Low-Overhead Beam Training Scheme for Extremely Large-Scale RIS in Near Field

Wang Liu, Cunhua Pan, Senior Member, IEEE, Hong Ren, Member, IEEE, Feng Shu, Member, IEEE, Shi Jin, Senior Member, IEEE, and Jiangzhou Wang, Fellow, IEEE

Abstract—Extremely large-scale reconfigurable intelligent surface (XL-RIS) has recently been proposed and is recognized as a promising technology that can further enhance the capacity of communication systems and compensate for severe path loss. However, the pilot overhead of beam training in XL-RIS-assisted wireless communication systems is enormous because the near-field channel model needs to be taken into account, and the number of candidate codewords in the codebook increases dramatically. To tackle this problem, we propose two deep learning-based near-field beam training schemes in XL-RIS-assisted communication systems, where deep residual networks are employed to determine the optimal near-field RIS codeword. Specifically, we first propose a far-field beam-based beam training (FBT) scheme in which the received signals of all far-field RIS codewords are fed into the neural network to estimate the optimal near-field RIS codeword. In order to further reduce the pilot overhead, a partial near-field beam-based beam training (PNBT) scheme is proposed, where only the received signals corresponding to the partial near-field XL-RIS codewords are input to the neural network. Moreover, we further propose an improved PNBT scheme to enhance the performance of beam training by fully exploring the neural network’s output. Finally, simulation results show that the proposed schemes outperform the existing beam training schemes and can reduce the beam sweeping overhead by approximately 95%.

Index Terms—Near-field, extremely large-scale RIS, beam training, deep residual learning.

I. INTRODUCTION

RECENTLY, reconfigurable intelligent surface (RIS) has received considerable research attention and is regarded as one of the key technologies for the next-generation communication systems since it can significantly improve the energy and spectrum efficiencies of the wireless communication systems [1], [2], [3], [4]. Specifically, an RIS consists of numerous passive and controllable reflecting elements which can independently adjust the phase shift of the incident electromagnetic signal with low power consumption. By properly tuning the coefficients of the RIS reflecting elements, RIS can improve the communication signal quality and establish an additional reliable reflecting link between the user equipment (UE) and the base station (BS). Furthermore, recent advances in materials enable RIS to tune the reflecting coefficients in real time to cope with the rapidly varying wireless propagation environment [5]. Reliable RIS-assisted communication systems rely heavily on accurate channel state information (CSI) for the phase shift design. However, it is challenging to obtain accurate CSI because of the passive property of the RIS.

Specifically, since the RIS lacks the ability to process the signal, only the cascade channel, i.e., the BS-RIS-UE channel, can be estimated. Accordingly, some channel estimation schemes, such as least square (LS) or minimum mean squared error (MMSE), have been proposed in [6], [7], and [8] for cascaded channel estimation. However, the pilot overhead required for cascaded channel estimation is proportional to the number of RIS elements and becomes excessive when the number of elements is large. In order to reduce the pilot overhead, the authors of [9] proposed a novel multi-user-based channel estimation scheme, where the correlation of each user’s cascade channel was explored. Furthermore, some compressive sensing (CS) based channel estimation schemes were proposed to reduce the pilot overhead, where the sparsity of the channel was leveraged [10], [11].

Compared to channel estimation, codebook-based beam training has been widely adopted for millimetre-wave communication systems owing to its lower operation complexity [12]. Furthermore, beam training techniques have recently been extended to RIS-assisted wireless communication systems [12], [13], [14]. Specifically, the RIS codebook is first designed based on the cascaded array steering vector. Then the RIS selects the optimal codeword from the codebook to form the phase shift based on the information acquired during
the training phase. For beam training, testing all codewords in the codebook during the training phase is the most straightforward scheme; however, this will incur an excessive pilot overhead [15]. To reduce the pilot overhead, a hierarchical codebook-based beam training scheme was proposed for millimetre wave (mmWave) communication systems, where the training phase was divided into multiple parts and the range of the optimal codeword was reduced in each part [16], [17], [18]. Moreover, the authors of [14] proposed a partial search-based beam training scheme to reduce the complexity of training.

Recently, deep learning (DL) as a branch of machine learning has received tremendous research attention and has been applied in beam training to reduce the pilot overhead since the neural networks possess a powerful ability to learn non-linear relations [19], [20], [21], [22], [23]. For example, the authors of [22] proposed to employ a deep neural network (DNN) to determine the optimal codeword based on the information of the previous low-frequency channel. In [20], a long and short-term memory (LSTM) network-based beam training scheme was proposed in which the received signals of the previous time slots are employed to predict the optimal beam for the next time slot. In addition, [23] demonstrated that the received signals corresponding to different codewords contain implicit non-linear relationships between them and these relationships can be explored by the DNN to determine the optimal codeword, where only portion of the codewords need to be tested.

However, the existing beam training schemes cannot be directly applied to the RIS-assisted wireless communication systems. Specifically, the RIS is expected to employ more reflecting elements to compensate for the severe "multiplicative fading" effect in the cascaded channel, where the path loss of the cascaded channel is equivalent to the product of the path loss of the two links. Due to the advantages of low power consumption and low cost of the RIS, it is feasible to deploy large-scale reflecting elements at the RIS to compensate for this effect [24]. Consequently, extremely large-scale RIS (XL-RIS) has been proposed and regarded as one of the prospective directions of the RIS [25]. However, the increased number of reflecting elements will lead to larger Rayleigh distance, which is the boundary between the near and far fields, and the range of the near field will expand accordingly. Consequently, the near-field domain becomes non-negligible, thus scatters, as well as users, are more likely to lie in the near-field domain. For the near-field domain, the electromagnetic field structure has fundamentally changed, and the spherical wave channel model should be considered when designing the RIS codebook. Since most of the existing beam training schemes were based on far-field codebooks, it is challenging to implement them in the near-field domain. In order to address this problem, the authors of [26] proposed a near-field codebook based on a spherical wave channel model at the BS, which increases distance sampling to accommodate the near-field domain. Furthermore, [25] designed a codebook for the XL-RIS, which was based on cascaded array steering vectors in the near-field domain. It can be observed from [25] and [26] that the near-field codebooks contain significantly more candidate codewords than the far-field codebooks, which makes beam training in the near-field much more complicated and entails increased pilot overhead. Although the authors of [25] proposed a hierarchical near-field RIS codebook and the corresponding hierarchical beam training scheme to reduce the pilot overhead, the required pilot overhead is still excessive, and the hierarchical training scheme is highly susceptible to noise. To the best of our knowledge, the existing beam training schemes cannot address the excessive pilot overhead in the XL-RIS-assisted communication systems.

In order to fill this gap, we propose two effective deep learning-based near-field beam training schemes for XL-RIS-assisted communication systems, where the pilot overhead is significantly reduced with excellent training performance. Our contributions are summarized as follows.

1) We propose a far-field beam-based beam training (FBT) scheme where a deep residual network (DRN) is employed to estimate the optimal near-field XL-RIS codeword. Specifically, the XL-RIS tests all far-field RIS codewords in the training phase and then the received signals corresponding to all far-field RIS codewords at the BS are input to the deep residual network (DRN). Next, the DRN estimates the optimal near-field XL-RIS codeword based on the input signals by exploring the implicit relationship between the received signals of different codewords. Since the number of candidate codewords in the far-field RIS codebook is much smaller than that in the near-field XL-RIS codebook, the pilot overhead of the proposed scheme is significantly reduced.

2) We propose a partial near-field beam-based beam training (PNBT) scheme to further reduce the pilot overhead. Compared to the FBT scheme, the scheme only tests partial near-field XL-RIS codewords and feeds the corresponding received signals to the DRN. Furthermore, based on the PNBT scheme, we further propose an improved PNBT scheme to enhance the performance of beam training. Specifically, the improved PNBT scheme will perform additional beam test of the more likely near-field XL-RIS codewords based on the neural network’s output.

3) Finally, we provide numerical simulations to validate the performance of our two proposed near-field XL-RIS beam training schemes. Compared with the sweeping scheme, which exhaustively tests all codewords, the proposed scheme can obtain similar achievable rate performance, but the pilot overhead is greatly reduced. Moreover, the DRN used in this paper can better extract the features of the received signal to improve the performance of beam training than ordinary convolutional networks or fully connected networks.

The paper is organized as follows. In Section II, we firstly introduce the signal model in the XL-RIS aided communication system and subsequently describe the far-field and near-field channel models. The corresponding beam training models are also presented in Section II. Section III first shows the problem formulation and the neural network structure. Then, two near-field RIS beam training schemes are presented.
In Section IV, we provide the simulation results. Conclusions are drawn in Section V.

In this paper, we adopt the following notations: Vectors and matrices are represented in bold lower case and bold upper case, respectively, e.g., $\mathbf{a}$ and $\mathbf{A}$, while $a$ and $A$ denote a scalar and a set. $(\cdot)^H$, $(\cdot)^T$ and $(\cdot)^H$ represent conjugate, transpose and conjugate transpose, respectively. $|\cdot|$ denotes absolute value. $[A]$ denotes the $i$-th element of $A$. $(\cdot)^*$ denote the order operation, e.g., for $A = \{a_1, a_2, \ldots, a_n\}$, $(A) = \{a_{\sigma_1}, a_{\sigma_2}, \ldots, a_{\sigma_n}\}$ with $a_{\sigma_1} \geq a_{\sigma_2} \geq \cdots \geq a_{\sigma_n}$. $\mathcal{CN}(\mu, \sigma)$ represents the Gaussian distribution with mean of $\mu$ and variance of $\sigma$. $\mathcal{U}(a, b)$ represents the uniform distribution between $a$ and $b$. Finally, diag($a$) represents the diagonal matrix using vector as the diagonal elements.

