Coded Beam Training for RIS Assisted Wireless Communications

Yuhao Chen, *Graduate Student Member, IEEE* and Linglong Dai, *Fellow, IEEE*

Abstract—Reconfigurable intelligent surface (RIS) is considered as one of the key technologies for future 6G communications. To fully unleash the performance of RIS, accurate channel state information (CSI) is crucial. Beam training is widely utilized to acquire the CSI. However, before aligning the beam correctly to establish stable connections, the signal-to-noise ratio (SNR) at UE is inevitably low, which reduces the beam training accuracy. To deal with this problem, we exploit the coded beam training framework for RIS systems, which leverages the error correction capability of channel coding to improve the beam training accuracy under low SNR. Specifically, we first extend the coded beam training framework to RIS systems by decoupling the base station-RIS channel and the RIS-user channel. For this framework, codewords that accurately steer to multiple angles is essential for fully unleashing the error correction capability. In order to realize effective codeword design in RIS systems, we then propose a new codeword design criterion, based on which we propose a relaxed Gerchberg-Saxton (GS) based codeword design scheme by considering the constant modulus constraints of RIS elements. In addition, considering the two dimensional structure of RIS, we further propose a dimension reduced encoder design scheme, which can not only guarentee a better beam shape, but also enable a stronger error correction capability. Simulation results reveal that the proposed scheme can realize effective and accurate beam training in low SNR scenarios.

Index Terms—RIS, beam training, channel coding, codeword design.

I. INTRODUCTION

Reconfigurable intelligent surface (RIS) is considered as a promising technology for future 6G wireless communications [1]. Thanks to the numerous low-cost reflecting elements, RIS can control the electromagnetic environment intelligently with low power consumption [2]–[4]. By properly controlling the phase shifts of RIS elements, directional beams with high array gain could be generated by beamforming to extend the signal coverage and improve the channel capacity [5], [6]. In order to realize effective beamforming so as to fully leverage the potential benefits of RIS, accurate channel state information (CSI) is essential [7], [8].

The CSI can be obtained by either explicit channel estimation or implicit beam training. For explicit channel estimation, since the elements on RIS can only reflect the incident signals, the base station (BS) needs to estimate the cascaded channel (the composite of user-RIS channel and RIS-BS channel) [9],

[10]. The size of the cascaded channel is the product of the number of BS antennas and the number of RIS elements. With the large number of RIS elements needed to generated high-gain beams, the size of cascaded channel is usually large, leading to an unacceptable pilot overhead for channel estimation. To avoid estimating the large cascaded channel matrix, the implicit beam training, which only aims to determine the angles of RIS and user equipment (UE), can be utilized. By searching the space with a series of pre-defined codewords, the angles of RIS and UE can be estimated based on the received power, according to which the beams at BS and RIS can be correctly aligned to UE.

A. Prior Works

To determine the angles of RIS and UE, the most intuitive scheme is the exhaustive beam training [11], [12]. During beam training, BS and RIS both generate narrow beams to sequentially search all possible angles in space. After transmitting all candidate beams, the angles of RIS and UE can be obtained by selecting the beams with the maximum received power. For this scheme, both BS and RIS generate high-gain narrow beams, so the angles can be accurately estimated. However, for this scheme, the number of candidate beams equals to the product of the number of BS antennas and the number of RIS elements. With the large number of BS antennas and RIS elements, the candidated beams are massive, resulting in an overwhelming beam training overhead.

In order to reduce the beam training overhead, researchers have developed various low overhead hierarchical beam training schemes. Existing works can be divided into two categories. The first category is single user hierarchical beam training, where the beam training overhead can be reduced by excluding a large range of impossible angles so as to narrow down the searching range effectively.

For example, in [13], beams generated by lower-layer codewords cover a wider range of angles compared with beams generated by higher-layer codewords. During beam training, lower-layer codewords are firstly transmitted. After transmitting the codewords in a certain layer, the index of the beam with the maximum received power is fed back to the BS and RIS, and the codewords in the next layer are decided accordingly. As the layer grows higher, the searching range gradually narrows down and the angular resolution increases continuously. After the highest-layer search, the angles of RIS and UE are then determined. By this means, a large range of wrong angles is excluded in lower-layer search, which avoids a lot of unnecessary high-resolution search and thus reduces the beam training overhead. However, since the higher-layer

The authors are with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China, and also with the Beijing National Research Center for Information Science and Technology (BNRist), Beijing 100084, China. (e-mails: chen-yh21@mails.tsinghua.edu.cn; daill@tsinghua.edu.cn).

codewords is determined by the result of lower-layer training, this category of schemes requires frequent feedback between UE and BS/RIS, which brings extra burden to the RIS systems. Moreover, since the searching range in a certain layer may vary for different UEs, it is hard to extend this category of schemes to multi-user systems, which severely limits the application of such schemes.

To realize effective beam training in multi-user systems, for the second category, all possible angles in space are divided into several disjoint subsets, and the beam generated by each codeword covers the angles in a certain subset simultaneously. After scanning the entire space by these codewords, the subset of the beam with the maximum received power is recorded. Then, the space is divided in a different way and the corresponding scanning and recording are conducted again. After a few rounds of scanning, the angles of RIS and multiple UEs can be determined independently based on the intersections of all recorded subsets.

Following this idea, researchers have studied several effective beam training schemes [14]–[17]. Specifically, authors in [14], [15] divided the whole space for every round of scanning in a random/hashed way. Authors in [16] further extend this hashing scheme to multi-RIS scenarios. By assigning different hashing functions to different RISs, the angles of different RISs can be simultaneously determined. For the above schemes, the choice of hashing function may affect the beam training accuracy, thus leading to an unstable performance. To deal with this, authors in [17] studied a full-coverage hierarchical beam training scheme. Different from the single user hierarchical beam training of the first category, in each layer, the beams also cover the whole space, but the angular resolution gradually increases. For this category of schemes, each beam can search multiple angles simultaneously, so the beam training overhead can also be reduced. In addition, since how to divide the entire space in each round does not depend on previous results, no extra feedback is needed during beam training, which makes this category of schemes more adaptive to different communication scenarios than the first category.

Unfortunately, before aligning the beam correctly to establish stable connections among BS, RIS and UE, the signal-to-noise ratio (SNR) at UE is inevitably low. What's worse, in RIS systems, there exists the "multiplicative fading" effect [18], which means the equivalent path loss of the BS-RIS-UE link is the product of the path loss of BS-RIS link and the path loss of RIS-UE link. Meanwhile, both categories of low overhead beam training schemes need to generate beams that cover a large range of angles, leading to a relatively low beamforming gain. These facts will result in a significantly low SNR at UE. As a result, the codeword may be mischosen, which leads to the "error propagation" phenomenon and greatly reduces the beam training accuracy. Therefore, how to realize accurate beam training in RIS systems under poor SNR conditions is crucial for the practical deployment of RIS in future communications.

B. Our Contributions

To improve the beam training accuracy under poor SNR scenarios, in this paper, we exploit the coded beam training framework in RIS systems. By applying the idea of channel coding in the beam training process, we can leverage the error correction capability of channel coding to enhance the reliability of beam training under low SNR[1]. The specific contributions are listed as follows.

- First, inspired by the coded beam training framework that is recently studied by us in multiple-input multiple-output (MIMO) systems [19], we design a coded beam training framework for RIS systems. Specifically, we map the angles in space to different beam patterns in space through the encoding function. Based on the intended beam patterns, we design the codewords, which is the foundation of the designed framework. After sequentially transmitting all codewords in the codebook, the UE can obtain the encoded transmitting sequence based on the received powers. Then, the decoding function is utilized to decode the received sequence and the angles of RIS and UE can be estimated. Thanks to the error correction capability of the encoding-decoding process, the error caused by low SNR during beam training can be corrected and the beam training accuracy improves accordingly.

- The most significant difference between the codeword design of RIS and the codeword design for the coded beam training framework in [19] is that, RIS is subject to constant modulus constraint, which makes it hard to generate ideal beams that cover a variety of angles. One of the efficient codeword design schemes is Gerchberg-Saxton (GS) based codeword design scheme [20]. To adapt the GS-based scheme to RIS, we first clarify that the criterion of minimizing the difference between the intended beam shape and the generated beam shape is actually unsuitable for beam training. Then, we propose that the criterion for codeword design should be distinguishing between the angles within the intended angle coverage range and the angles out of the angle coverage range. Based on this new criterion, we propose a relaxed GS-based codeword design scheme so as to improve the beam shape accuracy.

