Codebook Design and Beam Training for Extremely Large-Scale RIS: Far-Field or Near-Field?

Xiuhong Wei[1], Linglong Dai [1,*], Yajun Zhao[2], Guanghui Yu[2], Xiangyang Duan[2]

[1] Beijing National Research Center for Information Science and Technology (BNRist) as well as the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China

[2] ZTE Corporation, Shenzhen 518038, China

[*] The corresponding author, email: daill@tsinghua.edu.cn

Abstract: Reconfigurable intelligent surface (RIS) is more likely to develop into extremely large-scale RIS (XL-RIS) to efficiently boost the system capacity for future 6G communications. Beam training is an effective way to acquire channel state information (CSI) for XL-RIS. Existing beam training schemes rely on the far-field codebook. However, due to the large aperture of XL-RIS, the scatters are more likely to be in the near-field region of XL-RIS. The far-field codebook mismatches the near-field channel model. Thus, the existing far-field beam training scheme will cause severe performance loss in the XL-RIS assisted near-field communications. To solve this problem, we propose the efficient near-field beam training schemes by designing the near-field codebook to match the near-field channel model. Specifically, we firstly design the near-field codebook by considering the near-field cascaded array steering vector of XL-RIS. Then, the optimal codeword for XL-RIS is obtained by the exhausted training procedure. To reduce the beam training overhead, we further design a hierarchical near-field codebook and propose the corresponding hierarchical near-field beam training scheme, where different levels of sub-codebooks are searched in turn with reduced codebook size. Simulation results show the proposed near-field beam training schemes outperform the existing far-field beam training scheme.

Keywords: extremely large-scale RIS; near-field codebook design; beam training

Received: Sep. 30, 2021
Revised: Oct. 31, 2021
Editor: Honggang Zhang

I. INTRODUCTION

Recently, reconfigurable intelligent surface (RIS) has been proposed to improve the capacity of the wireless communication system [1, 2]. Specifically, RIS consists of a large number of reconfigurable elements (e.g., 256), which can be deployed between the base station (BS) and the user to establish an extra reflecting link. By properly reconfiguring the RIS elements, RIS can provide high reflecting beamforming gain with low cost and low consumption [3, 4]. The reliable RIS reflecting beamforming requires the accurate channel state information (CSI) [5, 6]. However, due to a large number of RIS elements, CSI acquisition is challenging for the RIS assisted system [7].

There are two typical categories of CSI acquisition methods, which are respectively the explicit CSI acquisition (i.e., channel estimation) and the implicit CSI acquisition (i.e., beam training). For the first category, the BS sends pilot signals to the user through the RIS, and the user directly estimates the channel based on the received pilot signals [8]. Since each RIS element is usually passive, only the cascaded channel, i.e., the cascade of the channel from the BS to the RIS and the channel from the RIS to the user, can be estimated by least squares (LS) [9] or minimum mean square error (MMSE) algorithm [10]. The high-dimensional cascaded channel estimation will lead to the unaffordable pilot overhead in RIS assisted systems. To solve this problem, two types of low-overhead cascaded channel

estimation schemes have been proposed [8]. On the one hand, some compressive sensing (CS) algorithms can be used to estimate the high-dimensional cascaded channel by leveraging the sparsity of the angular cascaded channel [11, 12]. On the other hand, the multi-user correlation is exploited to reduce the pilot overhead by considering that all users communicate with the BS via the same RIS [13]. However, for this category of method, since RIS cannot perform the reliable reflecting beamforming before channel estimation, the received signal-to-noise ratio (SNR) is usually low. It is difficult for channel estimation to achieve the satisfactory channel estimation accuracy with low received SNR.

The second category is beam training, where CSI can be obtained by estimating the physical directions of channel paths instead of the entire channel. This beam training method has been widely considered in the existing 5G system, especially for millimeter-wave communications [14, 15]. Specifically, the BS and the user perform the training procedure through multiple directional beams (codewords) predefined in the codebook to search the optimal directional beam. After beam training, the physical directions of channel paths can be effectively obtained [16]. Compared with channel estimation, beam training can directly achieve the reliable beamforming by the training procedure, which can avoid the estimation of the entire channel with low received SNR. Recently, the beam training method has been extended to the RIS assisted systems for CSI acquisition [17–19]. The basic idea is that, based on the cascaded array steering vector of the RIS cascaded channel, the codebook consisting of multiple RIS directional beams is firstly designed, and then the training procedure between the RIS and the user is performed to search the optimal RIS directional beam [17]. By considering the fact that the cascaded array steering vector is mainly determined by the angle differences at the RIS, the partial search based beam training scheme was further proposed to reduce the search complexity [18]. Moreover, inspired by the powerful learning ability of neural networks, a deep reinforcement learning based beam training scheme was proposed to reduce the beam training overhead [19].

However, the existing codebook and beam training schemes may not be applicable any more with the increasing number of RIS elements. Specifically, the RIS assisted system is faced with the "multiplicative fading" effect [20–23], where the equivalent path loss of the BS-RIS-user reflecting link is the product of (instead of the sum of) the path losses of the BS-RIS link and RIS-user link. Thanks to low cost and low power consumption, more and more RIS elements are expected to deploy to compensate for the severe path loss to ensure the capacity improvement [21]. Thus, RIS is more likely to develop into extremely large-scale RIS (XL-RIS) for future 6G communications, which will lead to the fundamental transformation of electromagnetic radiation field structure [24]. The electromagnetic radiation field can be divided into the far-field region and near-field region [25], where the far-field channel model and near-field channel model should be used, respectively. The boundary of these two regions is determined by the Rayleigh distance [25], which is proportional to the square of the array aperture. In the RIS assisted systems, if the array aperture of RIS is not very large, the Rayleigh distance is small, so the scatters are usually assumed in the far-field region of RIS. Correspondingly, existing RIS codebook for beam training is designed based on the far-field channel model [17–19]. With the increasing number of RIS elements from RIS to XL-RIS (e.g., from 256 to 1024), the array aperture of XL-RIS becomes very large, and the Rayleigh distance increases accordingly [24]. As a result, the scatters are more likely to be in the near-field region of XL-RIS, and the near-field channel model should be considered in the XL-RIS assisted system [26]. However, the existing far-field codebook mismatches the near-field channel model, and the corresponding far-field beam training will cause severe performance loss in the XL-RIS assisted near-field communications. Unfortunately, this important new problem has not been studied in the literature.

To fill in this gap, in this paper, we propose the efficient near-field beam training schemes for XL-RIS by designing the near-field codebook to match the near-field channel model. Our contributions are summarized as follows.

1. We design the near-field codebook to match the near-field channel model, and then propose the corresponding near-field beam training scheme for XL-RIS. Specifically, by considering the near-field cascaded array steering vector of the XL-RIS cascaded channel, the near-field codebook is

firstly designed, where each codeword is determined by a pair of sampled points in the x-y-z coordinate system. Then, the optimal codeword for XL-RIS is obtained by the exhausted training procedure between the XL-RIS and the user.

2. In order to reduce the beam training overhead, we further design a hierarchical near-field codebook and propose the corresponding hierarchical near-field beam training scheme for XL-RIS. Compared with the near-field codebook, the hierarchical near-field codebook consists of several different levels of sub-codebooks, which are determined by different sampling ranges and sampling steps. During the beam training process, we search from the first level sub-codebook to the last level sub-codebook in turn, where the sampling ranges and sampling steps gradually become smaller. Finally, the globally optimal codeword can be achieved in the last level sub-codebook associated with the minimum sampling ranges and sampling steps.