II. SYSTEM MODEL

A. Signal Model

As shown in Fig. 1, we consider an uplink time division duplexing (TDD) based orthogonal frequency division multiplexing (OFDM) communication system, where the direct link is blocked by obstacles, and an XL-RIS is employed between the BS and the $K$ users to assist communications [27]. The BS is equipped with $M$-antenna ($M = M_1 \times M_2$) uniform planar array (UPA), and each user is equipped with a single antenna. The XL-RIS with $N = N_1 \times N_2$ elements is deployed in the $x$-$y$ plane and its centre is located at the origin of the coordinate axis. It is noted that there is a separate control link between the BS and the RIS, which is utilized by the BS to adjust the phase shifts of the XL-RIS reflection elements [14], [28], [29].

Let us denote $\mathbf{G} \in \mathbb{C}^{M \times N}$ as the channel matrix from the BS to the XL-RIS, and $\mathbf{h}_k \in \mathbb{C}^{N \times 1}$ as the channel vector from the XL-RIS to the $k$-th user. During the uplink communication, the users transmit the pilot signals to the BS via the XL-RIS, and the processed pilot signal of the $k$-th user received at the BS side can be represented as

$$r_k = f_k^H \mathbf{G} \mathbf{h}_k x_k + f_k^H n_k,$$

where $f_k \in \mathbb{C}^{M \times 1}$ denotes the analog combining vector for the $k$-th user at the BS; $\mathbf{\Phi} = \text{diag}(\phi_1, \phi_2, \cdots, \phi_N)$ is a diagonal matrix and each element on the diagonal represents the phase shift of the reflecting elements of the XL-RIS; $x_k \in \mathbb{C}$ denotes the pilot signal transmitted by the $k$-th user, which satisfies $|x_k|^2 = 1$; $n_k \in \mathbb{C}^{M \times 1}$ is the noise vector at the BS, and each element follows the complex Gaussian distribution of zero mean and variance of $\sigma^2$.

B. Channel Model

The conversion of RIS to XL-RIS means not only the increase of the number of the reflecting elements but also the fundamental transformation of the channel model and electromagnetic field structure [26]. Specifically, the electromagnetic radiation field consists of two main regions, i.e., the far-field and near-field regions, and these two regions are divided by Rayleigh distance $Z = \frac{\lambda_c}{2D}$, where $D$ and $\lambda_c$ represent the array aperture and carrier wavelength, respectively [30].

In conventional RIS-aided communication systems, the array aperture of the RIS is not very large, and the corresponding Rayleigh distance is small, which results in a small near-field range, and most of the scatters and targets are in the far-field domain [31]. However, with the increase in the array aperture of the RIS, the range of the near-field also increases, which makes the spherical wave channel model no longer negligible. In the following, we briefly introduce the channel models in the far-field and near-field domains, respectively.

1) Far-Field Channel Model: In the far-field domain, the channels $\mathbf{G}$ and $\mathbf{h}_k$ can be represented by the widely used Saleh-Valenzuela channel model. Specifically, the channel $\mathbf{G}$ can be represented as

$$\mathbf{G} = \sqrt{\frac{MN}{L_G}} \sum_{l_1=1}^{L_G} \alpha_{l_1}^G \mathbf{a}(\vartheta_{l_1}^G, \psi_{l_1}^G) \mathbf{b}^H (\vartheta_{l_1}^{Gr}, \psi_{l_1}^{Gr}),$$

where $L_G$ denotes the number of paths between the BS and the XL-RIS; $\alpha_{l_1}^G$ represents the complex path gain of the $l_1$-th path; $\vartheta_{l_1}^G$ and $\vartheta_{l_1}^{Gr}$ denote the azimuth (elevation) angle at the BS and the azimuth (elevation) angle at the XL-RIS of the $l_1$-th path. In particular, the first channel path, i.e. $l_1 = 1$, represents the strongest channel path, which is generally the line-of-sight (LOS) path.

Similarly, the channel between the XL-RIS and the $k$-th user $\mathbf{h}_k$ can be formulated as

$$\mathbf{h}_k = \sqrt{\frac{N}{L_k}} \sum_{l_2=1}^{L_k} \alpha_{l_2}^k \mathbf{b}(\vartheta_{l_2}, \psi_{l_2}^k),$$

where $L_k$ denotes the number of paths between the XL-RIS and the $k$-th user; $\alpha_{l_2}^k$ represents the complex path gain and the azimuth (elevation) angle of the $l_2$-th path, respectively. In a similar way, $l_2 = 1$ represents the strongest path. Furthermore, $\mathbf{a}(\vartheta, \psi)$ and $\mathbf{b}(\vartheta, \psi)$ denote the far-field array steering vectors for the BS and the RIS, respectively. Specifically, $\mathbf{a}(\vartheta, \psi)$ can be formulated as [32]

$$\mathbf{a}(\vartheta, \psi) = \frac{1}{\sqrt{M}} \left[ e^{-j2\pi d \cos(\vartheta) m_1 / \lambda} \right] \otimes \left[ e^{-j2\pi d \sin(\vartheta) \cos(\psi) m_2 / \lambda} \right],$$

where $\lambda$ and $d$ denote the wavelength and the antenna spacing, which usually satisfy $d = \frac{\lambda}{3}$; $m_1 = [0, 1, \cdots, M_1 - 1]^T$ and $m_2 = [0, 1, \cdots, M_2 - 1]^T$. 

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
It should be noted that in mmWave communications, the received signal power is concentrated on the strongest path, and the purpose of beam training is to align the beam to the strongest path to maximize the beamforming gain while the other paths can be treated as the noise [33], [34]. Therefore, in this paper, we mainly consider the strongest path. Specifically, the received signal in (1) can be rewritten as
\[
r_k = \beta_k f_k^H a \left( \tilde{g}_1^c, \psi_1^c \right) b^H \left( \tilde{g}_1^G, \psi_1^G \right) \Phi b \left( \tilde{g}_1^k, \psi_1^k \right) x_k + n_{e,k},
\]
where \( \beta_k = \sqrt{\frac{M\alpha_k^2}{\sum N_k}} \) denotes the effective path gain of the \( k \)-th user, \( n_{e,k} \) denotes the effective noise, which consists of the combined channel noise \( n_k \) and combined received signals from the other paths. Since the locations of the BS and the RIS are generally fixed and the channel \( G \) is essentially constant over multiple channel coherence time intervals, we assume that the analog combining vector \( f_k \) at the BS has been designed to align to the strongest path, i.e. \( f_k = a \left( \tilde{g}_1^c, \psi_1^c \right) \). The received signal can be further formulated as [25]
\[
r_k = \beta_k b^H \left( \tilde{g}_1^G, \psi_1^G \right) \Phi b \left( \tilde{g}_1^k, \psi_1^k \right) x_k + n_{e,k}
= \beta_k b^H \left( \tilde{g}_1^G, \psi_1^G \right) \text{diag}(\Phi) b \left( \tilde{g}_1^k, \psi_1^k \right) x_k + n_{e,k}
= \beta_k \Phi^T \text{diag} \left( \text{b}^* \left( \tilde{g}_1^G, \psi_1^G \right) \right) b \left( \tilde{g}_1^k, \psi_1^k \right) x_k + n_{e,k}
= \beta_k \Phi^T b \left( \tilde{g}_1^k - \tilde{g}_1^G, \psi_1^k - \psi_1^G \right) x_k + n_{e,k},
\]
where \( \Phi = [\phi_1, \phi_2, \ldots, \phi_N]^T \) denotes the phase shift vector of the XL-RIS.

2) Near-Field Channel Model: In the near-field domain, the spherical wave channel model used [35], [36], which can be represented as
\[
G = \sqrt{\frac{M N L}{L_N}} \sum_{l=1}^{L_N} \alpha_l^G c_l \left( x_l^G, y_l^G, z_l^G \right) c_l^H \left( x_l^G, y_l^G, z_l^G \right),
\]
where \( \left( x_l^G, y_l^G, z_l^G \right) \) denotes the coordinate of the scatter on the \( l \)-th path between the XL-RIS and the BS. In particular, \( \left( x_l^G, y_l^G, z_l^G \right) \) represents the coordinate of the scatter on the strongest path. Similar to the array steering vectors in the far-field domain, \( c_l \left( x, y, z \right) \) and \( c_r \left( x, y, z \right) \) represent the near-field steering vectors at the XL-RIS and at the BS, respectively. Take \( c_r \left( x, y, z \right) \) as an example, the near-field steering vector can be formulated as
\[
c_r \left( x, y, z \right) = \frac{1}{\sqrt{N}} \left[ e^{-j \frac{2\pi}{D_{1,1}}}, \ldots, e^{-j \frac{2\pi}{D_{1,n_2}}}, \ldots, \right.
\left. e^{-j \frac{2\pi}{D_{N_1,1}}}, \ldots, e^{-j \frac{2\pi}{D_{N_1,n_2}}} \right]^T,
\]
where \( D_{n_1,n_2} = \sqrt{\left( x_{n_1,n_2} - x \right)^2 + y^2 + \left( z_{n_1,n_2} - z \right)^2} \) represents the distance from the element in the \( n_1 \)-th row and \( n_2 \)-th column on the XL-RIS to the scatter. \( \left( x_{n_1,n_2}, 0, z_{n_1,n_2} \right) \) represents the coordinate of the corresponding XL-RIS element.