- Apart from the constant modulus constraint, the structure of RIS is usually a 2-dimensional (2D) uniform planar array (UPA), which is also different from the uniform linear array (ULA) considered in [19]. The 2D structure leads to a poor orthogonality for the spatial sampling matrix in the proposed relaxed GS-based codeword design scheme, which also results in a non-ideal beam shape. To deal with this problem, we further propose a dimension reduced encoder design scheme. By decoupling the 2D codeword design problem into two 1D codeword design problems, the spatial sampling matrix degenerates to the 1D case and possesses a good orthogonality, thus improving the quality of the beam shape design. Moreover, since the encoder decouples the two dimensions of RIS, the error correction capability can also be improved. Then, we compare the necessary beam training overheads of the proposed framework and existing frameworks. Finally,

---

[1]Simulation codes will be provided to reproduce the results in this article: http://oa.ee.tsinghua.edu.cn/dailinglong/publications/publications.html.
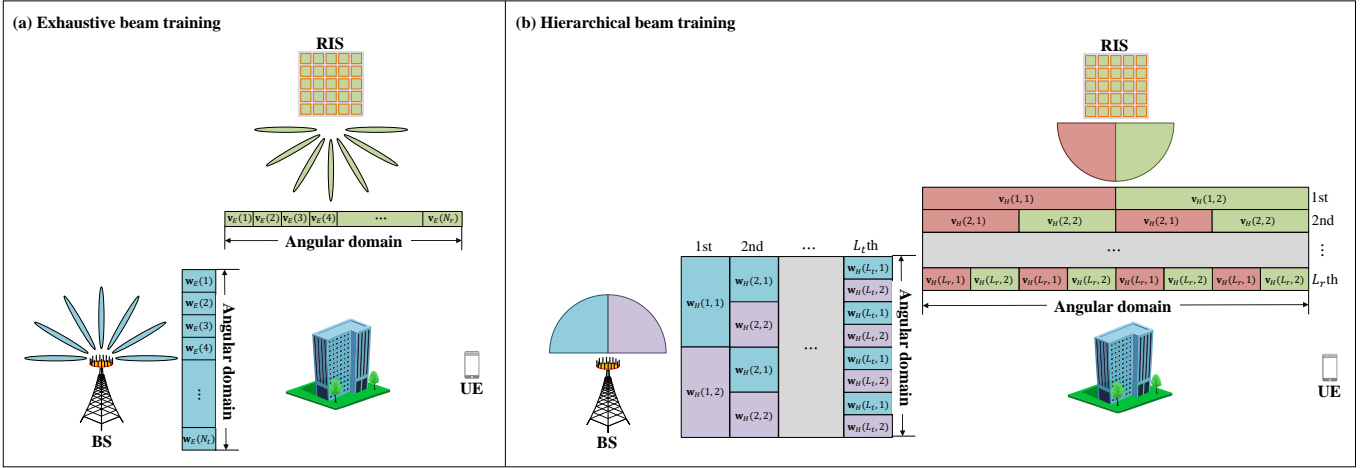
Fig. 1. Traditional beam training frameworks. (a) Exhaustive beam training; (b) Hierarchical beam training.

simulation results reveal that the proposed framework can realize efficient beam training in low SNR scenarios.

C. Organization and Notation

Organization: The remainder of this paper is organized as follows. In Section II, we first introduce the system model. Then, the traditional exhaustive beam training framework and hierarchical beam training framework are elaborated. In Section III, we introduce the proposed coded beam training framework in RIS systems. In Section IV, we first introduce the proposed relaxed GS-based codeword design scheme, followed by the proposed dimension reduced encoder design scheme. Then, the necessary beam training overheads for the proposed framework and traditional frameworks are analyzed. Simulation results are provided in Section V, and conclusions are finally drawn in Section VI.

Notation: Lower-case and upper-case boldface letters represent vectors and matrices, respectively; $\mathbf{v}(i)$ denotes the i-th element of the vector $\mathbf{v}$; $\mathbf{X}(i, j)$ denotes the (i, j)-th element of the matrix $\mathbf{X}$; $\mathbf{X}(i, :)$ and $\mathbf{X}(:, j)$ denote the i-th row and the j-th column of the matrix $\mathbf{X}$; $(\cdot)^T$ and $(\cdot)^H$ denote the transpose and conjugate transpose, respectively; $|\cdot|$ denotes the absolute operator; $\|\cdot\|_2$ denotes the l_2 norm operator; $\lceil \cdot \rceil$ denotes the ceiling operator; $\mathrm{mod}(\cdot)$ denotes the modulo operator; $\mathcal{CN}(\mu, \Sigma)$ and $\mathcal{U}(a, b)$ denote the Gaussian distribution with mean μ and covariance Σ, and the uniform distribution between a and b, respectively.

II. System Model and Background

In this section, we first introduce the system model of RIS assisted communication systems. Then, the traditional exhaustive and hierarchical beam training frameworks are reviewed.

A. System Model

We consider a downlink time division duplexing (TDD) RIS assisted communication system in this paper. The BS employs a uniform linear array (ULA) with N_t antennas and the RIS

employs a UPA with $N_{r_1} \times N_{r_2} = N_r$ antennas. The UE is equipped with a single antenna. We assume that the direct links between the BS and the UE are blocked by obstacles such as trees or buildings [10]. Then, the received signal $y_p \in \mathbb{C}$ in the p-th time slot at the UE can be represented as

$$y_p = \mathbf{h}_r \mathrm{diag}(\mathbf{v}_p)\mathbf{G}\mathbf{w}_p s_p + n_p, \qquad (1)$$

where $\mathbf{h}_r \in \mathbb{C}^{1 \times N_r}$ denotes the channel between UE and RIS; $\mathbf{v}_p \in \mathbb{C}^{N_r \times 1}$ denotes the reflecting vector of RIS at the p-th time slot; $\mathbf{G} \in \mathbb{C}^{N_r \times N_t}$ denotes the channel between RIS and BS; $\mathbf{w}_p$ denotes the beamforming vector of BS at the p-th time slot; $s_p \in \mathbb{C}$ denotes the signal sent by BS at the p-th time slot; and $n_p \sim \mathcal{CN}(0, \sigma^2)$ denotes the additive white Gaussian complex noise at the p-th time slot with σ^2 being the noise power, respectively. Due to the constant modulus constraint, RIS can only adjust the phase shift rather than the amplitude coefficient [21]. As a result, the reflecting vector of the RIS can be re-written as $\mathbf{v}_p = \left[e^{j\vartheta_1}, e^{j\vartheta_2}, \cdots, e^{j\vartheta_{N_r}} \right]$, where $\vartheta_n \in [0, 2\pi], n = 1, 2, \cdots, N_r$ represents the phase shift of the n-th element.

For the channel model, we apply the Saleh-Valenzuela channel model [22], so the channel $\mathbf{h}_r$ between UE and RIS can be written as

$$\mathbf{h}_r = \sqrt{\frac{N_r}{L_r}} \sum_{\ell=1}^{L_r} \alpha_\ell^r \mathbf{a}^T(\phi_\ell^r, \theta_\ell^r), \qquad (2)$$

where L_r denotes the number of paths between UE and RIS; α_ℓ^r denotes the path gain of the ℓ-th path; $\phi_\ell^r, \theta_\ell^r$ denote the azimuth angle and the elevation angle at RIS, respectively. The steering vector $\mathbf{a}(\phi, \theta)$ for a $N = N_1 \times N_2$-antenna UPA can be elaborated as

$$\mathbf{a}(\phi, \theta) = \frac{1}{\sqrt{N}} \left[e^{-j2\pi d \sin(\phi)\sin(\theta)\boldsymbol{\delta}/\lambda} \right] \otimes \left[e^{-j2\pi d \cos(\theta)\boldsymbol{\varsigma}\lambda} \right], \qquad (3)$$

where $\lambda = c/f_c$ denotes the wavelength of electromagnetic wave with f_c being the central frequency and c being the speed of light. The antenna spacing d is set to $d = \lambda/2$. The

antenna indices $\boldsymbol{\delta}$ and $\boldsymbol{\varsigma}$ can be represented as

$$\boldsymbol{\delta} = [\delta_1, \delta_2, \cdots, \delta_{N_1}]^T = \left[\frac{1-N_1}{2}, \frac{3-N_1}{2}, \cdots, \frac{N_1-1}{2}\right]^T$$
$$\boldsymbol{\varsigma} = [\varsigma_1, \varsigma_2, \cdots, \varsigma_{N_2}]^T = \left[\frac{1-N_2}{2}, \frac{3-N_2}{2}, \cdots, \frac{N_2-1}{2}\right]^T. \quad (4)$$

Simularly, the channel $\mathbf{G}$ between RIS and BS can be represented as

$$\mathbf{G} = \sqrt{\frac{N_t N_r}{L_G}} \sum_{\ell=1}^{L_G} \alpha_\ell^G \mathbf{a}(\phi_\ell^{G_r}, \theta_\ell^{G_r}) \mathbf{b}^T(\phi_\ell^{G_t}), \quad (5)$$

where L_G denotes the number of paths between RIS and BS; α_ℓ^G denotes the path gain of the ℓ-th path; $\phi_\ell^{G_r}, \theta_\ell^{G_r}, \phi_\ell^{G_t}$ denote the azimuth angle at RIS, the elevation angle at RIS and the azimuth angle at BS, respectively. The steering vector $\mathbf{b}(\phi)$ for a N-antenna ULA can be elaborated as

$$\mathbf{b}(\phi) = \frac{1}{\sqrt{N}}\left[1, e^{-j2\pi d \sin(\phi)/\lambda}, \cdots, e^{-j2(N-1)\pi d \sin(\phi)/\lambda}\right]^T. \quad (6)$$

Due to the severe loss incurred by the scattering, high-frequency communication heavily relies on the line-of-sight (LoS) path [23], so we set $L_r = L_G = 1$ in this paper. This means that the channel is determined by the angle at BS and the angle at RIS.