The rest of the paper is organized as follows. In Section II, we firstly introduce the signal model, and then review the existing far-field channel model and far-field codebook. The near-field channel model for the XL-RIS assisted system is also presented in Section II. In Section III, the near-field codebook is designed and the corresponding near-field beam training scheme is proposed, and then the hierarchical near-field codebook based beam training is further proposed to reduce the beam training overhead. Simulation results and conclusions are provided in Section IV and Section V, respectively.

Notation: Lower-case and upper-case boldface letters $\mathbf{a}$ and $\mathbf{A}$ denote a vector and a matrix, respectively; $\mathbf{a}^*$ and $\mathbf{a}^H$ denote the conjugate and conjugate transpose of vector $\mathbf{a}$, respectively; $\|\mathbf{a}\|_2$ denotes the l_2 norm of vector $\mathbf{a}$; $\mathbf{A}^H$ denotes the conjugate transpose of matrix $\mathbf{A}$. $\mathcal{CN}(\mu, \sigma)$ denotes the probability density function of the circularly symmetric complex Gaussian distribution with mean μ and variance σ^2. Finally, $\mathrm{diag}(\mathbf{a})$ denotes the diagonal matrix with the vector $\mathbf{a}$ on its diagonal.

II. SYSTEM MODEL

In this section, we will first introduce the signal model of the XL-RIS assisted communication system. Then, the existing far-field channel model and the far-field codebook for beam training will be briefly reviewed. Finally, the near-field channel model for XL-RIS is presented.

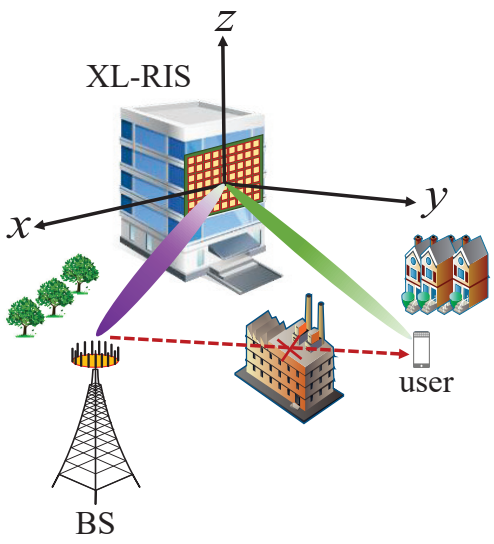

Figure 1. *The XL-RIS assisted wireless communication system.*

2.1 Signal Model

As shown in Figure 1, a XL-RIS is deployed between the BS with M-element antenna array and a single-antenna user to provide a reflecting link to assist communication, where the direct link between the BS and the user is blocked by obstacles [17, 27]. The XL-RIS consisting of $N = N_1 \times N_2$ elements is placed in the x-z plane, where the center of the XL-RIS is at the origin of the x-y-z coordinate system.

Let $\mathbf{G} \in \mathbb{C}^{N\times M}$ denote the channel from the BS to the XL-RIS and $\mathbf{h}_r \in \mathbb{C}^{1\times N}$ denote the channel from the XL-RIS to the user. By considering the downlink transmission, the received signal r at the user can be expressed by

$$r = \mathbf{h}_r \mathrm{diag}(\boldsymbol{\theta}) \mathbf{G}\mathbf{v}s + n, \tag{1}$$

where $\boldsymbol{\theta} = [\theta_1, \cdots, \theta_N] \in \mathbb{C}^{N\times 1}$ is the reflecting beamforming vector at the XL-RIS with θ_n representing the reflecting coefficient at the nth RIS element $(n = 1, \cdots, N)$, $\mathbf{v} \in \mathbb{C}^{M\times 1}$ represents the beamforming vector at the BS, s represents the symbol transmitted by the BS, $n \sim \mathcal{CN}(0, \sigma^2)$ represents

the received noise at the user with σ^2 representing the noise power.

To design the effective beamforming vectors $\mathbf{v}$ and $\boldsymbol{\theta}$, it is necessary to acquire the accurate CSI [5, 6]. Since there are extremely large number of RIS elements, channel estimation cannot achieve the satisfactory channel estimation with low pilot overhead. By contrast, beam training is a more effective way to acquire CSI [17, 19]. Specifically, since the XL-RIS is generally dominated by the main path (or a few paths), we only need to search the physical direction of the main path by beam training instead of explicitly estimating the entire channel. Thus, in the following description, only the main path is concerned, and the corresponding beam training method will be investigated to search the optimal directional beam to align with the main path.

Moreover, since the BS and the XL-RIS are generally deployed at fixed positions, the channel $\mathbf{G}$ from the BS to the XL-RIS has a much longer channel coherence time than the channel $\mathbf{h}_r$ from the XL-RIS to the user due to the user's mobility [17]. For simplicity, we assume that the beamforming vector $\mathbf{v}$ at the BS has been aligned with the main path of the channel $\mathbf{G}$ [17]. In this paper, we only focus on the beam training at the XL-RIS. Next, we will briefly review the existing far-field channel model and the corresponding far-field codebook for beam training.

2.2 Far-Field Channel Model and Far-Filed Codebook

Based on the far-field channel model, $\mathbf{G}$ and $\mathbf{h}_r$ can be respectively represented by

$$\mathbf{G}_{\text{far-field}} = \alpha_G \mathbf{a}\left(\phi_{G_r}, \psi_{G_r}\right) \mathbf{b}^T\left(\phi_{G_t}, \psi_{G_t}\right), \quad (2)$$

$$\mathbf{h}^r_{\text{far-field}} = \alpha_r \mathbf{a}^T\left(\phi_r, \psi_r\right), \quad (3)$$

where α_G and α_r represent the path gains, ϕ_{G_r} and ψ_{G_r} represent the spatial angles at the XL-RIS for the channel $\mathbf{G}$, ϕ_{G_t} and ψ_{G_t} represent the spatial angles at the BS for the channel $\mathbf{G}$, ϕ_r and ψ_r represent the spatial angles at the XL-RIS for the channel $\mathbf{h}_r$. $\mathbf{a}(\phi, \psi)$ and $\mathbf{b}(\phi, \psi)$ represent the far-field array steering vector associated to the XL-RIS and the BS, respectively. Take $\mathbf{a}(\phi, \psi)$ as an example, it can be expressed as [12]

$$\mathbf{a}\left(\phi, \psi\right) = \left[e^{-j2\pi\phi\mathbf{n}_1}\right] \otimes \left[e^{-j2\pi\psi\mathbf{n}_2}\right], \quad (4)$$

where $\mathbf{n}_1 = [0, 1, \cdots, N_1 - 1]^T$ and $\mathbf{n}_2 = [0, 1, \cdots, N_2 - 1]^T$. $\phi = d_f \sin\left(\vartheta\right) \cos\left(\upsilon\right) / \lambda$ and $\psi = d_f \sin\left(\upsilon\right) / \lambda$, where ϑ and υ respectively represent the physical angles in the azimuth and elevation, λ is the carrier wavelength, and d_f is the element spacing satisfying $d_f = \lambda/2$.