Similarly, the channel between the XL-RIS and the \( k \)-th user \( h_k \) can be formulated as
\[
h_k = \sqrt{\frac{N}{L_N}} \sum_{l=1}^{L_N} \alpha_l^k c_r \left( x_l^k, y_l^k, z_l^k \right),
\]
where \( \left( x_l^k, y_l^k, z_l^k \right) \) denotes the coordinate of the scatter between the XL-RIS and the \( k \)-th user of the \( l \)-th path and \( \left( x_l^k, y_l^k, z_l^k \right) \) denotes the coordinate of the scatter or the user on the strongest path.

Similar to (5) and (6), the received signal in the near-field domain can be rewritten as
\[
r_k = \beta_k c_r^H \left( x_l^G, y_l^G, z_l^G \right) \text{diag}(\Phi) c_r \left( x_l^k, y_l^k, z_l^k \right) x_k + n_{e,k}
= \beta_k \Phi^T \text{diag} \left( c_r^* \left( x_l^G, y_l^G, z_l^G \right) \right) c_r \left( x_l^k, y_l^k, z_l^k \right) x_k + n_{e,k}
= \beta_k \Phi^T c_r \left( x_l^G, y_l^G, z_l^G \right) \left( x_l^k, y_l^k, z_l^k \right) x_k + n_{e,k},
\]
where \( c_r \left( x_l^G, y_l^G, z_l^G \right) \left( x_l^k, y_l^k, z_l^k \right) \) denotes the effective near-field steering vector and can be formulated as
\[
c_r \left( x_l^G, y_l^G, z_l^G \right) \left( x_l^k, y_l^k, z_l^k \right) = \frac{1}{\sqrt{N}} \left[ e^{-j \frac{2\pi}{D_{1,1}} - D_{1,1}}, \ldots, e^{-j \frac{2\pi}{D_{1,n_2}} - D_{1,n_2}}, \ldots, e^{-j \frac{2\pi}{D_{N_1,1}} - D_{N_1,1}}, \ldots, e^{-j \frac{2\pi}{D_{N_1,n_2}} - D_{N_1,n_2}} \right],
\]
where \( D_{n_1,n_2} \) and \( D_{n_1,n_2} \) denote the distances from the element on the XL-RIS to the scatter \( \left( x_l^G, y_l^G, z_l^G \right) \) and to the scatter \( \left( x_l^k, y_l^k, z_l^k \right) \), respectively.

Since the propagation environment and scatter locations between the XL-RIS and the BS are generally fixed, and the channel \( G \) remains constant over long channel coherence time, we assume that the coordinates of the scatter on the strongest path between the XL-RIS and the BS, i.e. \( \left( x_l^G, y_l^G, z_l^G \right) \), are fixed and known. Therefore, for simplicity, the received signal can be further rewritten as
\[
r_k = \beta_k \Phi^T c_r \left( x_l^k, y_l^k, z_l^k \right) x_k + n_{e,k}.
\]

C. Beam Training Model

It is assumed that the phase shift vector \( \Phi = [\phi_1, \phi_2, \ldots, \phi_N] \) of the XL-RIS is selected from the predefined codebook, where each codeword in the codebook can form a passive beam. Based on the far-field steering vector in (4) and the received signal in (6), the existing codebooks for the RIS in the far-field domain are generally designed as [25]
\[
F = \left\{ b(\theta_1, \varphi_1), \ldots, b(\theta_1, \varphi_N), \ldots, b(\theta_N, \varphi_1), \ldots, b(\theta_N, \varphi_N) \right\}^T,
\]
where \( \theta_n = \frac{2n - N_1 - 1}{N_1} \) with \( n = 1, 2, \ldots, N_1 \) and \( \varphi_n = \frac{2n - N_1 - 1}{N_2} \) with \( n = 1, 2, \ldots, N_2 \). Each element of \( F \) is a candidate beam codeword for the RIS in the far-field domain. Such an angular domain-based far-field codebook can exploit the far-field path’s full angular information.

However, the array steering vector in the near-field domain depends on the coordinates of the scatter, which makes the angle-only far-field codebooks no longer applicable to the
channels in the near-field. For the application of the near-field domain, the design of the codebook for the XL-RIS should be based on the near-field array steering vector in (11) and the received signal model in (12). Furthermore, the strongest path from the XL-RIS to the user is generally the LOS path, where \( (x_1^k, y_1^k, z_1^k) \) denotes the coordinate of the user, and the height of the user is generally constant. The acquisition of information about \( z_1^k \) can be carried out after beam training by using some localization algorithms or sweep schemes [25], and the acquired \( z_1^k \) can be deployed in multiple channel coherence time intervals. Therefore, for the design of the near-field codebooks, we mainly consider the exploration of \( x_1^k \) and \( y_1^k \), which vary more frequently and whose beam training is the major cause of the high pilot overhead. Specifically, the near-field codebook for the XL-RIS is designed as [25]

\[
W = \{c_r(x_1, y_1), \ldots, c_r(x_{S_x}, y_1), \ldots, c_r(x_1, y_{S_y}), \ldots, c_r(x_{S_x}, y_{S_y})\},
\]

(14)

where \( S_x \) and \( S_y \) denote the numbers of points sampled on the x-axis and y-axis, respectively; \( x_{s_x} \) and \( y_{s_y} \) represent the sampled points on the x-axis and y-axis respectively. In addition, the set of all sampled points can be represented as

\[
\Omega = \{(x_{s_x}, y_{s_y}) | x_{s_x} = x_{\min} + (s_x - 1)\Delta_x, y_{s_y} = y_{\min} + (s_y - 1)\Delta_y\},
\]

(15)

where \( x_{\min} \) and \( y_{\min} \) denote the minimum value of sampling points; \( \Delta_x \) and \( \Delta_y \) denote the sampling intervals on x-axis and y-axis, respectively.

Based on the well-designed codebook, the beam training aims to search for the optimal codeword that enables the system to achieve the maximum achievable rate. In order to search for the optimal codeword, the users need to send the pilot signals to the BS via the XL-RIS over \( Q \) time slots, and the XL-RIS will select different codewords from the codebook to form the passive beam at each time slot. Based on the received signal in (12), the pilot signal of the \( k \)-th user received at the BS at the \( q \)-th time slot can be expressed as

\[
r_{k,q} = \beta_k \bar{\Theta} c_r(x_1^k, y_1^k, z_1^k) x_k[q] + n_{c,k,q},
\]

(16)

where \( x_{k,q} \) denotes the pilot signal sent by the \( k \)-th user to the BS at the \( q \)-th time slot; \( \bar{\Theta} \) and \( n_{c,k,q} \) represent the phase shift vector of the XL-RIS and effective channel noise at the \( q \)-th time slot, respectively.

After \( Q \) time slots, the BS can obtain a \( Q \)-dimensional vector of received signals. By assuming \( x_{k,q} = 1 \), the vector of received signals \( r_k = [r_{k,1}, r_{k,2}, \ldots, r_{k,Q}]^T \) can be expressed as

\[
r_k = \beta_k \bar{\Theta} c_r(x_1^k, y_1^k, z_1^k) x_k + n_{c,k},
\]

(17)

where \( \bar{\Theta} = [\phi_r^1, \phi_r^2, \ldots, \phi_r^Q]^T \) and \( n_{c,k} = [n_{c,k,1}, n_{c,k,2}, \ldots, n_{c,k,Q}]^T \).

Furthermore, the users are assumed to adopt orthogonal frequency subcarriers [37], which enables the beam training between users to be independent of each other. For simplicity, the index \( k \) in (17) can be omitted, and the received signal vector \( r_k \) can be further formulated as

\[
r = \beta \Theta c_r(x_1,y_1,z_1) + n_c.
\]

(18)

Based on the received signal vector in (18) and codebook in (14), the beam training problem in the near-field domain can be formulated as

\[
\phi^* = \arg \max_{\phi \in \mathcal{W}} \log_2 \left( 1 + \frac{|\phi^T c_r(x_1,y_1,z_1)|^2}{\sigma^2} \right).
\]

(19)

In practice, a straightforward approach to the problem in (19) is the sweeping scheme, i.e. testing all the codewords in the codebook. Although the sweeping scheme can achieve higher achievable rates, the pilot overhead associated with the sweeping scheme is unacceptable. Specifically, for the beam sweeping scheme, the number of time slots for beam training, \( Q \), is too large, which results in fewer time slots for transmitting useful information in each channel coherence time interval. It is even worse in the near-field domain since the near-field codebook possesses more candidate codewords than the far-field codebook [25].

In order to improve the efficiency of beam training, some beam training methods based on hierarchical codebooks have been proposed in the far-field domain [16], [17], [18]. Similarly, a recent hierarchical training scheme based on the near-field codebook was proposed to solve the problem of excessive pilot overhead in the near-field domain [25]. The hierarchical training schemes generally contain two types of codebooks, namely wide beam codebook and narrow beam codebook, where each wide beam can cover several narrow beams. During the hierarchical training procedure, all wide beam codewords will be tested in the first stage to find the optimal wide beam, and all narrow beams covered by the optimal wide beam will then be tested in the second stage to search for the final optimal narrow beam codeword. Compared with the exhaustive sweep search, the hierarchical search schemes significantly reduce the pilot overhead.

III. DEEP LEARNING FOR BEAM TRAINING

A. Problem Formulation

Although the hierarchical training scheme reduces the pilot overhead required for beam training to some extent, the hierarchical training scheme is susceptible to noise as well as multipath, and the pilot overhead is high, especially in the near-field domain. From the research work in [23] and [20], we know that the received pilot signals corresponding to different beam codewords contain implicit relationships with the optimal codeword, which provides additional information about the optimal codeword. By exploring these implicit relationships, we can reduce the pilot overhead and estimate the optimal codeword based on a small number of pilots.