B. Traditional Beam Training Frameworks

To determine the angle at BS and the angle at RIS, beam training is usually applied. By generating directional beams to search all angles in space, the above angles can be obtained according to the beam tuple with the maximum received power. Here, we introduce two types of traditional beam training frameworks in RIS assisted communication systems: exhaustive beam training and hierarchical beam training.

1) Exhaustive Beam Training: One intuitive way to estimate the angles is to exhaustively search all possible angles in space. As illustrated in Fig. 1(a), both BS and RIS apply codewords in exhaustive codebooks to generate narrow beams and sequentially search all possible angles in space. The codewords at the BS side and the RIS side are denoted as $\mathbf{w}_E(i), i = 1, 2, \cdots, N_t$ and $\mathbf{v}_E(j), j = 1, 2, \cdots, N_r$, respectively. After receiving and recording received powers from all beam tuples, the angle at BS and the angle at RIS are estimated according to the beam tuple with the maximum received power. Since narrow beams are applied in the exhaustive beam training framework, the codebook size is equal to the number of antenna elements at both sides. In our considered scenario, where BS is equipped with N_t antenna elements and RIS is equipped with N_r antenna elements, the necessary beam training overhead is $N_t N_r$. In future communication systems, the antenna number at both BS and RIS tend to be very large, which means the exhaustive beam training framework will suffer from an unacceptable beam training overhead.

2) Hierarchical Beam Training: In order to reduce the beam training overhead, we can apply the idea of hierarchical beam training framework in RIS assisted communication systems [17]. As illustrated in Fig. 1(b), both BS and RIS apply binary search based codebooks, so each layer contains two codewords. We denote the i-th codeword in the j-th layer

at the BS side and the RIS side as $\mathbf{w}_H(j, i)$ and $\mathbf{v}_H(j, i)$, respectively. According to the property of binary search, the numbers codebook layers at the BS side and the RIS side are $L_t = \log_2(N_t)$ and $L_r = \log_2(N_r)$, respectively. To gradually narrow down the possible range of UE, the beam patterns of codewords in higher layers possess higher angular resolutions compared to codewords in lower layers.

During the beam training procedure, the codewords are transmitted layer by layer. Specifically, at the first layer, the BS and RIS sequentially transmits four beam tuples to UE, which can be listed as $\{\mathbf{w}_H(1,1), \mathbf{v}_H(1,1)\}$, $\{\mathbf{w}_H(1,1), \mathbf{v}_H(1,2)\}$, $\{\mathbf{w}_H(1,2), \mathbf{v}_H(1,1)\}$ and $\{\mathbf{w}_H(1,2), \mathbf{v}_H(1,2)\}$. We set $\mathbf{u} \in \{\{0,0\}, \{0,1\}, \{1,0\}, \{1,1\}\}^{\max\{\log_2(N_t), \log_2(N_r)\}}$ as the tuple vector which describes the angle at BS and the angle at RIS. Then, we set $\mathbf{u}(1) = \{0,0\}$ if the received power of beam tuple $\{\mathbf{w}_H(1,1), \mathbf{v}_H(1,1)\}$ is the maximum, $\mathbf{u}(1) = \{0,1\}$ if the received power of beam tuple $\{\mathbf{w}_H(1,1), \mathbf{v}_H(1,2)\}$ is the maximum, $\mathbf{u}(1) = \{1,0\}$ if the received power of beam tuple $\{\mathbf{w}_H(1,2), \mathbf{v}_H(1,1)\}$ is the maximum and $\mathbf{u}(1) = \{1,1\}$ if the received power of beam tuple $\{\mathbf{w}_H(1,2), \mathbf{v}_H(1,2)\}$ is the maximum. After transmitting the beam tuples in all layers, the BS can decide the angle at BS and the angle at RIS based on the tuple vector $\mathbf{u}$. We take the first bit of each element in $\mathbf{u}$ as $\mathbf{u}_t$ and the second bit of each element in $\mathbf{u}$ as $\mathbf{u}_r$. The indices of the angle at BS and the angle at RIS can then be derived as $\text{bin2dec}(\mathbf{u}_t)$ and $\text{bin2dec}(\mathbf{u}_r)$, where $\text{bin2dec}(\cdot)$ denotes the operation of transforming a binary number to a decimal number.

By searching the entire space layer by layer, we can exclude many incorrect angles without searching them in an exhaustive way and thus greatly improve the beam training efficiency. The overall beam training overhead is $4\max\{L_t, L_r\} = 4\max\{\log_2(N_t), \log_2(N_r)\}$, which is far less than $N_t N_r$. However, during the binary hierarchical beam training, we need to generate wider beams than those in exhaustive beam training frameworks, so the beam gains are much smaller than those of narrow beams. In addition, in RIS assisted communication systems, there exists multiplicative fading effect [18], which means that the equivalent path loss of the BS-RIS-UE link is the product of the path loss of BS-RIS link and the path loss of RIS-UE link. These two factors lead to a SNR at UE during beam training, and thus severely limit the beam training accuracy in RIS assisted communication systems.

III. CODED BEAM TRAINING FRAMEWORK IN RIS SYSTEMS

To enhance the ability to realize accurate beam training under poor SNR conditions, we design a coded beam training framework for RIS assisted communication systems, which is inspired by the coded beam training framework for MIMO in [19]. By applying the idea of channel coding and adding redundant beam training pilots, the accidental error caused by random noise during the beam training can be corrected without feedback.

Different from the scenario considered in [19], in RIS assisted communication systems, we need to estimate the angles both at BS and at RIS, so the best beam tuple, rather

than the best beam, needs to be determined. Specifically, for BS, there are N_t candidate angles, and we need $k_t = \log_2(N_t)$ information bits to determine the angle at BS. Similarly, RIS has N_r candidate angles, and we need $k_r = \log_2(N_r)$ information bits to determine the angle at RIS. Similar to the hierarchical beam training framework in Section II-B, the information bits at BS and RIS are $\mathbf{u}_t$ and $\mathbf{u}_r$, respectively. To leverage the error correction capability of channel coding, we need to encode the effective information bits by mapping the information bits $\mathbf{u}_t \in \{0,1\}^{k_t}$ and $\mathbf{u}_r \in \{0,1\}^{k_r}$ to codewords $\mathbf{x}_t \in \{0,1\}^{n_t}$ and $\mathbf{x}_r \in \{0,1\}^{n_r}$, where $k_t \leq n_t, k_r \leq n_r$. We denote the encoding function at BS and RIS as f_t and f_r. Then, we have $\mathbf{x}_t = f_t(\mathbf{u}_t)$ and $\mathbf{x}_r = f_r(\mathbf{u}_r)$.

After encoding the information bits, we need to build the connection between the codewords $\mathbf{x}_t$ and $\mathbf{x}_r$ and the beam pattern in space during beam training. We denote the candidate angle list at BS as $\boldsymbol{\Omega}_t \in \mathbb{R}^{N_t}$, which can be expressed as

$$\boldsymbol{\Omega}_t(n) = \sin^{-1}\left(-\frac{N_t+1}{N_t} + \frac{2n}{N_t} \right), n = 1, 2, \cdots, N_t, \quad (7)$$

and the candidate angle list at RIS as $\boldsymbol{\Omega}_r \in \mathbb{R}^{N_t \times 2}$, which can be expressed as

$$\boldsymbol{\Omega}_r(n,1) = \sin^{-1}\left[\left(-\frac{N_{r_1}+1}{N_{r_1}} + \frac{2\lceil \frac{n}{N_{r_1}} \rceil}{N_{r_1}} \right) / \sin(\boldsymbol{\Omega}_r(n,2)) \right]$$

$$\boldsymbol{\Omega}_r(n,2) = \cos^{-1}\left(\frac{1-N_{r_2}}{N_{r_2}} + \frac{2\mathrm{mod}\,(n,N_{r_2})}{N_{r_2}} \right)$$

$$n = 1, 2, \cdots, N_r, \quad (8)$$

where $\boldsymbol{\Omega}_r(:,1)$ denotes the azimuth angles and $\boldsymbol{\Omega}_r(:,2)$ denotes the elevation angles. We denote the beam pattern at BS and at RIS as $\mathcal{V}_t \in \{0,1\}^{n_t \times N_t}$ and $\mathcal{V}_r \in \{0,1\}^{n_r \times N_r}$, which can be obtained by

$$\begin{aligned} \mathcal{V}_t(:,i) &= \mathbf{x}_t^{(i)}, i = 1, 2, \cdots, N_t \\ \mathcal{V}_r(:,j) &= \mathbf{x}_r^{(j)}, j = 1, 2, \cdots, N_r \end{aligned}, \quad (9)$$

where $\mathbf{x}_t^{(i)} = f_t(\mathbf{u}_t^{(i)})$ and $\mathbf{x}_r^{(j)} = f_r(\mathbf{u}_r^{(j)})$. Here, $\mathbf{u}_t^{(i)}$ and $\mathbf{u}_r^{(j)}$ denotes the information bits of different angle indices, which can be expressed as $\mathbf{u}_t^{(i)} = \mathrm{dec2bin}(i, N_t)$ and $\mathbf{u}_r^{(j)} = \mathrm{dec2bin}(j, N_r)$, where $\mathrm{dec2bin}(\cdot, \kappa)$ denotes the operation of transforming a decimal number to a binary number of κ bits.