By considering that the beamforming vector $\mathbf{v}$ at the BS has been designed, i.e., $\mathbf{v} = \frac{\mathbf{b}^*}{M}$, the receiver signal r in (1) can be further represented by

$$\begin{aligned} r =& \alpha \mathbf{a}^T\left(\phi_r, \psi_r\right) \operatorname{diag}\left(\boldsymbol{\theta}\right) \mathbf{a}\left(\phi_{G_r}, \psi_{G_r}\right) \bar{s} + n \\ =& \alpha \boldsymbol{\theta}^T \operatorname{diag}\left(\mathbf{a}\left(\phi_r, \psi_r\right)\right) \mathbf{a}\left(\phi_{G_r}, \psi_{G_r}\right) \bar{s} + n \\ =& \alpha \boldsymbol{\theta}^T \mathbf{a}\left(\phi_{G_r} + \phi_r, \psi_{G_r} + \psi_r\right) \bar{s} + n \\ =& \boldsymbol{\theta}^T \bar{\mathbf{h}}_{\text{far-field}} \bar{s} + n, \end{aligned} \quad (5)$$

where $\bar{\mathbf{h}}_{\text{far-field}} = \alpha \mathbf{a}\left(\phi_{G_r} + \phi_r, \psi_{G_r} + \psi_r\right)$ denotes the far-field cascaded channel, $\alpha = \alpha_G \alpha_r$ denotes the effective gain of $\bar{\mathbf{h}}_{\text{far-field}}$, and $\bar{s} = \mathbf{b}^T\left(\phi_{G_t}, \psi_{G_t}\right) \mathbf{v} s$ denotes the effective transmitted symbol.

For beam training, the entire procedure can be divided into multiple time slots. In different time slots, the reflecting beamforming vector $\boldsymbol{\theta}$ is set as different codewords in the predefined codebook, which will equivalently produce different directional beams. For each codeword, the user will measure the strength of the received signal r and feedback the optimal codeword index. Based on the far-field array steering vector in (4), the existing far-field codebook $\mathbf{F}$ is generally designed as [17]

$$\begin{aligned} \mathbf{F} = [&\mathbf{a}\left(\phi_1, \psi_1\right), \cdots, \mathbf{a}\left(\phi_1, \psi_{N_2}\right), \cdots, \\ &\mathbf{a}\left(\phi_{N_1}, \psi_1\right), \cdots, \mathbf{a}\left(\phi_{N_1}, \psi_{N_2}\right)]^*, \end{aligned} \quad (6)$$

where $\phi_n = \frac{2n - N_1 - 1}{N_1}$ with $n = 1, 2, \cdots, N_1$ and $\psi_n = \frac{2n - N_2 - 1}{N_2}$ for $n = 1, 2, \cdots, N_2$. Each column of $\mathbf{F}$ represents a codeword for $\boldsymbol{\theta}$.

The existing beam training schemes are mainly based on the above far-field codebook [17–19]. However, when the RIS develops into XL-RIS, the far-field codebook and beam training may not be applicable any more, which will be explained in detail in the next Section 2.3.

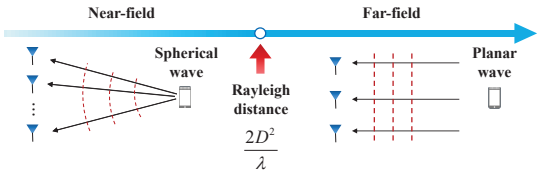

Figure 2. *The near-field region and the far-field region [24].*

2.3 Near-Field Channel Model

Specifically, as shown in Figure 2, the electromagnetic radiation field in wireless communications can be divided into far-field region and near-field region [25], where different fields will result in different channel models. The bound between these two fields is determined by the Rayleigh distance $Z = \frac{2D^2}{\lambda}$, where D represents the array aperture. In the conventional RIS assisted system, since the array aperture of the RIS is not too large, the corresponding Rayleigh distance is small. The scatters are in the far-field region of RIS [5, 17, 27], where the RIS channel can be modeled under the planar wave assumption, as described in (2) and (3). With the increase of the array aperture in the XL-RIS assisted system, the corresponding Rayleigh distance also increases. For example, we consider that the carrier frequency is 30 GHz and the corresponding carrier wavelength is $\lambda = 0.01$ meters. When the array aperture for RIS is $D = 0.1$ meters, the Rayleigh distance is only $Z = 2$ meters. When the array aperture for XL-RIS increases to $D = 1$ meter, the Rayleigh distance can reach $Z = 200$ meters. Thus, in the XL-RIS assisted system, the scatters are more likely to be in the near-field region, where the XL-RIS channel should be modeled under the spherical wave assumption. Next, we will introduce the near-field channel model of XL-RIS.

For convenience of description, the distances and coordinates are normalized by the carrier wavelength λ in the following of this paper [28]. The vertical or horizontal distance between two adjacent RIS elements is set to d. The coordinate of the RIS element can be represented as $(x_{n_1}, 0, z_{n_2})$ with $x_{n_1} = (n_1 - \frac{N_1+1}{2})d$ and $z_{n_2} = (n_2 - \frac{N_2+1}{2})d$ in the x-y-z coordinate system, where $n_1 = 1, \cdots, N_1$ and $n_2 = 1, \cdots, N_2$.

The near-field channel $\mathbf{h}^r_{\text{near-field}}$ from the XL-RIS to the user can be represented by [28]

$$\mathbf{h}^r_{\text{near-field}} = \alpha_r \mathbf{c}^T (x_r, y_r, z_r), \tag{7}$$

where (x_r, y_r, z_r) represents the coordinate of the scatter corresponding to the main path between the XL-RIS and user. Compared with the far-field channel model in (3), the array steering vector $\mathbf{c}(x_r, y_r, z_r)$ for the near-field channel model is derived based on the the spherical wave assumption, which can be represented by [28]

$$\mathbf{c}(x_r, y_r, z_r) = \Big[e^{-j2\pi D^r(1,1)}, \cdots, e^{-j2\pi D^r(1,N_2)}, \cdots, e^{-j2\pi D^r(N_1,1)}, \cdots, e^{-j2\pi D^r(N_1,N_2)}\Big]^T, \tag{8}$$

where $D^r(n_1, n_2) = \sqrt{(x_r - x_{n_1})^2 + y_r^2 + (z_r - z_{n_2})^2}$, which represents the distance from the (n_1, n_2)-th RIS element to (x_r, y_r, z_r).