Specifically, when the \( s = s_x s_y \)-th codeword in the codebook \( \mathcal{W} \) is applied, we can rewrite (10) in the following form:

\[
r_k = \beta_k \phi_r^c c_r(x_1^k, y_1^k) x_k + n_{c,k},
\]

(20)
Thereafter, we can define the beam gain function as

$$g(x_s, y_s) = c_r(x_s, y_s)^H \mathbf{c}_r(x^k, y^k)$$

$$= \frac{1}{N} \sum_{n_1=1}^{N_1} \sum_{n_2=1}^{N_2} e^{-j \pi d (D_{n_1, n_2} - D_{n_1, n_2}^*)}, \quad (21)$$

where $D_{n_1, n_2}$ denotes the distance from the user to the $(n_1, n_2)$-th element of the XL-RIS; $D_{n_1, n_2}^*$ denotes the distance from the position at which the beam is pointed to the $(n_1, n_2)$-th element of the XL-RIS. Obviously, $g(x_s, y_s)$ reflects the gain under the beams formed by different codewords and the relationship between them, and this relationship would be further reflected in the received signal. However, this relationship is highly non-linear and complicated, and an example is given in Fig. 2 for a more intuitive illustration, where it is assumed that the y-coordinates of the beams have been aligned to the user. In Fig. 2, the red dots represent the beam gain for different values of $x_s$, which is discrete and equally spaced. It can be seen that the gains of all beams are sampled under the same envelope. Since the beam gains are sampled at the same envelope with equal intervals, when the gain of one beam is determined, the gains of the remaining beams are successively limited or determined [21], [24]. Each beam gain therefore implicitly contains information about the gain of the remaining beams, which is non-linear and difficult to represent. Eventually, this underlying connection will be reflected in the received signals. By exploring this implicit connection, we can obtain the beam gain of the remaining codewords while testing only a portion of the codewords, and thus select the optimal codeword.

Therefore, in the near-field beam training, we propose to utilize the implicit relationship among codewords to estimate the optimal near-field codeword, in which only a small number of codewords need to be tested. Mathematically, the beam training model can be formulated as

$$s^* = f(\mathbf{r}), s^* \in \{1, 2, \ldots, s_x, s_y\}, \quad (22)$$

where $f(\cdot)$ denotes a mapping function, $s^*$ denotes the index of the optimal codeword in the near-field codebook.

However, the conventional estimation methods are difficult for dealing with these implicit relationships for two reasons. On the one hand, the relationships between the received signals and the optimal codeword are highly non-linear. On the other hand, the distribution of the equivalent noise $\mathbf{n}$ is difficult to deal with, which will seriously impact the estimation performance [20], [23]. Motivated by the application of deep learning methods, we propose to employ the neural network, which has a robust ability to learn non-linear relationships and deal with noise, to address this complex estimation problem. Specifically, we train two separate neural networks, i.e. an x-axis network and a y-axis network, to estimate the x-axis index $s_x$ and the y-axis index $s_y$ of the optimal near-field codeword, respectively. Therefore, the proposed deep learning-based beam training model can be further represented as

$$s^*_x = f_x(\mathbf{r}), s^*_x \in \{1, 2, \ldots, S_x\},$$

$$s^*_y = f_y(\mathbf{r}), s^*_y \in \{1, 2, \ldots, S_y\}, \quad (23)$$

where $f_x$ and $f_y$ denote the mapping functions of the x-axis network and the y-axis network, respectively; $s^*_x$ and $s^*_y$ denote the x-axis index and the y-axis index of the optimal codeword in the near-field codebook. According to (23), the index of the optimal near-field codeword can be represented as

$$s^* = (s^*_y - 1) S_x + s^*_x. \quad (24)$$

Furthermore, for the selection of the codewords to be tested, we propose to employ all far-field codewords in (13) and partial near-field codewords in (14), respectively. Accordingly, we propose two different beam training schemes, namely the FBT scheme and the PNB scheme. For the FBT scheme, all far-field beam codewords in (13) are tested, and the corresponding pilot received signals are used as the input to the neural network to estimate the optimal near-field codeword. Since far-field codebooks have far fewer codewords than near-field codebooks, the pilot overhead is effectively reduced compared to the sweeping and hierarchical schemes in the near-field domain. By contrast, for the PNB scheme, we select codewords from the near-field codebook at equal intervals for testing, which can also effectively reduce the pilot overhead.

It should be noted that the proposed deep learning-based beam training scheme consists of two phases, i.e. the neural network training phase and the estimation phase. For the training phase, training data containing the received pilot signal vectors and the corresponding optimal codeword index labels are collected to train the neural network, where the index labels of the optimal codewords can be obtained by conventional beam training methods. The specific neural network training steps are as follows:

1. For each user, the XL-RIS tests all near-field codewords and the BS receives the corresponding received signals.
2. The BS measures the quality of all received signals and obtains the optimal near-field codeword.
3. A dataset is constructed by storing received signals as features and their optimal codeword as labels. Repeat and obtain multiple sets of such data.
4. The obtained datasets are utilized to train the neural network.
For the estimation phase, the neural network has been fully trained and utilizes the pilot received signal vector as the input to estimate the index of the optimal codeword.

B. Deep Learning Model Design

In this section, we describe the structure of the proposed neural network and its principles in detail.

In the XL-RIS-assisted communication system, it is necessary to add distance sampling when designing the predefined codebook for the XL-RIS due to the spherical wave-based near-field channel, and the number of candidate codewords in the codebook increases dramatically. As a result, the input size of the neural network will inevitably increase as well, which makes the training of neural networks difficult along with overfitting problems. Fortunately, the convolutional neural networks (CNNs) employ the local sensing and parameter sharing approach, which effectively reduces the weight parameters and thus facilitates the training. Specifically, the CNNs only use a small number of convolutional kernels for feature extraction when faced with high-dimensional inputs, with each convolutional kernel representing the extraction of one feature. When extracting specific features, different positions of the input are convolved by the same convolution kernel, i.e. the weight parameters are shared. In addition, the CNNs possess many advantages in terms of good fault tolerance, parallel processing capability and high generalization ability, all of which motivates us to choose the CNNs to construct our near-field beam training algorithm.

Moreover, based on the CNN, we further adopt the deep residual network (DRN), which is a variation of CNN and can solve the performance degradation problem when the number of CNN’s layers is large. Specifically, in order to better extract the features in the input vector, the number of CNN’s layers should be increased. However, as the number of CNN’s layers increases to a certain extent, the performance of the neural network will degrade instead. This is due to the fact that the deeper the neural network is, the more pronounced the gradient disappearance is, which makes the parameters of the previous layers of the network not updated. To deal with this issue, residual networks are proposed, which consist of multiple residual blocks, to improve the performance of the network further.

1) Preprocessing Module: Before being fed into the neural network, the received signal vector is subject to data processing by the preprocessing module. Firstly, the received signal vector in complex form is divided into real part and imaginary part

\[ \mathbf{r} = \mathbf{r}_R + j \mathbf{r}_I \]

where \( \mathbf{r}_R \) and \( \mathbf{r}_I \) denote the real part and imaginary part of the received signal, respectively. After the signal is divided into real and imaginary parts, the two parts are transformed into a matrix with similar number of rows and columns. Next, the first convolutional layer would upgrade the input from the two feature channels to the sixty-four feature channels by convolving.

2) Convolution Module: In this module, massive convolution layers are deployed to extract features from the input received signal vector. For the residual networks, every two or three convolution layers would constitute the residual blocks first, and then the residual blocks are composed of the convolution module. As shown in Fig. 3, the convolution module of the residual network consists mainly of two basic

residual blocks, namely Basic Block and Bottle Block, while the latter is designed for deeper networks. In comparison to the regular convolution structure, the Basic Block and Bottle Block add identity mapping to the original base connection, which passes the current input directly to the next block. Such a structure, which is known as Short Connection, ensures that the gradient of the deeper network can be passed directly to the shallower network during backpropagation, thereby overcoming the problem of gradient disappearance to some extent.

For the configuration of the structure of the deep residual network, we refer to the classical Resnet18 and Resnet50 network, whose parameters are presented in Table II in Section IV. The structural parameters adopted in Table II are derived from classical residual network model in paper [38], which is the source of the study of residual networks and its proposed Table II-based network model proved to have more efficient mapping capabilities and the ability to mitigate gradient disappearance.

As an example, the structure of an 18-layer residual network, i.e. Resnet18, is shown in Fig. 4. From Fig. 4, it can be seen that the convolution module of Resnet18 consists of eight Basic Blocks, where the solid lines represent identity mappings of equal dimensions and the dashed lines represent identity mappings of different dimensions, which require additional convolution processing to convert to the same dimensions.

Note that each convolution layer is followed by a batch norm layer and a ReLU activation layer, whose purpose is to prevent overfitting and to increase the nonlinear mapping capability of the network, respectively. Furthermore, at the end of the convolution module, the average pooling layer is employed to transform the dimensionality of the output and map it to the fully connected layer.