Based on the beam pattern, the beam training codebook can be designed. We denote the beam training codebook at BS and at RIS as $\mathcal{C}_t \in \mathbb{C}^{n_t \times N_t \times 2}$ and $\mathcal{C}_r \in \mathbb{C}^{n_r \times N_r \times 2}$. Each layer has two codewords, and the beams generated by these two codewords should cover the entire space without overlapping. Specifically, for $\mathcal{C}_t$, at the i-th layer, the first codeword $\mathcal{C}_t(i,:,1)$ should cover the angles $\boldsymbol{\Omega}_t(\varrho)$, where $\mathcal{V}_t(i, \varrho) = 1$. On the contrary, the second codeword $\mathcal{C}_t(i,:,2)$ should cover the angles $\boldsymbol{\Omega}_t(\varrho)$, where $\mathcal{V}_t(i, \varrho) = 0$. For $\mathcal{C}_r$, at the j-th layer, the first codeword $\mathcal{C}_r(j,:,1)$ should cover the angles $\boldsymbol{\Omega}_r(\varepsilon,:)$, where $\mathcal{V}_r(j, \varepsilon) = 1$, while the second codeword $\mathcal{C}_r(j,:,2)$ should cover the angles $\boldsymbol{\Omega}_r(\varepsilon)$, where $\mathcal{V}_r(j, \varepsilon) = 0$.

After designing the beam training codebook, we can start the training process. Similar to the hierarchical beam training framework, at each layer (i.e., the i-th layer), BS and

RIS sequentially transmits four beam tuples to UE, which can be listed as $\{\mathcal{C}_t(i,:,1), \mathcal{C}_r(i,:,1)\}$, $\{\mathcal{C}_t(i,:,1), \mathcal{C}_r(i,:,2)\}$, $\{\mathcal{C}_t(i,:,2), \mathcal{C}_r(i,:,1)\}$ and $\{\mathcal{C}_t(i,:,2), \mathcal{C}_r(i,:,2)\}$. Based on the received power of these four beam tuples, we set the received tuple vector $\hat{\mathbf{x}}(i) = \{0,0\}$ if the received power of the first beam tuple is the maximum, $\hat{\mathbf{x}}(i) = \{0,1\}$ if the received power of the second beam tuple is the maximum, $\hat{\mathbf{x}}(i) = \{1,0\}$ if the received power of the third beam tuple is the maximum and $\hat{\mathbf{x}}(i) = \{1,1\}$ if the received power of the fourth beam tuple is the maximum. After transmitting all the beam training pilots, the received tuple vector is denoted as $\hat{\mathbf{x}}$. By seperating each tuple in $\hat{\mathbf{x}}$ into two bits, the received codewords corresponding to BS and RIS can be denoted as $\hat{\mathbf{x}}_t$ and $\hat{\mathbf{x}}_r$. Finally, the decoding function g_t and g_r can be applied to recover the original information bits by $\hat{\mathbf{u}}_t = g_t(\hat{\mathbf{x}}_t)$ and $\hat{\mathbf{u}}_r = g_r(\hat{\mathbf{x}}_r)$. Since we introduce redundant bits through encoding, the error in $\hat{\mathbf{x}}$ can be corrected, thus improving the beam training accuracy under poor SNR conditions.

IV. CODEWORD DESIGN FOR THE PROPOSED CODED BEAM TRAINING FRAMEWORK

The above framework endows the RIS system with the self-correction ability during beam training and will potentially improve the beam training accuracy. To fully unleash the performance of the proposed framework, the accurate beam shape design is essential.

For BS, the codeword design is straightforward [24] . We can generate the multi-mainlobe codeword with a weighted summation of several array response vectors as

$$\mathcal{C} = \sum_{\phi_i \in \tilde{\boldsymbol{\Omega}}_t} e^{j\psi_i} \mathbf{b}(\phi_i), \quad (10)$$

where $\tilde{\boldsymbol{\Omega}}_t$ denotes the set of angles that the codeword needs to cover, and the auxiliary phase off-set ψ_i can help guarantee a high gain within the intended angle range, which can be elaborated as $\psi_i = i\pi(-1 + \frac{1}{N_t}), i = 1, 2, \cdots, \left| \tilde{\boldsymbol{\Omega}}_t \right|$.

However, for RIS, to design a codeword that evenly covers a set of randomly distributed angles in space is not easy due to the constant modulus constraint of RIS elements. Therefore, in this section, we propose a relaxed Gerchberg-Saxton-based codeword design scheme and a dimension reduced encoder design scheme to approach the desired beam pattern at RIS from two aspects.

A. Proposed Relaxed Gerchberg-Saxton-based Codeword Design Scheme

As for the beam training problem, the amplitude of the designed codeword at different angles greatly affects the beam training accuracy, while the specific phase is not that important. This characteristic makes the codeword design problem similar to the phase retrieval problem in the field of digital holography imaging. Gerchberg-Saxton (GS) algorithm is widely used to solve phase retrieval problem [25], [26]. By iteratively imposing the two amplitude measurements in the object plane and diffraction pattern plane, the phase information of the image can be recovered. Following the idea
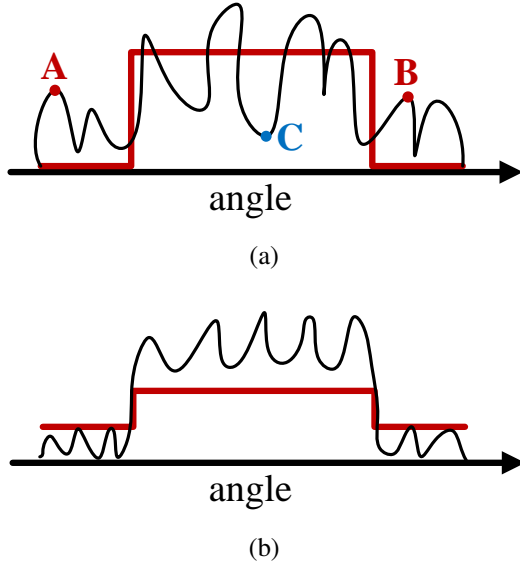
(a)

(b)

Fig. 2. The designed beam shape when applying (a) GS-based codeword design scheme in [20]; (b) proposed relaxed GS-based codeword design scheme.

of GS algorithm, authors in [20] studied a GS-based codeword design scheme at BS by applying the power normalization to one updating process to satisfy the power constraint of the codeword.

However, due to the constant modulus constraint at RIS, directly applying the codeword design scheme in [20] cannot yield an ideal beam shape and would possibly lead to some intrinsic errors even in high SNR scenarios. Specifically, as illustrated in Fig. 2(a), the red line represents the intended beam shape, the black line represents the generated beam shape, and the axis represents angles in space. We can see that due to the constant modulus constraint, the generated beam shape is not ideal and there exists oscillation. At angles out of the angle coverage range, we want the amplitude to be 0 (i.e., point A and point B), but the amplitude is actually larger than some angles within the angle coverage range (i.e., point C). As a consequence, when UE is located in point A or point B, the received power of this codeword is actually high even though it should be near 0. In this case, error could happen even there is no noise in the system.

The root cause of this kind of error is that, the objective of the GS-based codeword design scheme in [20] is to minimize the difference between the intended beam shape and the generated beam shape, which can be formulated as

$$
\begin{aligned}
\min \ & \left\| \mathbf{A}^H \mathbf{v} - \mathbf{s} \right\|_2^2 \\
\text{s.t.} \ & \mathbf{v}(i) = e^{j\vartheta_i}, i = 1, 2, \cdots, N_r,
\end{aligned}
\tag{11}
$$

where $\mathbf{A} \in \mathbb{C}^{N_r \times N_r}$ denotes the array response vectors at different angles, $\mathbf{v} \in \mathbb{C}^{N_r \times 1}$ denotes the designed codeword, and $\mathbf{s} \in \mathbb{C}^{N_r \times 1}$ denotes the intended beam shape. Such an objective is reasonable in the realm of codeword design, but is not suitable for beam training. For beam training, the top priority is to distinguish between the angles within the angle coverage range and the angles out of the angle coverage

range. To solve this problem. we propose a relaxed GS-based codeword design scheme. By relaxing the requirements of approaching the ideal beam pattern, we may not get the most similar beam shape, but we can distinguish the angles within the angle coverage range and the angles out of the angle coverage range more clearly. The procedure of the proposed relaxed GS-based codeword design scheme is elaborated in **Algorithm 1**.