Similarly, the array steering vector at the RIS of the channel $\mathbf{G}$ from the BS to the XL-RIS should be also near-field. Thus, the near-field cascaded channel $\bar{\mathbf{h}}_{\text{near-field}}$ can be represented by

$$\bar{\mathbf{h}}_{\text{near-field}} = \alpha \bar{\mathbf{c}}\left((x_{G_r}, y_{G_r}, z_{G_r}), (x_r, y_r, z_r)\right), \tag{9}$$

where $(x_{G_r}, y_{G_r}, z_{G_r})$ represents the coordinate of the scatter corresponding to the main path between the BS and XL-RIS. The near-field cascaded array steering vector $\bar{\mathbf{c}}\left((x_{G_r}, y_{G_r}, z_{G_r}), (x_r, y_r, z_r)\right)$ can be represented by

$$\bar{\mathbf{c}}\left((x_{G_r}, y_{G_r}, z_{G_r}), (x_r, y_r, z_r)\right) = \Big[e^{-j2\pi D(1,1)}, \cdots, e^{-j2\pi D(1,N_2)}, \cdots, e^{-j2\pi D(N_1,1)}, \cdots, e^{-j2\pi D(N_1,N_2)}\Big]^T, \tag{10}$$

where $D(n_1, n_2) = D^{G_r}(n_1, n_2) + D^r(n_1, n_2)$, represents the effective distance of $\bar{\mathbf{h}}_{\text{near-field}}$, and $D^{G_r}(n_1, n_2) = \sqrt{(x_{G_r} - x_{n_1})^2 + y_{G_r}^2 + (z_{G_r} - z_{n_2})^2}$. It is noted that $(x_{G_r}, y_{G_r}, z_{G_r})$ satisfies $X^{G_r}_{\min} \le x_{G_r} \le X^{G_r}_{\max}$, $Y^{G_r}_{\min} \le y_{G_r} \le Y^{G_r}_{\max}$, $Z^{G_r}_{\min} \le z_{G_r} \le Z^{G_r}_{\max}$, and (x_r, y_r, z_r) satisfies $X^r_{\min} \le x_r \le X^r_{\max}$,$Y^r_{\min} \le y_r \le Y^r_{\max}$, and $Z^r_{\min} \le z_r \le Z^r_{\max}$.

Compared with the far-field cascaded array steering vector only associated with the angles in (4), the near-field cascaded array steering vector is determined by a pair of points in the x-y-z coordinate system, i.e., $(x_{G_r}, y_{G_r}, z_{G_r})$ and (x_r, y_r, z_r). The existing far-field codebook mismatches the near-field channel model. Thus, the corresponding far-field beam training will cause severe performance loss in the XL-RIS assisted near-field communications. In this paper, we design the near-field codebook to match the near-field channel model, and then propose the corresponding near-field beam training for XL-RIS, which will be introduced in the next Section III.

III. PROPOSED NEAR-FIELD CODEBOOK DESIGN AND BEAM TRAINING SCHEME FOR XL-RIS

In this section, we will introduce the proposed near-field codebook and the corresponding near-field beam training. Then, a hierarchical near-field codebook and the corresponding beam training will be further proposed to reduce the beam training overhead.

3.1 Proposed Near-Field Codebook Design and Beam Training

Inspired by the near-field dictionary matrix design for the XL-MIMO channel estimation [28], we design a near-field codebook for the XL-RIS beam training. In the near-field dictionary matrix proposed in [28] for the extremely large-scale linear antenna array, the considered entire two-dimensional (2D) plane is divided by several sampled points in the x-y coordinate system. Each column of the dictionary matrix is the corresponding near-field array steering vector for the linear array associated with one sampled point. If the near-field codebook for the planar array is required to design in the XL-MIMO system, we only need to extend the 2D plane to the three-dimensional (3D) space, where each codeword can be generated based on the near-field array steering vector for the planar array associated with one sampled point in the x-y-z coordinate system. However, the near-field cascaded channel brings new challenges to the codebook design for XL-RIS.

Specifically, since the near-field cascaded array steering vector $\bar{\mathbf{c}}$ of the near-field cascaded channel $\bar{\mathbf{h}}_{\text{near-field}}$ is determined by the sum of the distance from $(x_{G_r}, y_{G_r}, z_{G_r})$ to the XL-RIS and the distance from (x_r, y_r, z_r) to the XL-RIS, each codeword for XL-RIS should be related to a pair of sampled points in the x-y-z coordinate system, instead of only one sampled point [28]. Next, we will introduce the designed near-field codebook based on the near-field cascaded array steering vector.

3.1.1 Near-Field Codebook Design

Let Ξ^{G_r} and Ξ^{r} denote the two collections of sampled points corresponding to $(x_{G_r}, y_{G_r}, z_{G_r})$ and (x_r, y_r, z_r), which can be represented as

$$\begin{aligned}\Xi^{G_r} = \big\{&(x_s^{G_r}, y_s^{G_r}, z_s^{G_r})| \\ x_s^{G_r} &= X_{\min}^{G_r}, X_{\min}^{G_r} + \Delta x^{G_r}, \cdots, X_{\max}^{G_r}; \\ y_s^{G_r} &= Y_{\min}^{G_r}, Y_{\min}^{G_r} + \Delta y^{G_r}, \cdots, Y_{\max}^{G_r}; \\ z_s^{G_r} &= Z_{\min}^{G_r}, Z_{\min}^{G_r} + \Delta z^{G_r}, \cdots, Z_{\max}^{G_r}\big\},\end{aligned} \tag{11}$$

$$\begin{aligned}\Xi^{r} = \big\{&(x_s^{r}, y_s^{r}, z_s^{r})| \\ x_s^{r} &= X_{\min}^{r}, X_{\min}^{r} + \Delta x^{r}, \cdots, X_{\max}^{r}; \\ y_s^{r} &= Y_{\min}^{r}, Y_{\min}^{r} + \Delta y^{r}, \cdots, Y_{\max}^{r}; \\ z_s^{r} &= Z_{\min}^{r}, Z_{\min}^{r} + \Delta z^{r}, \cdots, Z_{\max}^{r}\big\},\end{aligned} \tag{12}$$

where Δx^{G_r}, Δy^{G_r} and Δz^{G_r} represent the sampling step on the x-axis, y-axis and z-axis for Ξ^{G_r}, respectively. Δx^{r}, Δy^{r} and Δz^{r} represent the sampling step on the x-axis, y-axis and z-axis for Ξ^{r}, respectively. Let $\Delta = [\Delta x^{G_r}, \Delta y^{G_r}, \Delta z^{G_r}, \Delta x^{r}, \Delta y^{r}, \Delta z^{r}]$ denote all sampling steps. Given a pair of sampled points $(x_s^{G_r}, y_s^{G_r}, z_s^{G_r})$ and $(x_s^{r}, y_s^{r}, z_s^{r})$, the effective sampled distance $D_s(n_1, n_2)$ can be presented by

$$D_s(n_1, n_2) = D_s^{G_r}(n_1, n_2) + D_s^{r}(n_1, n_2), \tag{13}$$

where $D_s^{G_r}(n_1, n_2) = \sqrt{(x_s^{G_r} - x_{n_1})^2 + {y_s^{G_r}}^2 + (z_s^{G_r} - z_{n_2})^2}$ and $D_s^{r}(n_1, n_2) = \sqrt{(x_s^{r} - x_{n_1})^2 + y_s^{r2} + (z_s^{r} - z_{n_2})^2}$.

Algorithm 1 shows the specific near-field codebook design procedure. From Step 3, we can find that the near-field codeword is generated based on the near-field cascaded array steering vector, which is related to a pair of sampled points $(x_s^{G_r}, y_s^{G_r}, z_s^{G_r})$ and $(x_s^{r}, y_s^{r}, z_s^{r})$. It is noted that different pairs of sampled points may produce the same effective sampled distance, which will result in the same codeword. In order

Algorithm 1. *Near-field codebook design.*

Inputs: The two collections of sampled points Ξ^{G_r} and Ξ^r, the number of RIS elements N_1 and N_2.

Initialization: $\mathbf{W} = \emptyset$, $L = 0$.