3) Output Module: Since the near-field deep learning-based beam training is similar to a multi-classification problem, a fully connected layer and a Softmax activation layer are deployed to perform the estimation of the optimal codeword. The function of the fully connected layer is to refine the features extracted by the convolution module and estimate the probability of each codeword being the optimal codeword, while the function of the Softmax activation layer is to normalize the output of the fully connected layer into probability values. Specifically, the function of the Softmax activation layer can be formulated as

\[
p_i = \frac{e^{a_i}}{\sum_{i=1}^{S} e^{a_i}},
\]

where \( I \) denotes the output dimension of the fully connected layer; \( p_i \) and \( a_i \) represent the \( i \)-th output of the Softmax activation layer and the fully connected layer, respectively. Since the proposed scheme employs the dual neural network structure, i.e. x-axis network and y-axis network, to estimate the optimal near-field codeword separately, two probability distribution vectors would be obtained, which can be formulated as

\[
\hat{P}^x = [\hat{p}_1^x, \hat{p}_2^x, \cdots, \hat{p}_{S_x}^x]^T, \\
\hat{P}^y = [\hat{p}_1^y, \hat{p}_2^y, \cdots, \hat{p}_{S_y}^y]^T,
\]

(26)
where $p_{sx}^*$ and $p_{sy}^*$ denote the probabilities that the x-axis index of the optimal near-field codeword is $s_x$ and the y-axis index is $s_y$, respectively. The greater the values of $p_{sx}^*$ and $p_{sy}^*$, the higher the probability that the optimal near-field codeword is $\tilde{c}_r(x_{s_x}, y_{s_y})$.

Furthermore, the cross-entropy loss function will be used as an evaluation criterion, which is widely applied in multi-classification problems. Specifically, the cross-entropy loss function of the x-axis network and the y-axis network can be expressed as

$$\text{Loss}^x = -\sum_{s_x=1}^{S_x} p_{sx}^* \log_{10} \hat{p}_{sx}^*,$$

$$\text{Loss}^y = -\sum_{s_y=1}^{S_y} p_{sy}^* \log_{10} \hat{p}_{sy}^*, \quad (27)$$

where $p_{sx}^* = 1$ and $p_{sy}^* = 1$ if the actual optimal near-field codeword is $c_r(x_{s_x}, y_{s_y})$. Otherwise, $p_{sx}^* = 0$ and $p_{sy}^* = 0$.

### C. Far-Field Beam-Based Beam Training

Given the detailed design of the residual network structure in the previous section, two primary beam training schemes are developed, namely the FBT scheme and PNBT scheme, which differ in the choice of the codewords to be tested in the pilot transmission phase. In this section, the detailed steps of the FBT scheme will be provided.

The framework of the proposed FBT scheme is shown in Fig. 5. For the FBT scheme, the user first sends the pilot signals to the BS via the XL-RIS, during which the XL-RLS selects codewords from the far-field RIS codebook to form the phase shifts at different time slots. By defining the receive signal of the $n$-th far-field RIS codeword as $y_{F_n}$, the BS can obtain the receive signal vector $r_{F_n} = [r_{F_1}^T, r_{F_2}^T, \cdots, r_{F_N}^T]^T$.

According to (16), the received signal $r_{F_n}^T$ can be expressed as

$$r_{F_n}^T = \beta u_n^T \tilde{c}_r(x_{s_x}, y_{s_y}) x + n_{e,n}, \quad (28)$$

where $u_n$ denotes the $n$-th codeword in the far-field RIS codebook $\mathcal{F}$. Note that since all far-field RIS codewords are tested, the total time slot length $Q$ is equal to the number of far-field RIS codewords $N$.

Then, the received signal vector $r_{F_n}^T$ are fed into the x-axis network and the y-axis network respectively and the probability distribution vectors $\hat{P}_x$ and $\hat{P}_y$ can be obtained according to (26). Based on $\hat{P}_x$ and $\hat{P}_y$, the x-axis index and y-axis index of the optimal near-field codeword can be expressed as

$$\hat{s}_x = \arg \max_{s_x=1,2,\cdots,S_x} \hat{P}_x_{s_x},$$

$$\hat{s}_y = \arg \max_{s_y=1,2,\cdots,S_y} \hat{P}_y_{s_y}. \quad (29)$$
Then the index of the optimal near-field codeword in the codebook can be formulated as
\[
\hat{s}^* = (\hat{s}^*_y - 1) S_x + \hat{s}^*_x.
\] (30)

Finally, the optimal near-field RIS codeword can be obtained by
\[
\hat{w}^* = [W]_{\hat{s}^*}.
\] (31)

**Algorithm 1** The Far-Field Beam-Based Training

**Input:**
Well-trained x-axis and y-axis residual networks;

**Output:**
Optimal near-field codeword \( w^* \);

1: XL-RIS implements far-field beam test and the received signal vectors \( r^F \) can be obtained at the BS according to (28);
2: Obtain the probability distribution vector \( \hat{P}^x \) and \( \hat{P}^y \) by inputting \( r^F \) to the x-axis and y-axis residual networks, respectively;
3: The x-axis index and y-axis index of the optimal near-field codeword can be obtained via (29);
4: The optimal near-field codeword \( \hat{w}^* \) can be obtained via (31);
5: return \( \hat{w}^* \);

**D. Partial Near-Field Beam-Based Training**

1) Partial Near-field Beam-based Training Scheme: Compared to the FBT scheme, the XL-RIS in the PNBT scheme selects the partial near-field codewords for test at step 2. Specifically, the XL-RIS selects the codewords from the near-field RIS codebook at equally spaced intervals according to the index to form the phase shifts at different pilot time slots. By defining \( D \) as the sampling interval and \( I = \left\lfloor \frac{S_y S_x}{D} \right\rfloor \) as the number of codewords picked, the set of selected near-field RIS codewords can be represented as
\[
D = \{ w_{t_i} | t_i = i \times D, i = 1, 2, \cdots, I \},
\] (32)

where \( w_{t_i} \) denotes the \( t_i \)-th codeword in the near-field RIS codebook \( W^N \).

Then, according to (16), the received signal of the \( t_i \)-th near-field codeword can be represented as
\[
r^N_{t_i} = \beta w^T w_{t_i} c_r (x^k, y^k, z^k) x + n_{e,t_i},
\] (33)

and the received signal vector \( r^N = [r^N_{t_1}, r^N_{t_2}, \cdots, r^N_{t_I}]^T \) can be obtained at the BS.

Furthermore, the remaining steps for obtaining the optimal near-field codeword in the PNBT scheme are the same as the FBT scheme and will not be repeated here.

**Algorithm 2** The Partial Near-Field Beam-Based Training

**Input:**
Well-trained x-axis and y-axis residual networks;

**Output:**
Optimal near-field codeword, \( w^* \);

1: XL-RIS implements partial near-field codewords test and the received signal vectors \( r^N \) can be obtained at the BS according to (33);
2: Obtain the probability distribution vector \( \hat{P}^x \) and \( \hat{P}^y \) by inputting \( r^N \) to the x-axis and y-axis residual networks, respectively;
3: The x-axis index and y-axis index of the optimal near-field codeword can be obtained via (29);
4: The optimal near-field codeword \( \hat{w}^* \) can be obtained via (31);
5: return \( \hat{w}^* \);

2) The Improved PNBT Scheme: The PNBT scheme can significantly reduce the pilot overhead; however, the estimation performance will degrade when the sampling interval \( D \) is large. In order to further improve the performance of the PNBT scheme, we propose an improved PNBT scheme in which the output of the neural network, i.e. \( \hat{P}^x \) and \( \hat{P}^y \), will be further exploited and additional codeword tests are required. The framework of the improved PNBT scheme is shown in Fig. 6.
Fig. 6. The framework of the improved PNBT scheme in which only partial near-field XL-RIS codewords are tested in the first phase and only codewords in set $B'$ are tested in the second phase.

As shown in the diagram of Algorithm 3, the steps to obtain the probability distribution vectors $\hat{P}_x$ and $\hat{P}_y$ are the same for both the PNBT scheme and the improved PNBT scheme. Based on the $\hat{P}$ and $\hat{P}'$, the improved PNBT scheme performs additional tests on some of the more possible near-field codewords rather than directly selecting the most possible codeword, which differs from the original scheme. Specifically, the BS can obtain the first $K$ maximum possible $x$-axis indices and $L$ maximum possible $y$-axis indices according to $\hat{P}$ and $\hat{P}'$, which can be clearly represented in the set $\mathcal{L}_x$ and $\mathcal{L}_y$ as

$$\mathcal{L}_x = \{\sigma_1, \sigma_2, \ldots, \sigma_K\}$$

and

$$\mathcal{L}_y = \{\gamma_1, \gamma_2, \ldots, \gamma_L\}.$$ 

(36)

(37)

where $\langle \cdot \rangle$ denotes the order operation, e.g., for $A = \{a_1, a_2, \ldots, a_n\}$, $\langle A \rangle = \{a_{\sigma_1}, a_{\sigma_2}, \ldots, a_{\sigma_n}\}$ with $a_{\sigma_1} \geq a_{\sigma_2} \geq \ldots \geq a_{\sigma_n}$. Based on $\mathcal{L}_x$ and $\mathcal{L}_y$, the codewords that point to the intersection of these indices can be represented by the set $B$ as

$$B = \{w_j | b_j = (\gamma - 1)S_x + \sigma, \sigma \in \mathcal{L}_x, \gamma \in \mathcal{L}_y, j = 1, 2, \ldots, KL\}.$$ 

(38)

In order to avoid duplicate test, the set $B$ requires to exclude the near-field codewords from the first test, and eventually, the additional codewords that need to be tested can be represented as

$$B' = B - B \cap D.$$ 

(39)