Algorithm 1 Proposed relaxed GS-based codeword design scheme

Input: $\tilde{\mathbf{\Omega}}_r$, $\mathbf{A}$, K_{iter}, Δ
Output: Designed codeword $\mathbf{v}$
 1: Initialize the intended beam shape $\mathbf{s}$ by (14)
 2: Obtain the initial designed codeword $\hat{\mathbf{v}}_{(1)}$ by $\hat{\mathbf{v}}_{(1)} = \frac{1}{\sqrt{N_r}} e^{j\angle\left(\mathbf{A}^\dagger \mathbf{s}_{(0)}\right)}$
 3: **for** $k = 1$ to K_{iter} **do**
 4:     Update $\mathbf{s}_{(k)}$ by (15)
 5:     Update Υ by (16)
 6:     Update $\hat{\mathbf{s}}_{(k)}$ by (17)
 7:     Update $\hat{\mathbf{v}}_{(k+1)}$ by (18)
 8: **end for**
 9: Obtain the designed codeword $\mathbf{v}$ by $\mathbf{v} = \hat{\mathbf{v}}_{(K_{\text{iter}}+1)}$

Here, $\tilde{\mathbf{\Omega}}_r$ denotes the angle coverage range of the intended beam shape, which is obtained based on $\mathcal{V}_r$. Matrix $\mathbf{A}$ transforms the codeword to the beam shape at the entire space, which can be expressed as

$$
\mathbf{A}(:,n) = \mathbf{a}\left(\mathbf{\Omega}_r(n,1), \mathbf{\Omega}_r(n,2)\right).
\tag{12}
$$

We denote the intended beam shape as

$$
\mathbf{s} = \left[s(\phi_1, \theta_1), \cdots, s(\phi_1, \theta_{N_{r_2}}), \cdots, s(\phi_{N_{r_1}}, \theta_{N_{r_2}})\right], \tag{13}
$$

where $s(\phi, \theta) = |s(\phi, \theta)| e^{j\varphi(\phi, \theta)}$ with $|s(\phi, \theta)|$ being the intended amplitude and $\varphi(\phi, \theta)$ being the phase information. We hope angles within the angle coverage range can receive signals and angles out of the angle coverage range cannot receive signals, so the intended amplitude $|s(\phi, \theta)|$ should be

$$
|s(\phi, \theta)| = \begin{cases} \mathcal{P} & (\phi, \theta) \in \tilde{\mathbf{\Omega}}_r \\ 0 & (\phi, \theta) \notin \tilde{\mathbf{\Omega}}_r \end{cases}, \tag{14}
$$

where $\mathcal{P}$ is the constant decided by the codeword power. At the beginning of the iteration, we initialize the intended beam shape $\mathbf{s}_{(0)}$ by (13), where the amplitude is generate by (14) and the phase information $\varphi(\phi, \theta)$ is generated randomly. Based on $\mathbf{s}_{(0)}$, we initialize the designed codeword $\hat{\mathbf{v}}_{(1)}$ as $\hat{\mathbf{v}}_{(1)} = \frac{1}{\sqrt{N_r}} e^{j\angle\left(\mathbf{A}^\dagger \mathbf{s}_{(0)}\right)}$, where $\angle(\cdot)$ denotes the phase operator.

In the k-th round of iteration, we first calculate the beam shape realized by the designed codeword $\hat{\mathbf{v}}_{(k)}$ as

$$
\mathbf{s}_{(k)} = \mathbf{A}^H \hat{\mathbf{v}}_{(k)}. \tag{15}
$$

Then, different from the scheme in [20], where the intended amplitude $|s(\phi, \theta)|$ is directly assigned to $\mathbf{s}_{(k)}$, we divide the points in $\mathbf{s}_{(k)}$ into two categories. For the first category, the corresponding amplitude can distinguish the angles within the angle coverage range from the angles out of the angle
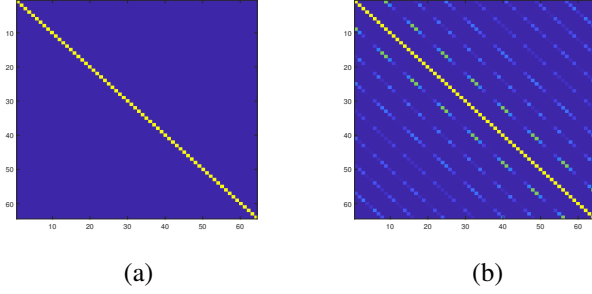
(a)   (b)

Fig. 3. The orthogonality of $\mathbf{A}$ when the size of RIS is (a) 64×1; (b) 8×8.

coverage range. The set of the points in the first category can be expressed as

$$\Upsilon = \Big\{ (\phi, \theta) \mid \big((\phi, \theta) \in \tilde{\boldsymbol{\Omega}}_r \ \&\& \ \mathbf{s}_{(k)}(\phi, \theta) \geq \mathcal{P}(1 - \Delta) \big)$$
$$\| \big((\phi, \theta) \notin \tilde{\boldsymbol{\Omega}}_r \ \&\& \ \mathbf{s}_{(k)}(\phi, \theta) \leq \mathcal{P}\Delta \big) \Big\}, \tag{16}$$

where $\Delta \in [0, 0.5]$ is the dividing factor. For points in the first category, the amplitudes in this round have already revealed the difference between angles within the angle coverage range and angles out of the angle coverage range, so we relax the requirements on them by not assigning the exact amplitude in $\mathbf{s}$ to them. On the contrary, for the second category, where $(\phi, \theta) \notin \Upsilon$, we assign new amplitude to them, so $\hat{\mathbf{s}}_{(k)}$ can be expressed as

$$\hat{\mathbf{s}}_{(k)}(\phi, \theta) =$$
$$\begin{cases} \mathbf{s}_{(k)}(\phi, \theta) & (\phi, \theta) \in \Upsilon \\ \mathcal{P}(1 - \Delta) e^{j\angle(\mathbf{s}_{(k)}(\phi, \theta))} & (\phi, \theta) \notin \Upsilon \ \&\& \ (\phi, \theta) \in \tilde{\boldsymbol{\Omega}}_r \\ \mathcal{P}\Delta e^{j\angle(\mathbf{s}_{(k)}(\phi, \theta))} & (\phi, \theta) \notin \Upsilon \ \&\& \ (\phi, \theta) \notin \tilde{\boldsymbol{\Omega}}_r \end{cases} \tag{17}$$

Based on $\hat{\mathbf{s}}_{(k)}$, the designed codeword for the next round of iteration can be obtained by

$$\hat{\mathbf{v}}_{(k+1)} = \frac{1}{\sqrt{N_r}} e^{j\angle(\mathbf{A}^\dagger \hat{\mathbf{s}}_{(k)})}. \tag{18}$$

After K_{iter} rounds of iteration, the designed codeword $\mathbf{v}$ can finally be obtained as $\mathbf{v} = \hat{\mathbf{v}}_{(K_{\text{iter}}+1)}$.

With the proposed relaxed GS-based codeword design scheme, we can generate beam shape like Fig. 2(b), where the angles within the angle coverage range and the angles out of the angle coverage range can be clearly distinguished.

B. Proposed Dimension Reduced Encoder Design Scheme

The above scheme works well for RIS with ULA, but for a RIS with UPA, the generated beam shape is still highly non-ideal. This is because the orthogonality of $\mathbf{A}$ in the two dimensional case is bad. As illustrated in Fig. 3, for the case where RIS is equipped with a 64×1 ULA, the orthogonality of $\mathbf{A}$ is good, while for case were RIS is equipped with a 8×8 UPA, the orthogonality of $\mathbf{A}$ is bad. Given the fact that RIS usually possesses a UPA structure to guarantee enough reflection area, how to generate good beam shape for a RIS with UPA is crucial to ensure a high beam training accuracy. Since the above scheme works well for RIS with ULA, can
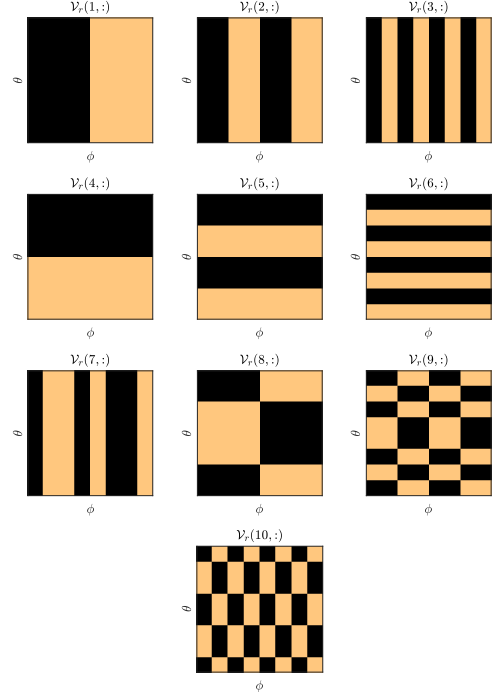

Fig. 4. Beam pattern generated by a random $\mathbf{G}_{\text{Ham}}$.

we decouple the 2D beam shape design problem into two 1D beam shape design problems? To enable this decoupling, the intended beam shape in space should be independent in the two dimensions. As discussed in Section III, the intended beam shape is determined by the encoding function $f_r(\cdot)$. Therefore, we propose a dimension reduced encoder design scheme to decouple the 2D beam shape design problem into two 1D beam shape design problem so as to improve the beam training accuracy.