1. **for** $(x_s^{G_r}, y_s^{G_r}, z_s^{G_r}) \in \Xi^{G_r}$ **do**
2. **for** $(x_s^r, y_s^r, z_s^r) \in \Xi^r$ **do**
3. $\bar{\mathbf{c}}_s = \left[e^{-j2\pi D_s(1,1)}, \cdots, e^{-j2\pi D_s(1,N_2)}, \cdots, e^{-j2\pi D_s(N_1,1)}, \cdots, e^{-j2\pi D_s(N_1,N_2)}\right]^H$
4. **if** $\bar{\mathbf{c}}_s \notin \mathbf{W}$ **then**
5. $\mathbf{W} = [\mathbf{W}, \bar{\mathbf{c}}_s]$
6. $L = L + 1$
7. **end if**
8. **end for**
9. **end for**

Output: The designed near-field XL-RIS codebook $\mathbf{W}$, and the codebook size L.

to solve this problem, we need to ensure that each new codeword is different from all previous codewords, as shown in Steps 4-7. Finally, the designed near-field codebook $\mathbf{W}$ is obtained, where each column represents one codeword for the reflecting beamforming vector $\boldsymbol{\theta}$ at the XL-RIS. The corresponding codebook size L is also obtained.

After designing the near-field codebook for XL-RIS, the beam training procedure between the XL-RIS and the user can be performed to search the optimal codeword for the reflecting beamforming vector $\boldsymbol{\theta}$ at the XL-RIS. Next, we will introduce the corresponding near-field beam training scheme.

3.1.2 Near-Field Beam Training

The specific near-field beam training procedure is summarized in **Algorithm 2**, where all the codewords in the designed near-field codebook $\mathbf{W}$ need to be traversed. The entire training procedure can be divided into L time slots. In l-th time slots, the BS transmits the effective symbol $\bar{s}$ to the user, where the reflecting beamforming vector $\boldsymbol{\theta}_l$ is set as the l-th codeword in the designed near-field codebook $\mathbf{W}$ at the XL-RIS, as shown in Step 3. After L time slots, the user can obtain the optimal codeword based on all received signals $\{r_l\}_{l=1}^L$ with the help of Steps 4-7. Finally, the optimal codeword index l_{opt} is fed back from the user to XL-RIS.

Since the near-field cascaded array steering vector

Algorithm 2. *Near-field beam training.*

Inputs: The designed near-field XL-RIS codebook $\mathbf{W}$, and the effective transmitted symbol $\bar{s}$.

Initialization: $l = 0$, $|r|_{\text{opt}} = 0$, $l_{\text{opt}} = 0$.

1. **for** $\bar{\mathbf{c}}_s \in \mathbf{W}$ **do**
2. $l = l + 1$
3. $r_l = \boldsymbol{\theta}_l^T \bar{\mathbf{h}}_{\text{near-field}} \bar{s} + n_l$, where $\boldsymbol{\theta}_l = \bar{\mathbf{c}}_s$
4. **if** $|r_l| > |r|_{\text{opt}}$ **then**
5. $l_{\text{opt}} = l$
6. $|r|_{\text{opt}} = |r_l|$
7. **end if**
8. **end for**

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

for XL-RIS is jointly determined by a pair of sampled points, the codebook size L is usually large. This exhausted training procedure will lead to huge beaming training overhead. In order to reduce the beam training overhead, we further design a hierarchical near-field codebook and propose the corresponding hierarchical beam training scheme.

3.2 Proposed Hierarchical Near-Field Codebook Design and Beam Training

To reduce the beam training overhead, one effective way is to reduce the codebook size. By referring to (11) and (12), we can find the codebook size for the entire sampling space is mainly determined by the sampling steps on the x-axis, y-axis and z-axis. If the sampling steps are increased, the codebook size will be reduced. But the performance of the beam training will be also degraded accordingly, since it is difficult to accurately locate the scatters corresponding to the main paths with the reduced codebook. In order to solve this problem, we design a hierarchical near-field codebook, which consists of several different levels of sub-codebooks. These different levels of sub-codebooks are determined by different sampling ranges and sampling steps.

Specifically, let K denote the number of different levels of sub-codebooks. In the k-th sub-codebook ($k = 1, 2, \cdots, K$), the corresponding collections of

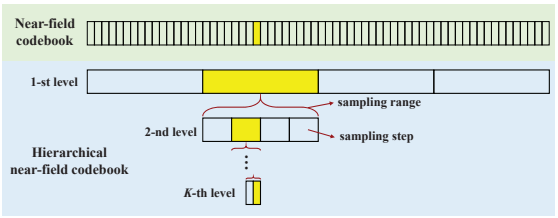

Figure 3. *Comparison between the near-field codebook and the hierarchical near-field codebook.*

sampled points $\Xi_k^{G_r}$ and Ξ_k^r can be defined as

$$\begin{aligned}\Xi_k^{G_r} = \big\{&(x_s^{G_r,k}, y_s^{G_r,k}, z_s^{G_r,k})| \\ &x_s^{G_r,k} = X_{\min}^{G_r,k}, X_{\min}^{G_r,k} + \Delta x^{G_r,k}, \cdots, X_{\max}^{G_r,k}; \\ &y_s^{G_r,k} = Y_{\min}^{G_r,k}, Y_{\min}^{G_r,k} + \Delta y^{G_r,k}, \cdots, Y_{\max}^{G_r,k}; \\ &z_s^{G_r,k} = Z_{\min}^{G_r,k}, Z_{\min}^{G_r,k} + \Delta z^{G_r,k}, \cdots, Z_{\max}^{G_r,k}\big\},\end{aligned} \tag{14}$$

$$\begin{aligned}\Xi_k^{r} = \big\{&(x_s^{r,k}, y_s^{r,k}, z_s^{r,k})| \\ &x_s^{r,k} = X_{\min}^{r,k}, X_{\min}^{r,k} + \Delta x^{r,k}, \cdots, X_{\max}^{r,k}; \\ &y_s^{r,k} = Y_{\min}^{r,k}, Y_{\min}^{r,k} + \Delta y^{r,k}, \cdots, Y_{\max}^{r,k}; \\ &z_s^{r,k} = Z_{\min}^{r,k}, Z_{\min}^{r,k} + \Delta z^{r,k}, \cdots, Z_{\max}^{r,k}\big\}.\end{aligned} \tag{15}$$

Take the sampling on the x-axis for $\Xi_k^{G_r}$ as an example, $[X_{\min}^{G_r,k}, X_{\max}^{G_r,k}]$ and $\Delta x^{G_r,k}$ represent the sampling range and sampling step, respectively. Further, let $R^k = \big\{[X_{\min}^{G_r,k}, X_{\max}^{G_r,k}], [Y_{\min}^{G_r,k}, Y_{\max}^{G_r,k}], [Z_{\min}^{G_r,k}, Z_{\max}^{G_r,k}], [X_{\min}^{r,k}, X_{\max}^{r,k}], [Y_{\min}^{r,k}, Y_{\max}^{r,k}], [Z_{\min}^{r,k}, Z_{\max}^{r,k}]\big\}$ and $\Delta^k = \{\Delta x^{G_r,k}, \Delta y^{G_r,k}, \Delta z^{G_r,k}, \Delta x^{r,k}, \Delta y^{r,k}, \Delta z^{r,k}\}$ respectively denote all sampling ranges and all sampling steps for the k-th level sub-codebook. In this way, $\Xi_k^{G_r}$ and Ξ_k^r can be completely determined by R^k and Δ^k. Given $\Xi_k^{G_r}$ and Ξ_k^r, the k-th level sub-codebook $\mathbf{W}_k$ and the corresponding codebook size L_k can be further obtained by referring to **Algorithm 1**. Figure 3 shows the comparison between the near-field codebook and the hierarchical near-field codebook. In the hierarchical near-field codebook, from the 1-st level sub-codebook to the K-th level sub-codebook, both the corresponding sampling ranges and sampling steps gradually become smaller. Thus, the codebook size of each level of sub-codeboook is not large.