Then the BS delivers the set $B'$ to the XL-RIS via the direct control link, and the user needs to transmit a few pilot signals again for additional near-field codewords test, whose procedures are the same as that for the first test. By implementing additional near-field codewords test, the received signal vector corresponding to the near-field codewords with the maximum
probability can be obtained as
\[
\begin{align*}
W^N &= \beta w^b_1 e_r (x_1^k, y_1^k, z_1^k) x + n_{r,n}, \quad (40) \\
W^N &= [r_{b_1}, r_{b_2}, \ldots, r_{b_{K'}}]^T. \quad (41)
\end{align*}
\]

Finally, based on the measurement of the received signal vector, the index of the optimal near-field codeword in the codebook \(W\) can be represented as
\[
s^* = \arg \max_{b_j \in B} \left| r_{b_j} \right|, \quad (42)
\]
and the optimal near-field codeword can be obtained as
\[
\hat{w}^* = |W|_{s^*}. \quad (43)
\]

**Algorithm 3 The Improved PNBT**

**Input:**
- Well-trained x-axis and y-axis residual networks;
- The number of additional codewords to be tested \(K, L\);

**Output:**
- Optimal near-field codeword, \(w^*\);

1. XL-RIS implements partial near-field codewords test and the received signal vectors \(r^N\) can be obtained at the BS according to (33);
2. Obtain the probability distribution vector \(\hat{P}^x\) and \(\hat{P}^y\) by inputting \(r^N\) to the x-axis and y-axis residual network, respectively;
3. The set of near-field codewords that require additional test \(B'\) can be obtained by (38) and (39);
4. The XL-RIS performs additional beam test and \(r^N\) can be obtained by (40);
5. The index of the optimal near-field codeword \(s^*\) can be obtained by (42);
6. The optimal near-field codeword \(\hat{w}^*\) can be obtained by (43)
7. **return** \(\hat{w}^*\);

**IV. SIMULATION RESULTS**

In order to evaluate the performances of the two DRN-based near-field beam training schemes, numerical simulations are implemented, and the simulation results are presented in this section. Firstly, the parameters of the simulation system are presented, which include the parameters of the channel, the codebook and the neural networks. Then, the simulation results of the proposed scheme are given and compared with the existing schemes. In addition, three performance metrics are proposed to assess the performances of beam training schemes.

**A. System Setup**

In our simulations, XL-RIS-assisted mmWave communication systems are considered, where the number of antennas of the BS and the number of elements of the XL-RIS are set to \(M = 512\) and \(N = 512\), respectively. The number of channel paths from the BS to the XL-RIS and from the XL-RIS to the user are set to \(L_G = 3\) and \(L_k = 3\) respectively, where the channel gains of the strongest path are set to \(\alpha_{0}^G, \alpha_{0}^k \sim CN(0, 1)\) and the gains of the remaining channels are set to \(\alpha_{l}^G, \alpha_{l}^k \sim CN(0, 0.001)\) for \(l_1, l_2 = 2, 3\). Furthermore, the range of the user’s plane coordinate \((x_l, y_l)\) is given by \((50 m, 50 m)\) and \((30 m, 30 m)\), which are within the limits of the near-field domain. The sampling intervals of the near-field RIS codebook in the x-axis and y-axis are set to \(\Delta_x = 1m\) and \(\Delta_y = 1m\), which results in the numbers of sampling points being \(S_x = 100\) and \(S_y = 60\), respectively. Consequently, the number of candidate codewords in the near-field XL-RIS codebook reaches \(S_xS_y = 6,000\), which is far more than the number of far-field codewords \(N = 512\). The transmit SNR is defined as \(10 \log \left( \frac{1}{\alpha} \right)\), where the energy of the transmit signal is assumed to be 1. Finally, the selection interval for the PNBT scheme is fixed to \(D = 20\), which means that only 1 in 20 codewords in the PTBT scheme is tested. Based on the above channel and codebook parameters, we generate 20,000 channel samples and corresponding labels, of which 90 percentage are used for training the neural network and 10 percentage for evaluating the performance.

For the proposed neural network, 18-layer and 50-layer residual networks, i.e. ResNet18 and ResNet50, are adopted, and the specific structural parameters are shown in Table II. In order to accelerate the convergence of the neural networks, the learning rate decay strategy is employed, where the initial learning rate is set to 0.005, and the learning rate decays by half when the estimation accuracy does not improve within two training epochs. Moreover, the mini-batch strategy is also adopted, where the neural network is trained for 120 epochs with 2000 batches in each epoch. Fig. 7 shows the variation of the loss function with training epochs during the training phase. The loss gradually decreases and flattens out at approximately epoch 40. Then, the model parameters of the well-trained x-axis network and y-axis network are saved and used for subsequent performance tests.

**B. Metrics and Baselines**

In order to better evaluate and compare the performances of different near-field beam training schemes, three performance metrics are adopted.
are the most straightforward approach, of which the most schemes, the following five baselines are adopted: 

1) Exhaustive Beam Search: The exhaustive beam search are the most straightforward approach, of which the most representative schemes are sweeping scheme and maximum likelihood (ML) scheme. For the exhaustive search schemes, the XL-RIS requires to test all codewords and then the BS selects the codeword corresponding to the received signal with the highest energy or the codeword corresponding to the maximum probability. The exhaustive search schemes tend to provide the best performance but with an

2) Hierarchical Beam Search: The authors of [25] proposed a hierarchical search scheme for the XL-RIS, and a multi-layer codebook was employed, where the range of the higher-level codewords can cover several lower-level codewords. In the hierarchical search scheme, the optimal higher-level codeword is obtained by the exhaustive test, and then the lower-level codewords covered by that codeword are exhaustively tested to find the optimal lower-level codeword.

3) DL-based beam training in [23]: [23] proposed a beam training scheme for the far-field mmWave systems, where the fully connected (FC) neural network was employed. We adopt the same network structure as well as the beam training strategy and apply them to the near-field XL-RIS-assisted wireless communication system.

4) DL-based beam training in [20]: A CNN based far-field beam training scheme was proposed in [20]. Similarly, the same network structure and a close beam training strategy will be applied to the near-field for the XL-RIS, after which its performance will be compared with that of the proposed schemes in this paper.

5) Orthogonal matching pursuit (OMP) algorithm: The classical compressive sensing (CS) algorithm is widely used for the estimation of millimetre wave channels owing to its accuracy and low pilot overhead [26], [39]. Therefore, we decide to adopt the OMP algorithm, which is one of the CS algorithms that exploit the sparsity of the angular domain, to recover the channel and form the phase shift matrix of the XL-RIS for downlink transmission with the maximum ratio combining (MRC) principle. Specifically, by using the OMP algorithm and the codebook of the near-field XL-RIS in 14, the $\mathbf{h}_k^c$ can be obtained, which is the estimated value of the cascaded channel $\text{diag}(\mathbf{f}_k^c)\mathbf{g}_k$. Then, the phase shift matrix of the XL-RIS is set to $\Phi = \text{diag}(\phi)$, where $\phi$ is set as $\phi = ||\mathbf{h}_k^c||_2$. Finally, the achievable rate obtained after implementing the OMP algorithm can be expressed as

$$E_k = \log_2 \left(1 + \frac{\phi^T \mathbf{h}_k^c}{\sigma^2} \right), \quad (47)$$

C. Simulation Result

1) How Is the Performance of the Proposed Solution?:
In Fig. 8, the performance of the proposed scheme in terms of the normalized RIS beam gain is compared at different transmission SNRs. From Fig. 8, it can be seen that at higher transmit SNRs, e.g. SNR > 5dB, the improved PNBT scheme with $K = 2$ and $L = 5$ can achieve more than 90% of the perfect performance and even outperforms the sweeping scheme, which demonstrates the powerful non-linear relationship learning capability of the residual network. Furthermore, the FBT scheme can also have a similar beam gain performance to the sweeping scheme at different transmission SNRs. It can
also be observed that the FBT scheme outperforms the PNBT scheme due to the fact that the FBT scheme tests more codewords, and the neural network can extract more information about the channel. However, with additional beam tests being implemented, the improved PNBT scheme gradually performs better than the FBT scheme because the knowledge of the probability distribution vector is fully exploited.

2) How Does the Proposed Scheme Compare With the Existing Schemes?: Fig. 9 compares the performances of the proposed schemes in terms of the achievable rate, where the existing beam training scheme is also included for comparison. As shown in Fig. 9, both the proposed near-field beam training schemes exhibit similar performance to the sweeping scheme and the ML algorithm in terms of the achievable rate at all transmit SNRs; however, the pilot overhead is significantly less than that of the sweeping scheme and the ML algorithm. Furthermore, the proposed two near-field beam training schemes provide a significant performance improvement over the near-field hierarchical beam scheme because the hierarchical beam training is highly sensitive to noise and lacks the ability to handle it like the neural network. It can be also seen from Fig. 9 that the beam training algorithm proposed in this paper can more accurately choose the optimal beam to obtain a higher achievable rate, compared with the conventional OMP estimation algorithms. In fact, since there is no codebook sampling design specifically applicable to the XL-RIS cascade channel, which can degrade the correlation between codewords, the CS-based channel estimation algorithm suffers a severe performance loss. As a result, the existing conventional estimation techniques can be hardly applied directly to the XL-RIS-assisted communication system, where both the spherical wave model and the characteristics of the cascade channel have to be considered. At the same time, the proposed schemes also outperform two ML-based beam training schemes, which are mainly suitable for the far-field domain and have difficulty in addressing the large amount of multi-dimensional near-field codewords.

