions to orthogonal frequency division multiplexing (OFDM) for wireless communications, established a framework for resource cooperation in

wireless networks, and introduced deep learning to communications. In these areas, he has published more than 600 journal and conference papers in addition to more than 40 granted patents. His publications have been cited more than 66,000 times with an H-index of 117. He has been listed as a Highly Cited Researcher by Clarivate/Web of Science almost every year.

He was elected an IEEE Fellow and an IET Fellow for his contributions to signal processing for wireless communications. He received the 2024 IEEE Eric E. Sumner Award, the 2019 IEEE ComSoc Edwin Howard Armstrong Achievement Award, and several awards from the IEEE Signal Processing, Vehicular Technology, and Communications Societies.