Based on the hierarchical near-field codebook, we further propose the hierarchical near-field beam training scheme. The basic idea is to search from the 1-st level sub-codebook to the K-th level sub-codebook in turn, where the sampling ranges of the latter level sub-codebook is determined by the optimal codeword searched and the sampling steps by the former level sub-codebook, as shown in Figure 3. By assuming that the searched optimal codeword for the k-th level sub-codebook is $\bar{\mathbf{c}}_{s,k,\text{opt}}$ with the corresponding sampled points $(x_{s,k,\text{opt}}^{G_r}, y_{s,k,\text{opt}}^{G_r}, z_{s,k,\text{opt}}^{G_r})$ and $(x_{s,k,\text{opt}}^{r}, y_{s,k,\text{opt}}^{r}, z_{s,k,\text{opt}}^{r})$, the sampling ranges of the $(k+1)$-th level sub-codebook can be determined accordingly. Take $[X_{\min}^{G_r,k+1}, X_{\max}^{G_r,k+1}]$ as an example, it can be represented by

$$\begin{aligned}&[X_{\min}^{G_r,k+1}, X_{\max}^{G_r,k+1}] = \\ &[x_{s,k,\text{opt}}^{G_r} - \frac{\Delta x^{G_r,k}}{2}, x_{s,k,\text{opt}}^{G_r} + \frac{\Delta x^{G_r,k}}{2}].\end{aligned} \tag{16}$$

For the 1-st level sub-codebook, the corresponding sampling ranges can be set as $R^1 = \big\{[X_{\min}^{G_r}, X_{\max}^{G_r}], [Y_{\min}^{G_r}, Y_{\max}^{G_r}], [Z_{\min}^{G_r}, Z_{\max}^{G_r}], [X_{\min}^{r}, X_{\max}^{r}], [Y_{\min}^{r}, Y_{\max}^{r}], [Z_{\min}^{r,k}, Z_{\max}^{r,k}]\big\}$. The initial

Algorithm 3. *Hierarchical near-field beam training.*

Inputs: The number of different levels of sub-codebooks K, the initial sampling ranges R^1, the initial sampling steps Δ^1, the step control parameter δ, and the effective transmitted symbol $\bar{s}$, the number of RIS elements N_1 and N_2.

Initialization: $l_k = 0$, $|r|_{k,\text{opt}} = 0$ and $l_{k,\text{opt}} = 0$ for $\forall k$

1. **for** $k = 1, 2, \cdots, K$ **do**
2. generate $\Xi_k^{G_r}$ and Ξ_k^r based on R^k and Δ^k by (14) and (15)
3. generate $\mathbf{W}_k$ based on $\Xi_k^{G_r}$ and Ξ_k^r by **Algorithm 1**
4. **for** $\bar{\mathbf{c}}_{s,k} \in \mathbf{W}_k$ **do**
5. $l_k = l_k + 1$
6. $r_{l,k} = \boldsymbol{\theta}_{l,k}^T \bar{\mathbf{h}}_{\text{near-field}} \bar{s} + n_{l,k}$, where $\boldsymbol{\theta}_{l,k} = \bar{\mathbf{c}}_{s,k}$
7. **if** $|r_{l,k}| > |r|_{k,\text{opt}}$ **then**
8. $l_{k,\text{opt}} = l_k$
9. $|r|_{k,\text{opt}} = |r_{l,k}|$
10. **end if**
11. **end for**
12. generate R^{k+1} based on $l_{k,\text{opt}}$ by (16)
13. $\Delta^{k+1} = \delta \Delta^k$
14. **end for**

Output: The feedback optimal codeword index $l_{K,\text{opt}}$ from the user.

sampling steps Δ^1 should be set as bigger values to reduce the codebook size. In this paper, we set $\Delta^1 = A\Delta$, where Δ is the sampling steps for the designed near-field codebook in Section 3.1, and A is a scalar greater than 1. Moreover, we define a step control parameter δ ($0 < \delta < 1$) to gradually decrease the sampling steps of the sub-codebooks.

The specific hierarchical near-field beam training procedure is summarized in **Algorithm 3**, where the entire beam training procedure is divided into K stages. In the k-th stage, the k-th level sub-codebook $\mathbf{W}_k$ is firstly generated by Steps 2-3. Then, Steps 4-11 are performed to search the optimal codeword in $\mathbf{W}_k$. It is noted that the searched optimal codeword index $l_{k,\text{opt}}$ should be fed back from the user to the XL-RIS to generate the sampling ranges R^{k+1} and sampling steps Δ^{k+1} for the $(k+1)$-th level sub-codebook, as shown in Steps 12-13. Finally, the optimal codeword for the K-th level sub-codebook is regarded as the searched globally optimal codeword in the hierarchical near-field codebook. It is noted that when the user's mobility is considered, the proposed beam training scheme can be extended to the beam tracking scheme, where the proposed near-field codebook is still effective.

Compared with the beam training overhead L of the near-field beam training scheme, the beam training overhead of the hierarchical near-field beam training scheme is $\sum_{k=1}^{K} L_k$. Specifically, L is determined by $\left[\left\lfloor\frac{(X_{\max}^{G_r}-X_{\min}^{G_r})}{\Delta x^{G_r}}\right\rfloor+1\right]\cdot\left[\left\lfloor\frac{(Y_{\max}^{G_r}-Y_{\min}^{G_r})}{\Delta y^{G_r}}\right\rfloor+1\right]\cdot\left[\left\lfloor\frac{(Z_{\max}^{G_r}-Z_{\min}^{G_r})}{\Delta z^{G_r}}\right\rfloor+1\right]\cdot\left[\left\lfloor\frac{(X_{\max}^{r}-X_{\min}^{r})}{\Delta x^{r}}\right\rfloor+1\right]\cdot\left[\left\lfloor\frac{(Y_{\max}^{r}-Y_{\min}^{r})}{\Delta y^{r}}\right\rfloor+1\right]\cdot\left[\left\lfloor\frac{(Z_{\max}^{r}-Z_{\min}^{r})}{\Delta z^{r}}\right\rfloor+1\right]$, where $\lfloor x\rfloor$ denotes the maximum integer not greater than x. L_1 is determined by $\left[\left\lfloor\frac{(X_{\max}^{G_r}-X_{\min}^{G_r})}{A\Delta x^{G_r}}\right\rfloor+1\right]\cdot\left[\left\lfloor\frac{(Y_{\max}^{G_r}-Y_{\min}^{G_r})}{A\Delta y^{G_r}}\right\rfloor+1\right]\cdot\left[\left\lfloor\frac{(Z_{\max}^{G_r}-Z_{\min}^{G_r})}{A\Delta z^{G_r}}\right\rfloor+1\right]\cdot\left[\left\lfloor\frac{(X_{\max}^{r}-X_{\min}^{r})}{A\Delta x^{r}}\right\rfloor+1\right]\cdot\left[\left\lfloor\frac{(Y_{\max}^{r}-Y_{\min}^{r})}{A\Delta y^{r}}\right\rfloor+1\right]\cdot\left[\left\lfloor\frac{(Z_{\max}^{r}-Z_{\min}^{r})}{A\Delta z^{r}}\right\rfloor+1\right]$, and L_k is determined by $\left(\left\lfloor\frac{1}{\delta}\right\rfloor+1\right)^6$ for $k > 1$. It is noted that $A > 1$ and $0 < \delta < 1$. By carefully designing the sampling steps controlled by A and δ for the hierarchical near-field codebook, $\sum_{k=1}^{K} L_k$ can be much smaller than L.