Next, the performances of the different schemes in terms of achievable rate and pilot overhead, i.e. the effective achievable rate, are presented in Fig. 10. As shown in Fig. 10, although
the sweeping scheme offers a high achievable rate performance, it suffers from a high pilot overhead, which results in a dramatic reduction in data communication time, as confirmed by its low effective achievable rate. For the proposed near-field beam training scheme, the FBT and PNBT schemes reduce the pilot overhead by 91% and 95%, compared to the sweeping scheme, where they only test 512 and 300 codewords, respectively. Moreover, the improved PNBT scheme with \( K = 2 \) and \( L = 5 \) only needs to test \( I + KL = 310 \) codewords, thus reducing the pilot overhead by approximately 95%. Consequently, the proposed near-field beam training scheme achieves excellent performance in terms of the effective achievable rates.

3) How About the Impact of the Number of Codewords for Additional Test?: Fig. 11 illustrates the impact of different values of \( K \) and \( L \) on the normalized RIS beam gain of the improved PNBT scheme. It can be seen that the performance of the improved PNBT scheme improves as \( K \) and \( L \) increase because the knowledge of the probability distribution vectors is increasingly exploited. Additionally, the performance of the improved PNBT scheme tends to converge and approaches the perfect performance when \( K > 3 \) and \( L > 3 \), which means that the probability distribution vector has been fully exploited.

4) How Is the Performance of Different Neural Network Structures?: In Fig. 12, we focus on the impact of different neural network structures on the performance of the proposed near-field beam training scheme. Specifically, we employ different neural network structures with the same beam training strategy and compare their performances. It can be seen that the residual neural network-based near-field beam training schemes achieve more significant RIS beam gain than the conventional CNN, which further demonstrates the superiority of residual networks. Additionally, with the increase of the residual network from 18 to 50 layers, the performance of the residual neural network-based near-field beam training scheme does not degrade like the traditional convolutional network but further improves, which also proves the superior performance of the residual structure in solving the gradient disappearance problem.

5) Why We Set the Sampling Interval \( D \) to 20?: Fig. 13 and Fig. 14 compare the performances of the proposed PNBT scheme with the improved PNBT scheme in terms of achievable rate and the effective achievable rate at different sampling intervals, respectively, where the Resnet18 network is employed and the transmit SNR is set to 20. From Fig. 13, it can be seen that the achievable rate of the PNBT scheme decreases correspondingly as the sampling interval increases, which is due to the fact that an increase in the sampling interval reduces the input to the neural network and thus the information acquired is reduced. Furthermore, the improved PNBT scheme receives less impact when the sampling interval is increased, especially when \( K = 2 \) and \( L = 5 \), the achievable rate of the improved PNBT scheme remains more or less the same and close to that of the sweeping scheme. In terms of the effective achievable rate, both the PNBT scheme and the improved PNBT scheme improve as the sampling interval increases, which is because the amount of pilot overhead is
Fig. 14. Effective achievable rate for the proposed schemes with different intervals $D$.

Fig. 15. Normalized SNR of the proposed schemes with different neural network structures.

significantly reduced. Moreover, when the sampling interval reaches 20, the effective reachable rate of both the PNBT scheme and the improved PNBT scheme reaches their maximum value and starts to converge, which means that the pilot overhead and achievable rate are in excellent balance. Therefore, we set the sampling interval to 20 in order to achieve a good performance of the proposed schemes.

6) How Is the Pilot Overhead of the Proposed Scheme?:

Fig. 15 compares the pilot overhead of the proposed schemes with the existing beam training schemes. It can be seen that the proposed schemes significantly reduce the pilot overhead in the near-field XL-RIS-assisted wireless communication systems. When the sampling interval is 20, the proposed schemes are capable of reducing the pilot overhead by 91% or even 95%, and less than that of the hierarchical search scheme.

V. CONCLUSION

In this paper, we proposed two deep learning-based near-field beam training schemes in XL-RIS-assisted wireless communication systems to reduce the pilot overhead, where the deep residual networks are employed to determine the optimal near-field RIS codeword. Our proposed scheme mainly utilizes neural networks to extract the implicit information between the received signals of different codewords, thus reducing the reliance on the information obtained from the codeword test. Specifically, we first proposed a FBT scheme in which the received signals of all far-field RIS codewords are fed into the neural network to estimate the optimal near-field RIS codeword. In order to further reduce the pilot overhead, the PNBT scheme was proposed, where only the received signals corresponding to the partial near-field RIS codewords were served as input to the neural network. Furthermore, we further proposed the improved PNBT scheme to enhance the performance of the beam training by fully exploring the neural network’s output. Note that in both schemes a dual neural network structure was proposed, i.e., an x-axis network and a y-axis network to estimate the index of the two dimensions of the optimal codeword. Simulation results showed that our proposed scheme can approach the performance of the sweeping scheme in terms of the achievable rate, however, the pilot overhead was significantly less than that of the sweeping scheme. Furthermore, the proposed schemes are more compatible with the XL-RIS-assisted wireless communication systems than the existing beam-training schemes, where better performance was obtained and the pilot overhead was reduced.

REFERENCES

[1] C. Pan et al., “Reconfigurable intelligent surfaces for 6G systems: Principles, applications, and research directions,” IEEE Commun. Mag., vol. 59, no. 6, pp. 14–20, Jun. 2021.
[2] C. Pan et al., “Intelligent reflecting surface aided MIMO broadcasting for simultaneous wireless information and power transfer,” IEEE J. Sel. Areas Commun., vol. 38, no. 8, pp. 1719–1734, Aug. 2020.
[3] C. Pan et al., “Multicell MIMO communications relying on intelligent reflecting surfaces,” IEEE Trans. Wireless Commun., vol. 19, no. 8, pp. 5218–5233, Aug. 2020.
[4] S. Shen, B. Clerckx, and R. Murch, “Modeling and architecture design of reconfigurable intelligent surfaces using scattering parameter network analysis,” IEEE Trans. Wireless Commun., vol. 21, no. 2, pp. 1229–1243, Feb. 2022.
[5] T. J. Cui, M. Q. Qi, X. Wan, J. Zhao, and Q. Cheng, “Coding metamatamaterials, digital metamaterials and programmable metamaterials,” Light. Sci. Appl., vol. 3, no. 10, p. e218, Oct. 2014.
[6] 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. (ICASSP), May 2019, pp. 4659–4663.
[7] T. L. Jensen and E. De Carvalho, “An optimal channel estimation scheme for intelligent reflecting surfaces based on a minimum variance unbiased estimator,” in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), May 2020, pp. 5000–5004.
[8] Q. Nadeem, H. Alwazani, A. Kamoun, A. Chaaban, M. Debbah, and M. Alouini, “Intelligent reflecting surface-assisted multi-user MISO communication: Channel estimation and beamforming design,” IEEE Open J. Commun. Soc., vol. 1, pp. 661–680, 2020.
[9] G. Zhou, C. Pan, H. Ren, P. Popovski, and A. L. Swindlehurst, “Channel estimation for RIS-aided multiuser millimeter-wave systems,” IEEE Trans. Signal Process., vol. 70, pp. 1478–1492, 2022.
[10] P. Wang, J. Fang, H. Duan, and H. Li, “Compressed channel estimation for intelligent reflecting surface-assisted millimeter wave systems,” IEEE Signal Process. Lett., vol. 27, pp. 905–909, 2020.
[11] X. Wei, D. Shen, and L. Dai, “Channel estimation for RIS assisted wireless communications—Part II: An improved solution based on double-structured sparsity,” IEEE Commun. Lett., vol. 25, no. 5, pp. 1403–1407, May 2021.
[12] C. You, B. Zheng, and R. Zhang, “Fast beam training for IRS-assisted multiuser communications,” IEEE Wireless Commun. Lett., vol. 9, no. 11, pp. 1845–1849, Nov. 2020.

[13] J. Zhang, Y. Huang, J. Wang, X. You, and C. Masouros, “Intelligent interactive beam training for millimeter wave communications,” IEEE Trans. Wireless Commun., vol. 20, no. 3, pp. 2034–2048, Mar. 2021.

[14] B. Ning, Z. Chen, W. Chen, Y. Du, and J. Fang, “Terahertz multiuser massive MIMO with intelligent reflecting surface: Beam training and hybrid beamforming,” IEEE Trans. Veh. Technol., vol. 70, no. 2, pp. 1376–1393, Feb. 2021.

[15] T. Bai and R. W. Heath Jr., “Analysis of beam sweep channel estimation in mmWave massive MIMO networks,” in Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP), Dec. 2016, pp. 615–619.

[16] Z. Xiao, T. He, P. Xia, and X. Xia, “Hierarchical codebook design for beamforming training in millimeter-wave communication,” IEEE Trans. Wireless Commun., vol. 15, no. 5, pp. 3380–3392, May 2016.

[17] Z. Xiao, H. Dong, L. Bai, P. Xia, and X. 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] 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.

[19] A. Alkhateeb, S. Alex, P. Varkey, Y. Li, Q. Qu, and D. Tujkovic, “Deep learning coordinated beamforming for highly-mobile millimeter wave systems,” IEEE Access, vol. 6, pp. 37328–37348, 2018.

[20] K. Ma, D. He, H. Sun, Z. Wang, and S. Chen, “Deep learning assisted calibrated beam training for millimeter-wave communication systems,” IEEE Trans. Commun., vol. 69, no. 10, pp. 6706–6721, Oct. 2021.

[21] K. Ma, D. He, H. Sun, and Z. Wang, “Deep learning assisted mmWave beam prediction with prior low-frequency information,” in Proc. IEEE Int. Conf. Commun., Jun. 2021, pp. 1–6.

[22] K. Ma, P. Zhao, and Z. Wang, “Deep learning assisted beam prediction using out-of-band information,” in Proc. IEEE 91st Veh. Technol. Conf. (VTC-Spring), May 2020, pp. 1–5.