We choose the Hamming code as the coding scheme for our proposed coded beam training framework because of the high degree of freedom of Hamming code regarding the encoder design. Specifically, a (n_r, k_r) Hamming code can encode a k_r-bit bitstream $\mathbf{u}_r^{(i)}$ into a n_r-bit codeword $\mathbf{x}_r^{(i)}$ with a generator matrix $\mathbf{G}_{\text{Ham}} \in \{0, 1\}^{k_r \times n_r}$ by $\mathbf{x}_r^{(i)} = f_r(\mathbf{u}_r^{(i)}) = \mathbf{u}_r^{(i)} \mathbf{G}_{\text{Ham}}$. The generator matrix $\mathbf{G}_{\text{Ham}}$ has the structure

$$\mathbf{G}_{\text{Ham}} = \begin{bmatrix} \mathbf{I}_{k_r} & \mathbf{Q} \end{bmatrix}, \tag{19}$$

where $\mathbf{I}_{k_r}$ denotes the $k_r \times k_r$ identical matrix. Submatrix $\mathbf{Q} \in \{0, 1\}^{k_r \times (n_r - k_r)}$ is designed artificially. To guarantee the error correction ability, each row of $\mathbf{Q}$ should contain at least two "1". We consider a 8×8 RIS, the necessary number of information bits $k_r = \log_2(8 \times 8) = 6$. The codeword length n_r should thus be at least $n_r = 10$ [27]. We first randomly generate $\mathbf{Q}$ as

$$\mathbf{Q} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}^T . \tag{20}$$

Then, the n_r beam patterns $\mathcal{V}_r$ are depicted in Fig. 4. The

first 6 beam patterns are the same as those in hierarchical beam training frameworks. For these beam patterns, the two dimension (i.e., ϕ and θ) can be decoupled. For example, to generate beam pattern $\mathcal{V}_r(1,:)$, we can design a 1D beam $\mathbf{v}_\phi \in \mathbb{C}^{8 \times 1}$ that covers $\phi \in [0, \pi]$ and a 1D beam $\mathbf{v}_\theta \in \mathbb{C}^{8 \times 1}$ that covers $\theta \in [-\pi, \pi]$. The 2D beam can be realized by codeword $\mathbf{v} = \mathbf{v}_\phi \otimes \mathbf{v}_\theta$. However, for the last 4 beam patterns (the redundant beam patterns for error correction), only the 7th beam pattern can be decoupled into two 1D beams, and the other 3 beam patterns cannot be decoupled since the ϕ-axis and the θ-axis are interwoven with each other.

What leads to this interweave? Since the first 6 columns of $\mathbf{G}_{\text{Ham}}$ is an identical matrix, we can actually view the first 6 beam patterns as the **basis** patterns. Therefore, for the 7th beam pattern, according to the first column of $\mathbf{Q}$, it is obtained by adding up the first three basis (i.e., $[\mathcal{V}_r(1,:) + \mathcal{V}_r(2,:) + \mathcal{V}_r(3,:)]_2$), where $[\cdot]_2$ denotes the mod-2 arithmetic. Since the first three beam patterns are all consistent at θ-axis and varying at ϕ-axis, so the beam pattern $\mathcal{V}_r(7,:)$ can still be decoupled. However, for the 8th beam pattern, it is obtained by $[\mathcal{V}_r(1,:) + \mathcal{V}_r(4,:) + \mathcal{V}_r(5,:)]_2$. Since $\mathcal{V}_r(4,:)$ and $\mathcal{V}_r(5,:)$ are consistent at ϕ-axis and varying at θ-axis, they will interweave with $\mathcal{V}_r(1,:)$ and make $\mathcal{V}_r(8,:)$ unable to be decoupled. Similarly, $\mathcal{V}_r(9,:)$ and $\mathcal{V}_r(10,:)$ are also unable to be decoupled.

The above analysis inspires us that if we need to decouple the redundant beam patterns, we need to design the matrix $\mathbf{Q}$ so that only the basis with the same consistency is added together. For the simplicity of description, beam patterns that are consistent at θ-axis and varying at ϕ-axis are defined as **Type I** pattern, while beam patterns that are consistent at ϕ-axis and varying at θ-axis are defined as **Type II** pattern. To guarantee the error correction ability, we have the following **Lemma 1**.

Lemma 1. *To guarantee the error correction ability at RIS, the number of RIS elements at each dimension should be strictly larger than* 4 *and the number of redundant beam patterns for each dimension should be at least* 3.

Proof. For a block code such as Hamming code, the error correction ability is related to its minimum hamming distance $d_{\min}$, which is equal to the minimum Hamming weight of its nonzero codewords [27]. In order to correct t bits errors, $d_{\min}$ should satisfy $d_{\min} \geq 2t + 1$. In our framework, we hope the Hamming code can correct 1 bit error, so $d_{\min} \geq 3$. As a result, for each row in $\mathbf{G}_{\text{Ham}}$, we should have at least three "1", which means each row of $\mathbf{Q}$ should have at least two "1". As discussed in Section IV-B, the two types of beam patterns cannot co-exist in the same column of $\mathbf{Q}$, so $\mathbf{Q}$ can be written as a block matrix as

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_{\text{I}} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_{\text{II}} \end{bmatrix}, \qquad (21)$$

where $\mathbf{Q}_{\text{I}}$ and $\mathbf{Q}_{\text{II}}$ denote the submatrix related to **Type I** pattern and **Type II** pattern, respectively. Therefore, $\mathbf{Q}_{\text{I}}$ and $\mathbf{Q}_{\text{II}}$ should both have at least two "1". Since **Type I** pattern and **Type II** pattern are homogeneous, we will only discuss **Type I** pattern and $\mathbf{Q}_{\text{I}}$ in the following discussion.

If $N_{r_1} \leq 4$, $\mathbf{Q}_{\text{I}}$ has at most two rows. To avoid repeated beam pattern, $\mathbf{Q}_{\text{I}}$ would only be $\begin{bmatrix} 1 & 1 \end{bmatrix}^T$. In this case, the Hamming weights of first two rows of $\mathbf{G}_{\text{Ham}}$ are only 2, which means $d_{\min} = 2$, and the Hamming code can no longer correct 1 bit error. As a result, the number of RIS elements at each dimension should be strictly larger than 4.

Next, we need to prove that the number of columns in $\mathbf{Q}_{\text{I}}$ should be at least 3. If $\mathbf{Q}_{\text{I}}$ only has two columns, in order for $d_{\min} \geq 3$, all columns in $\mathbf{Q}_{\text{I}}$ should be "1". In this case, if we calculate the difference of these two rows in $\mathbf{G}_{\text{Ham}}$, we can get a codeword with Hamming weight $d_{\min} = 2$. As a result, the number of columns in $\mathbf{Q}_{\text{I}}$ should be at least 3. Similarly, the number of columns in $\mathbf{Q}_{\text{II}}$ should also be at least 3, which completes the proof. $\square$

Based on the above analyses, we now introduce the steps of the proposed dimension reduced encoder design scheme. Since RIS is equipped with $N_r = N_{r_1} \times N_{r_2}$ antenna elements, we have $\log_2(N_{r_1})$ **Type I** patterns and $\log_2(N_{r_2})$ **Type II** patterns. We denote the number of redundant beam patterns for **Type I** patterns and **Type II** patterns as m_{r_1} and m_{r_2} respectively, then they should satisfy

$$\begin{cases} m_{r_1} = \max\{3, m_{r_1,\text{int}}\} \\ m_{r_2} = \max\{3, m_{r_2,\text{int}}\} \end{cases}, \qquad (22)$$

where $m_{r_1,\text{int}}$ denotes the minimum integer that satisfies $2^{m_{r_1,\text{int}}} - m_{r_1,\text{int}} - 1 \geq \log_2(N_{r_1})$ and $m_{r_2,\text{int}}$ denotes the minimum integer that satisfies $2^{m_{r_2,\text{int}}} - m_{r_2,\text{int}} - 1 \geq \log_2(N_{r_2})$ [27]. For redundant beam patterns corresponding to **Type I** patterns, each row of $\mathbf{Q}_{\text{I}} \in \{0,1\}^{\log_2(N_{r_1}) \times m_{r_1}}$ should be composed of m_{r_1}-tuples of weight 2 or more. There are a total of $\sum_{i=2}^{m_{r_1}} C(m_{r_1}, i) = 2^{m_{r_1}} - m_{r_1} - 1$ types of m_{r_1}-tuples of weight 2 or more, so we can always fill $\mathbf{Q}_{\text{I}}$ without repeating existing tuples. Meanwhile, $\mathbf{Q}_{\text{II}} \in \{0,1\}^{\log_2(N_{r_2}) \times m_{r_2}}$ can be generated by the same way. Finally, the submatrix $\mathbf{Q}$ can be composed by

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_{\text{I}} & \mathbf{0}_{\log_2(N_{r_1}) \times m_{r_2}} \\ \mathbf{0}_{\log_2(N_{r_2}) \times m_{r_1}} & \mathbf{Q}_{\text{II}} \end{bmatrix}, \qquad (23)$$

where $\mathbf{0}_{\iota \times \gamma}$ denotes the all-zero matrix with dimension $\iota \times \gamma$.