IV. SIMULATION RESULTS

In this section, we provide the simulation results to verify the performance of the two proposed near-field beam training schemes.

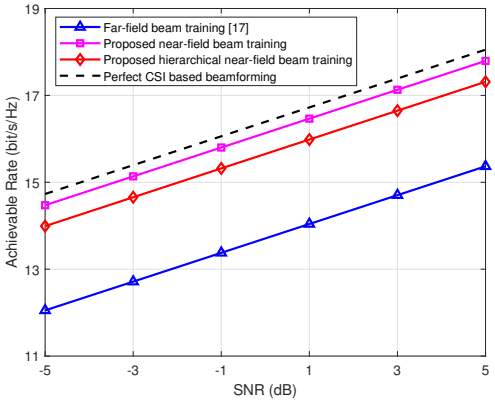

Figure 4. *Achievable rate performance comparison against the SNR.*

For simulations, we consider that the number of antennas at the BS is $M = 64$, and the number of RIS elements at the XL-RIS is $N = 512$ ($N_1 = 128$ and $N_2 = 4$). The path gains are generated by $\alpha_G \sim \mathcal{CN}(0,1)$ and $\alpha_r \sim \mathcal{CN}(0,1)$. The vertical or horizontal distance between two adjacent RIS elements is set as $d = 1/2$ [28]. The distance from the XL-RIS to the scatter is limited to $X_{\min}^{G_r} = X_{\min}^{r} = -1200d$, $X_{\max}^{G_r} = X_{\max}^{r} = 1200d$, $Y_{\min}^{G_r} = Y_{\min}^{r} = 10d$, $Y_{\max}^{G_r} = Y_{\max}^{r} = 200d$, $Z_{\min}^{G_r} = Z_{\min}^{r} = -400d$, $Z_{\max}^{G_r} = Z_{\max}^{r} = 400d$. The symbol transmitted by the BS is set as $s = 1$, and the beamforming vector at the BS is set as $\mathbf{v} = \frac{\mathbf{b}^*}{M}$. Thus, the effective transmitted symbol is $\bar{s} = 1$. The SNR is defined as $1/\sigma^2$.

We compare the proposed near-field beam training scheme and the hierarchical near-field beam training scheme with the existing far-field beam training scheme [17]. In the near-field beam training scheme, the sampling steps on all x-axis, y-axis and z-axis are set to the same, i.e., $\Delta x^{G_r} = \Delta y^{G_r} = \Delta z^{G_r} = \Delta x^{r} = \Delta y^{r} = \Delta z^{r} = \Delta_s$. That is to say $\Delta = [\Delta_s, \Delta_s, \Delta_s, \Delta_s, \Delta_s, \Delta_s]$. In the hierarchical near-field beam training scheme, we set $A = 4$, and the initial sampling steps $\Delta^1 = A\Delta$. The step control parameter is set as $\delta = 0.25$. The number of different levels of sub-codebooks is set as $K = 2$. In the far-field beam training scheme, the far-field codebook $\mathbf{F}$ defined in (6) is adopted. Moreover, we provide the beamforming scheme with perfect CSI as the upper bound of performance, i.e., $\boldsymbol{\theta} = \bar{\mathbf{c}}^*$.

Figure 4 shows the achievable rate performance

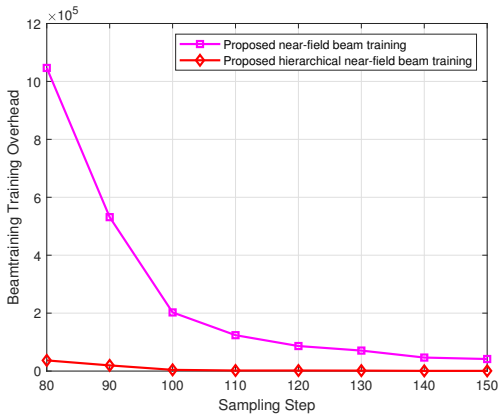

Figure 5. *Beam training overhead comparison against the sampling step* Δ_s.

comparison against the SNR, where $\Delta_s = 100d$. We can find that compared with the existing far-field beam training scheme, the two proposed near-field beam training schemes can achieve better achievable rate performance. Due to the error propagation among different levels of sub-codebooks search, the performance of the hierarchical near-field beam training scheme is slightly worse than that of the near-field beam training scheme. Specifically, the hierarchical near-field beam training scheme can achieve about 95% achievable rate performance of the near-field beam training scheme.

Figure 5 further shows the beam training overhead comparison of the two proposed near-field beam training schemes against the sampling step Δ_s. We can find that the hierarchical near-field beam training scheme can greatly reduce the beam training overhead. When $\Delta_s = 100d$, the beam training overhead of the near-field beam training scheme and the hierarchical near-field beam training scheme are respectively 147628 and 15927, where the latter is only about 10% of the former.

V. CONCLUSION

In this paper, we have proposed two near-field beam training schemes by designing the near-field codebook for the XL-RIS assisted systems. Simulation results show that the proposed near-field beam training schemes can achieve better performance than the existing far-field beam training scheme. Particularly, compared with the near-field beam training scheme, the hierarchical near-field beam training scheme can reduce the beam training overhead by about 90% with about 95% achievable rate performance. For future works, the multi-beam training method can be used for the near-field XL-RIS beam training to further reduce the beam training overhead.

ACKNOWLEDGEMENT

This work was supported in part by the National Key Research and Development Program of China (Grant No. 2020YFB1807205), in part by the National Natural Science Foundation of China (Grant No. 62031019), and in part by the European Commission through the H2020-MSCA-ITN META WIRELESS Research Project under Grant 956256.

References

[1] E. Basar, M. Di Renzo, *et al.*, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, Aug. 2019, pp. 116 753–116 773.

[2] L. Dai, B. Wang, *et al.*, "Reconfigurable intelligent surface-based wireless communications: Antenna design, prototyping, and experimental results," *IEEE Access*, vol. 8, Mar. 2020, pp. 45 913–45 923.

[3] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, Nov. 2019, pp. 5394–5409.

[4] C. Jia, H. Gao, *et al.*, "Machine learning empowered beam management for intelligent reflecting surface assisted mmWave networks," *China Commun.*, vol. 17, no. 10, Oct. 2020, pp. 100–114.

[5] P. Wang, J. Fang, *et al.*, "Intelligent reflecting surface-assisted millimeter wave communications: Joint active and passive precoding design," *IEEE Trans. Veh. Technol.*, vol. 69, no. 12, Dec. 2020, pp. 14 960–14 973.