[23] C. Qi, Y. Wang, and G. Y. Li, “Deep learning for beam training in millimeter wave massive MIMO systems,” IEEE Trans. Wireless Commun., early access, Sep. 22, 2020, doi: 10.1109/TWC.2020.3024279.

[24] W. Tang et al., “Wireless communications with reconfigurable intelligent surface: Path loss modeling and experimental measurement,” IEEE Trans. Wireless Commun., vol. 20, no. 1, pp. 421–439, Jan. 2021.

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

[26] 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.

[27] H. Zhu and J. Wang, “Chunk-based resource allocation in OFDMA systems—Part II: Joint chunk, power and bit allocation,” IEEE Trans. Commun., vol. 60, no. 2, pp. 499–509, Feb. 2012.

[28] 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, pp. 5394–5409, Nov. 2019.

[29] Q.-U.-A. Nadeem, H. Alwazani, A. Kammoun, A. Chaaban, M. Debbah, and M.-S. Alouini, “Intelligent reflecting surface-assisted multi-user MISO communication: Channel estimation and beamforming design,” IEEE Open J. Commun. Soc., vol. 1, pp. 661–680, 2020.

[30] 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.

[31] Z. Zhou, X. Gao, J. Fang, and Z. Chen, “Spherical wave channel and analysis for large linear array in LoS conditions,” in Proc. IEEE GlobalSIP Workshops, Dec. 2015, pp. 1–6.

[32] C. Hu, L. Dai, T. Mir, Z. Gao, and J. Fang, “Super-resolution channel estimation for mmWave massive MIMO with hybrid precoding,” IEEE Trans. Veh. Technol., vol. 67, no. 9, pp. 8954–8958, Sep. 2018.

[33] W. Roh et al., “Millimeter-wave beamforming as an enabling technology for 5G cellular communications: Theoretical feasibility and prototype results,” IEEE Commun. Mag., vol. 52, no. 2, pp. 106–113, Feb. 2014.

[34] A. L. Swindlehurst, E. Ayanoglu, P. Heydari, and F. Capolino, “Millimeter-wave massive MIMO: The next wireless revolution?” IEEE Trans. Commun., vol. 52, no. 9, pp. 56–62, Sep. 2014.

[35] J. Sherman, “Properties of focused apertures in the Fresnel region,” JRE Trans. Antennas Propag., vol. 10, no. 4, pp. 399–408, Jul. 1962.

[36] H. Yan, S. Jin, C. Wen, and X. Ma, “Channel estimation for extremely large-scale massive MIMO systems,” IEEE Wireless Commun. Lett., vol. 9, no. 5, pp. 633–637, May 2020.

[37] X. Wei, C. Hu, and L. Dai, “Deep learning for beamspace channel estimation in millimeter-wave massive MIMO systems,” IEEE Trans. Commun., vol. 69, no. 1, pp. 182–193, Jan. 2021.

[38] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 770–777.

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

Wang Liu received the B.S. degree from the School of Electronics and Information Engineering, Nanjing University of Information Science and Technology, Nanjing, China, in 2021. He is currently pursuing the Ph.D. degree with the School of Information Science and Engineering, Southeast University, Nanjing. His research interests include massive MIMO, reconfigurable intelligent surfaces (RIS), and machine learning.

Cunhua Pan (Senior Member, IEEE) received the B.S. and Ph.D. degrees from the School of Information Science and Engineering, Southeast University, Nanjing, China, in 2010 and 2015, respectively.

From 2015 to 2016, he was a Research Associate with the University of Kent, U.K. From 2016 to 2018, he was a Post-Doctoral Researcher with the Queen Mary University of London, U.K., where he was a Lecturer from 2019 to 2021. Since 2021, he has been a Full Professor with Southeast University. He has published more than 120 IEEE journal articles. His research interests include reconfigurable intelligent surfaces (RIS), intelligent reflection surface (IRS), ultra-reliable low latency communication (URLLC), machine learning, 5G, the Internet of Things, and mobile edge computing. He is currently an Editor of IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, IEEE WIRELESS COMMUNICATIONS LETTERS, IEEE COMMUNICATIONS LETTERS and IEEE ACCESS. He serves as the Guest Editor for IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS on the Special Issue on xURLLC in 6G: Next Generation Ultra-Reliable and Low-Latency Communications. He also serves as the Leading Guest Editor for the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING (JSTSP) Special Issue on Advanced Signal Processing for Reconfigurable Intelligent Surface-Aided 6G Networks, the IEEE Vehicular Technology Magazine on the Special Issue on Backscatter and Reconfigurable Intelligent Surface Empowered Wireless Communications in 6G, the IEEE OPEN JOURNAL OF VEHICULAR TECHNOLOGY on the Special Issue of Reconfigurable Intelligent Surface Empowered Wireless Communications in 6G and Beyond, and IEEE ACCESS Special Issue on Reconfigurable Intelligent Surface Aided Communications for 6G and Beyond. He is a Workshop Organizer in IEEE ICC 2021 on the Topic of Reconfigurable Intelligent Surfaces for Next Generation Wireless Communications (RIS for 6G Networks) and a Workshop Organizer in IEEE Globecom 2021 on the Topic of Reconfigurable Intelligent Surfaces for Future Wireless Communications. He is currently the Workshops and Symposia Officer of Reconfigurable Intelligent Surfaces Emerging Technology Initiative. He is the Workshop Chair of IEEE WCNC 2024 and the TPC Co-Chair for IEEE ICC 2022. He serves as a TPC Member for numerous conferences, such as ICC and GLOBECOM and the Student Travel Grant Chair for ICC 2019. He received the IEEE ComSoc Leonard G. Abraham Prize in 2022 and IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award in 2022.
Hong Ren (Member, IEEE) received the B.S. degree in electrical engineering from Southwest Jiaotong University, Chengdu, China, in 2011, and the M.S. and Ph.D. degrees in electrical engineering from Southeast University, Nanjing, China, in 2014 and 2018, respectively. From 2016 to 2018, she was a Visiting Student with the School of Electronics and Computer Science, University of Southampton, U.K. From 2018 to 2020, she was a Post-Doctoral Scholar with the Queen Mary University of London, U.K. She is currently an Associate Professor with Southeast University. Her research interests include communication and signal processing, including ultra-low latency and high reliable communications, massive MIMO, and machine learning. She received the IEEE ComSoc Leonard G. Abraham Prize in 2022 and the IEEE WCSP 2022 Best Paper Award.

Feng Shu (Member, IEEE) received the Ph.D. degree in information and communication engineering from Southeast University, China, in 2002. From September 2009 to September 2010, he was a Visiting Post-Doctoral Researcher with The University of Texas at Dallas, Richardson, TX, USA. From October 2005 to November 2020, he was with the School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing, where he was promoted from an Associate Professor to a Full Professor of supervising Ph.D. students in 2013. Since December 2020, he has been with the School of Information and Communication Engineering, Hainan University, Haikou, China, where he is currently a Professor and a Supervisor of Ph.D. and graduate students. He has authored or coauthored more than 300 articles in archival journals with more than 150 articles in IEEE journals and 210 SCI-indexed papers. His citations are 5655. He holds more than 36 Chinese patents and also a PI or Co-PI for seven national funding projects. His research interests include wireless networks, wireless location, array signal processing, and machine learning in those areas. He was awarded with the Leading-Talent Plan of Hainan Province in 2020, the Fujian Hundred-Talent Plan of Fujian Province in 2018, and the Mingjian Scholar Chair Professor in 2015. He was an Editor of the IEEE SYSTEMS JOURNAL from 2019 to 2021 and IEEE ACCESS from 2016 to 2018. He was an IEEE TRANSACTIONS ON COMMUNICATIONS Exemplary Reviewer in 2020. He is currently an Editor of the IEEE WIRELESS COMMUNICATIONS LETTERS.

Shi Jin (Senior Member, IEEE) received the B.S. degree in communications engineering from the Guilin University of Electronic Technology, Guilin, China, in 1996, the M.S. degree from the Nanjing University of Posts and Telecommunications, Nanjing, China, in 2003, and the Ph.D. degree in information and communications engineering from Southeast University, Nanjing, in 2007. From June 2007 to October 2009, he was a Research Fellow with the University College London, Adastra Park Research Campus, London, U.K. He is currently a Faculty Member with the National Mobile Communications Research Laboratory, Southeast University. His research interests include space time wireless communications, random matrix theory, and information theory. He and his coauthors have been awarded the 2011 IEEE Communications Society Stephen O. Rice Prize Paper Award in the field of communication theory, the 2010 Young Author Best Paper Award by the IEEE Signal Processing Society, and the 2022 Best Paper Award. He was an Associate Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE COMMUNICATIONS LETTERS, and IET Communications. He is serving as an Area Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and IET Electronics Letters.

Jiangzhou Wang (Fellow, IEEE) is currently a Professor with the University of Kent, U.K. He has published more than 400 articles and four books. His research interests include mobile communications. He is a fellow of the Royal Academy of Engineering (U.K.) and IET. He was a recipient of the 2022 IEEE Communications Society Leonard G. Abraham Prize and the IEEE Globecon2012 Best Paper Award. He was the Technical Program Chair of the 2019 IEEE International Conference on Communications (ICC2019), Shanghai, the Executive Chair of the IEEE ICC2015, London, and the Technical Program Chair of the IEEE WCNC2013. He is/was an Editor of a number of international journals, including IEEE TRANSACTIONS ON COMMUNICATIONS from 1998 to 2013.