With the proposed scheme, now we get back to the above example where the RIS is equipped with 8×8 elements. In this case, $m_{r_1} = m_{r_2} = 3$, and submatrix $\mathbf{Q}$ can be generated as

$$\mathbf{Q} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}. \qquad (24)$$

Then, the $n_r = k_r + m_{r_1} + m_{r_2}$ beam patterns $\mathcal{V}_r$ are depicted in Fig. 5. Through the proposed dimension reduced encoder design scheme, although we need two more redundant beam patterns, the two dimensions of RIS are properly decoupled and the quality of generated beam shape can be guarenteed.

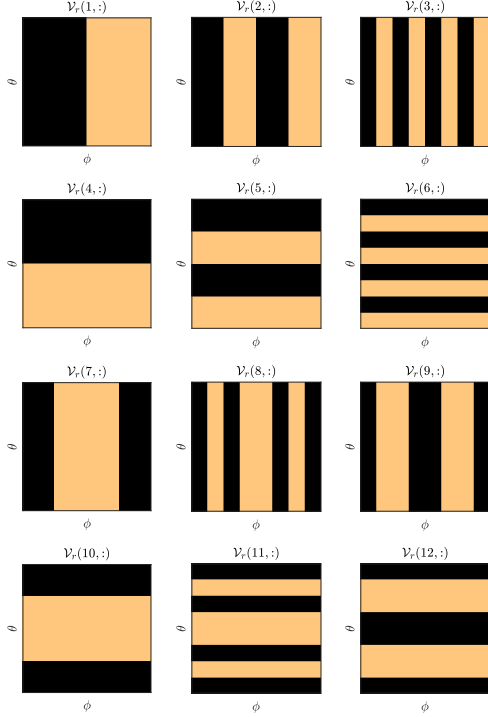Based on the designed $\mathbf{Q}$, the check matrix $\mathbf{H}$ can be

Fig. 5. Beam pattern generated by proposed encoder.

expressed as

$$\mathbf{H} = \begin{bmatrix} \mathbf{Q}^T & \mathbf{I}_{m_{r_1}+m_{r_2}} \end{bmatrix}, \quad (25)$$

based on which we can determine whether the received codeword $\hat{\mathbf{x}}_r$ contains error bits by calculating the syndrome $\mathbf{c}_r$ as

$$\mathbf{c}_r = \hat{\mathbf{x}}_r \mathbf{H}^T. \quad (26)$$

When all bits in $\mathbf{c}_r$ equal to zero, the received $\hat{\mathbf{x}}_r$ is a normal codeword and there is no error. On the contrary, when $\mathbf{c}_r \neq \mathbf{0}_{1\times(m_{r_1}+m_{r_2})}$, $\hat{\mathbf{x}}_r$ is not a normal codeword generated by the designed codebook and there exists error in $\hat{\mathbf{x}}_r$. Since $d_{\min} = 3$, all 1-bit error has a unique syndrome and can be corrected. In addition, when there exists 1-bit error in **Type I** pattern and 1-bit error in **Type II** pattern simultaneously, these 2-bit errors can also be corrected according to the following **Lemma 2**.

Lemma 2. *By the proposed dimension reduced encoder design scheme, the error in **Type I** patterns and the error in **Type II** patterns are independent with each other.*

Proof. For a certain information bit stream $\mathbf{u}_r^{(j)}$, it can be divided into two parts: the bits corresponding to the ϕ dimension with length $\log_2(N_{r_1})$, denoted as $\mathbf{u}_{r,\mathrm{I}}^{(i)}$, and the bits corresponding to the θ dimension with length $\log_2(N_{r_2})$, denoted as $\mathbf{u}_{r,\mathrm{II}}^{(i)}$. The generated codeword $\mathbf{x}_r^{(i)}$ can then be derived as

$$\mathbf{x}_r^{(i)} = \begin{bmatrix} \mathbf{u}_{r,\mathrm{I}}^{(i)} & \mathbf{u}_{r,\mathrm{II}}^{(i)} \end{bmatrix} \begin{bmatrix} \mathbf{I}_{\log_2(N_{r_1})} & \mathbf{0} & \mathbf{Q}_{\mathrm{I}} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{\log_2(N_{r_2})} & \mathbf{0} & \mathbf{Q}_{\mathrm{II}} \end{bmatrix}$$
$$= \begin{bmatrix} \mathbf{u}_{r,\mathrm{I}}^{(i)} & \mathbf{u}_{r,\mathrm{II}}^{(i)} & \mathbf{u}_{r,\mathrm{I}}^{(i)}\mathbf{Q}_{\mathrm{I}} & \mathbf{u}_{r,\mathrm{II}}^{(i)}\mathbf{Q}_{\mathrm{II}} \end{bmatrix}. \quad (27)$$

TABLE I
BEAM TRAINING OVERHEADS FOR DIFFERENT FRAMEWORKS

Frameworks	Training Overheads
Exhaustive beam training	$N_t N_r$
Hierarchical beam training	$4\max\{\log_2(N_t), \log_2(N_r)\}$
Coded beam training	$4\max\{n_t, n_r\}$

After the transmission, we denote the received codeword as

$$\hat{\mathbf{x}}_r^{(i)} = \begin{bmatrix} \hat{\mathbf{u}}_{r,\mathrm{I}}^{(i)} & \hat{\mathbf{u}}_{r,\mathrm{II}}^{(i)} & \widehat{\mathbf{u}_{r,\mathrm{I}}^{(i)}\mathbf{Q}_{\mathrm{I}}} & \widehat{\mathbf{u}_{r,\mathrm{II}}^{(i)}\mathbf{Q}_{\mathrm{II}}} \end{bmatrix}. \quad (28)$$

Then, the syndrome $\mathbf{c}_r$ can be derived as

$$\mathbf{c}_r = \hat{\mathbf{x}}_r^{(i)}\mathbf{H}^T = \begin{bmatrix} \mathbf{c}_{r,\mathrm{I}} & \mathbf{c}_{r,\mathrm{II}} \end{bmatrix}$$
$$= \begin{bmatrix} \hat{\mathbf{u}}_{r,\mathrm{I}}^{(i)} & \hat{\mathbf{u}}_{r,\mathrm{II}}^{(i)} & \widehat{\mathbf{u}_{r,\mathrm{I}}^{(i)}\mathbf{Q}_{\mathrm{I}}} & \widehat{\mathbf{u}_{r,\mathrm{II}}^{(i)}\mathbf{Q}_{\mathrm{II}}} \end{bmatrix} \begin{bmatrix} \mathbf{Q}_{\mathrm{I}} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_{\mathrm{II}} \\ \mathbf{I}_{m_{r_1}} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{m_{r_2}} \end{bmatrix} \quad (29)$$
$$= \begin{bmatrix} \hat{\mathbf{u}}_{r,\mathrm{I}}^{(i)}\mathbf{Q}_{\mathrm{I}} + \widehat{\mathbf{u}_{r,\mathrm{I}}^{(i)}}\mathbf{Q}_{\mathrm{I}} & \hat{\mathbf{u}}_{r,\mathrm{II}}^{(i)}\mathbf{Q}_{\mathrm{II}} + \widehat{\mathbf{u}_{r,\mathrm{II}}^{(i)}}\mathbf{Q}_{\mathrm{II}} \end{bmatrix}.$$

We can see from (29) that $\mathbf{c}_r$ has two parts, and for the first part $\mathbf{c}_{r,\mathrm{I}}$, it is only related to the error happened to **Type I** patterns, and for the second part $\mathbf{c}_{r,\mathrm{II}}$, it is only related to the error happened to **Type II** patterns, which completes the proof. □

From the above analyses, the proposed dimension reduced encoder design scheme can not only improve the quality of beam shape by enabling the coupling of two dimensions of RIS, but also enhance the error correction capability of traditional Hamming code, thus further improving the beam training accuracy.

C. Beam Training Overhead Analysis

In this subsection, we will analyze the necessary beam training overheads of the traditional exhaustive beam training framework and traditional hierarchical beam training framework and compare them with that of the proposed coded beam training framework. The results are listed in Table I.