[6] Z. Zhang and L. Dai, "A joint precoding framework for wideband reconfigurable intelligent surface-aided cell-free network," *IEEE Trans. Signal Process.*, vol. 69, Aug. 2021, pp. 4085–4101.

[7] Q. Wu, S. Zhang, *et al.*, "Intelligent reflecting surface-aided wireless communications: A tutorial," *IEEE Trans. Commun.*, vol. 69, no. 5, May 2021, pp. 3313–3351.

[8] X. Wei, D. Shen, *et al.*, "Channel estimation for RIS assisted wireless communications: Part I - fundamentals, solutions, and future opportunities," *IEEE Commun. Lett.*, vol. 25, no. 5, May 2021, pp. 1398–1402.

[9] D. Mishra and H. Johansson, "Channel estimation and low-complexity beamforming design for passive intelligent surface assisted MISO wireless energy transfer," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (IEEE ICASSP'19)*, Brighton, UK, May 2019, pp. 4659–4663.

[10] Q. Nadeem, H. Alwazani, *et al.*, "Intelligent reflecting surface-assisted multi-user MISO communication: Channel estimation and beamforming design," *IEEE Open J. Commun. Soc.*, vol. 1, May 2020, pp. 661–680.
[11] P. Wang, J. Fang, *et al.*, "Compressed channel estimation for intelligent reflecting surface-assisted millimeter wave systems," *IEEE Signal Process. Lett.*, vol. 27, May 2020, pp. 905–909.
[12] X. Wei, D. Shen, *et al.*, "Channel estimation for RIS assisted wireless communications: Part II - an improved method based on double-structured sparsity," *IEEE Commun. Lett.*, vol. 25, no. 5, May 2021, pp. 1402–1407.
[13] Z. Wang, L. Liu, *et al.*, "Channel estimation for intelligent reflecting surface assisted multiuser communications," in *Proc. IEEE Wireless Commun. and Networking Conf. (IEEE WCNC'20)*, Seoul, Korea, May 2020, pp. 1–6.
[14] W. Wu, D. Liu, *et al.*, "Low-complexity beam training for 5G millimeter-wave massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, Jan. 2020, pp. 361–376.
[15] Z. Xiao, T. He, *et al.*, "Hierarchical codebook design for beamforming training in millimeter-wave communication," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, May 2016, pp. 3380–3392.
[16] X. Gao, L. Dai, *et al.*, "Near-optimal beam selection for beamspace mmWave massive MIMO systems," *IEEE Commun. Lett.*, vol. 20, no. 5, May 2016, pp. 1054–1057.
[17] C. You, B. Zheng, *et al.*, "Fast beam training for IRS-assisted multiuser communications," *IEEE Wireless Commun, Lett.*, vol. 9, no. 11, Nov. 2020, pp. 1845–1849.
[18] B. Ning, Z. Chen, *et al.*, "Terahertz multi-user massive MIMO with intelligent reflecting surface: Beam training and hybrid beamforming," *IEEE Trans. Veh. Technol.*, vol. 70, no. 2, Feb. 2021, pp. 1376–1393.
[19] J. Zhang, Y. Huang, *et al.*, "Intelligent interactive beam training for millimeter wave communications," *IEEE Trans. Wireless Commun.*, vol. 20, no. 3, Mar. 2021, pp. 2034–2048.
[20] M. Najafi, V. Jamali, *et al.*, "Physics-based modeling and scalable optimization of large intelligent reflecting surfaces," *IEEE Trans. Commun.*, vol. 69, no. 4, Apr. 2021, pp. 2673–2691.
[21] W. Tang, M. Z. Chen, *et al.*, "Wireless communications with reconfigurable intelligent surface: Path loss modeling and experimental measurement," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, Jan. 2021, pp. 421–439.
[22] M. Di Renzo, A. Zappone, *et al.*, "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, Nov. 2020, pp. 2450–2525.
[23] Z. Chu, P. Xiao, *et al.*, "Intelligent reflecting surfaces enabled cognitive internet of things based on practical pathloss model," *China Commun.*, vol. 17, no. 12, Dec. 2020, pp. 1–16.
[24] M. Cui and L. Dai, "Channel estimation for extremely large-scale MIMO: Far-field or near-field?" *IEEE Trans. Commun.*, vol. 70, no. 4, Apr. 2022, pp. 2663–2677.
[25] K. T. Selvan and R. Janaswamy, "Fraunhofer and fresnel distances: Unified derivation for aperture antennas." *IEEE Ant. Propag. Mag.*, vol. 59, no. 4, Aug. 2017, pp. 12–15.
[26] M. Cui, Z. Wu, *et al.*, "Near-field communications for 6G: Fundamentals, challenges, potentials, and future directions," *arXiv preprint arXiv:2203.16318*, Mar. 2022.
[27] L. Wei, C. Huang, *et al.*, "Channel estimation for RIS-empowered multi-user MISO wireless communications," *IEEE Trans. Commun.*, vol. 69, no. 6, Jun. 2021, pp. 4144–4157.
[28] Y. Han, S. Jin, *et al.*, "Channel estimation for extremely large-scale massive MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 9, no. 5, May 2020, pp. 633–637.

Biographies

Xiuhong Wei received the B.E. degree in School of Mechanical, Electrical and Information Engineering from Shandong University, Weihai, China, in 2019. She is currently working toward the M.S. degree in electronic engineering from Tsinghua University, Beijing, China. Her research interests are massive MIMO, RIS assisted wireless communications, and AI based wireless communication.

Linglong Dai is an Associate Professor of Department of Electronic Engineering, Tsinghua University, China. His current research interests include massive MIMO, reconfigurable intelligent surface (RIS), millimeter-wave/Terahertz communications, and machine learning for wireless communications. He has received the IEEE Communications Society Leonard G. Abraham Prize in 2020, and the IEEE Communications Society Stephen O. Rice Prize in 2022. He was listed as a Highly Cited Researcher by Clarivate in 2020 and 2021. He was elevated as an IEEE Fellow in 2022.

Yajun Zhao was born in 1976. He has B.E. and Master degrees. Since 2010, he has acted as a radio expert in the wireless advanced research department at the ZTE Corporation. Prior to this, he worked for Huawei on wireless technology research in the wireless research department. Currently, he is mainly engaged in research on 5G standardization technology and future mobile communication technology (6G). His research interests include unlicensed spectrum, flexible duplex, CoMP, and interference mitigation.

Guanghui Yu received the Master and Ph.D. degrees in automatic control from Beijing Institute of Technology, China, in 1998 and 2003, respectively. Since 2003, he has acted as a radio expert in the Wireless Advanced Research Department of ZTE Corporation. His research interests include Wimax, 2G, 3G, 4G, 5G and 6G design in RAN, especially involved in multiplexing and access, MIMO, interference management, and channel modeling, as well as network architecture. He is also involved as one of the main researchers in the link and system simulation platform and takes part in all kinds of 3GPP RAN1activities.

Xiangyang Duan received his B.E. and M.E. degrees in mechanical engineering and material processing engineering from Tsinghua University, Beijing, China, in 1999. He works at ZTE Corporation, Shenzhen, China, where he researches the universal mobile telecommunications system, baseband units, and wireless system architecture. His research interests include 5G technology and mobile edge computing as well as novel services and applications for future mobile networks.