Specifically, For the traditional exhaustive beam training framework, each possible beam tuples in space should be sequentially explored before determining the beam best tuple. Given the fact that the number of candidate narrow beams is equal to the number of antenna elements, the total beam training overhead of the exhaustive beam training framework should be $N_t N_r$. In this case, when the number of RIS elements is large, an unacceptable beam training overhead will severely limit the system performance. On the other hand, for the traditional hierarchical beam training framework, we need $2 \times 2 = 4$ beams at each layer. Since the numbers of layers at BS and RIS are $\log_2(N_t)$ and $\log_2(N_r)$, respectively, the total beam training overhead is $4\max\{\log_2(N_t), \log_2(N_r)\}$. Through the hierarchical beam training framework, a lot of incorrect angles are excluded at low layers, so the number of necessary beam training overhead is greatly reduced.

For the proposed coded beam training framework, the codewords are composed of basis patterns and the redundant

TABLE II
SIMULATION PARAMETERS

BS antenna number N_t	64
RIS antenna number $N_{r_1} \times N_{r_2} = N_r$	$16 \times 16 = 256$
Central frequency f_c	28 GHz
The distribution of ϕ, θ	$\mathcal{U}(-\pi, \pi)$
Iteration number K_{iter}	100
Threshold Δ	0.3
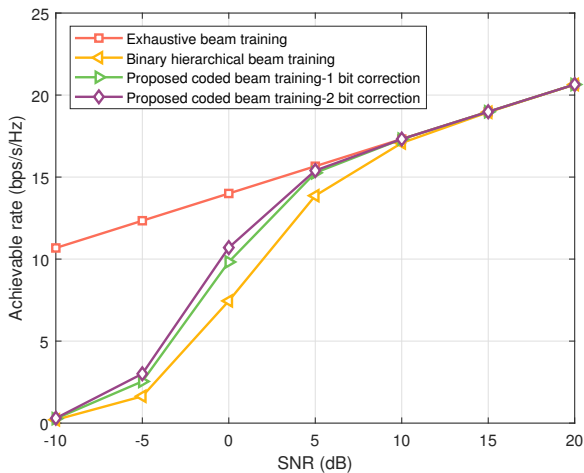
Fig. 6. Achievable rate performance v.s. SNR.

Fig. 7. Success rate v.s. SNR.

patterns. Similarly, we need four beams in each layer. Therefore, the necessary beam training overhead for our proposed framework is $4 \max \{n_t, n_r\}$, where $n_t = \log_2(N_t) + m_t$ and $n_r = \log_2(N_r) + m_r$. Since m is the minimum integer that satisfies $2^m - m - 1 \leq \log_2(N)$, when $N \leq 3$, $m_r \leq \log_2(N)$, which means that the proposed scheme will not introduce a large extra beam training overhead compared to hierarchical beam training framework.

V. SIMULATION RESULTS

In this section, we evaluate the performance of the proposed coded beam training framework through numerical experiments. The simulation parameters are listed in **Table II** The antenna spacing is set to $d = \frac{c}{2f_c}$. We compare the achievable rate performance of the proposed coded beam training framework with both the traditional exhaustive beam training framework and the traditional hierarchical beam training framework. The achievable rate is obtained by

$$R = \log_2 \left(1 + \frac{P_t}{\sigma^2} \mathbf{h}_r \text{diag}(\mathbf{v}) \mathbf{G} \mathbf{w} \mathbf{w}^H \mathbf{G}^H \text{diag}(\mathbf{v}^H) \mathbf{h}_r^H \right), \tag{30}$$

where P_t denotes the transmission power at BS and σ^2 denotes the noise power. The reflecting vector of RIS $\mathbf{v}$ and the beamforming vector of BS $\mathbf{w}$ are both determined through the corresponding beam training frameworks.

Fig. 6 depicts the achievable rate performance of different beam training frameworks against the SNR. We assume that the beam training overheads for all frameworks are all sufficient. In this case, the traditional exhaustive beam training
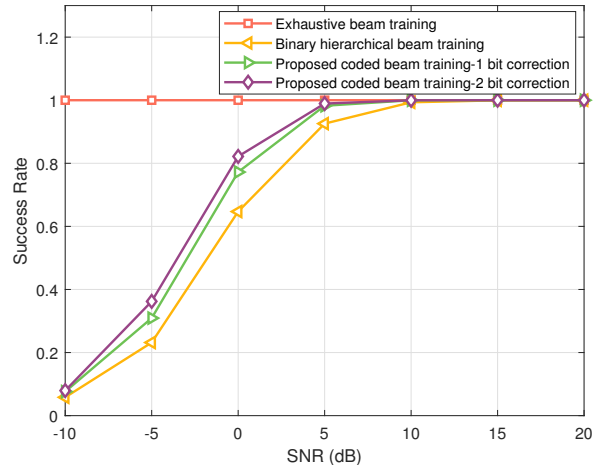
framework can always detect the best beam tuples for BS and RIS thanks to the high beamforming gain realized by the large antenna array. We can observe that compared to the traditional compared to the traditional hierarchical the proposed coded beam training framework can realize a higher achievable rate performance in low SNR scenarios. The suffix "*1 bit correction*" means that we only utilize the check matrix **H** to correct 1-bit error in received codewords as traditional Hamming code, while the suffix "*2 bit correction*" means that we exploit the property of the designed encoder to enable some 2-bit errors to be corrected. We can see that since the proposed dimension reduced encoder can decouple the two dimensions of RIS, the error correction capability can also be enhanced compared to traditional Hamming code.

To evaluate the probability of different frameworks to select the best tuple, we compare the success rate of different frameworks against SNR in Fig. 7. Here, we also assume that the beam training overhead for all frameworks are all sufficient. Similar to the achievable rate performace, the proposed coded beam training framework are more likely to detect the best tuple for BS and RIS successfully compared to traditional hierarchical beam training framework thanks to the error correction capability brought by the encoding-decoding process. In addition, through the decoupling ability enabled by the proposed dimension reduced encoder design scheme, the proposed framework embraces a higher success rate compared to the framework based on traditional Hamming code.

Furthermore, to reveal the impact of beam training overhead on different frameworks, we compare the achievable rate performance of different beam training frameworks against the number of pilot overheads in Fig. 8. The beam training SNR is set to 10 dB and the pilot overhead is increasing from 4 to 100. In our considered system, the necessary beam training overhead for the traditional hierarchical beam training framework should be $4 \max \{\log_2(N_t), \log_2(N_r)\} = 4 \times 8 = 32$. The necessary beam training overhead for the proposed coded beam training framework is $4 \max \{n_t, n_r\} = 4 \times 14 = 56$. We can observe that when the pilot number is insufficient for all
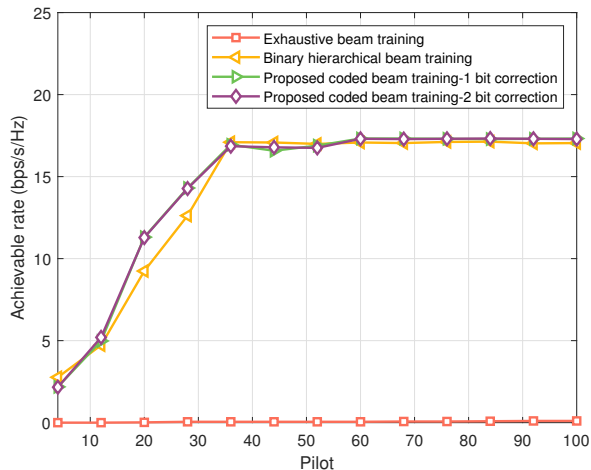
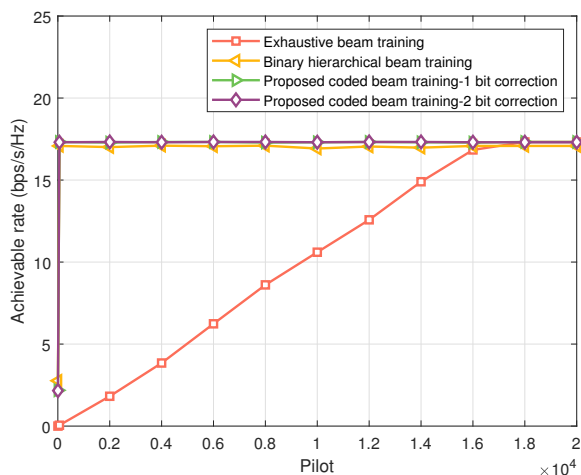Fig. 8. Achievable rate performance v.s. beam training overhead, $P \in [4, 100]$.

Fig. 9. Achievable rate performance v.s. beam training overhead, $P \in [4, 20000]$.

frameworks, the proposed coded beam training framework still outperforms existing hierarchical beam training framework since it also have certain error correction capabilities. When the pilot number is sufficient for the hierarchical framework but insufficient for the proposed framework, the achievable rate of the proposed scheme is slightly lower than that of the hierarchical framework. This is because the redundant beams have not been entirely transmitted, so the error correction can sometimes be misleading. When the pilot number is sufficient for the proposed framework, it can reach the maximum achievable rate thanks to the error correction ability. We also notice that in this scenario, the trend of the "1 bit correction" and the "2 bit correction" is nearly the same, this is because when SNR= 10 dB, both schemes can determine the best beam tuple, which is consistent with the results in Fig. 6 and Fig. 7. The traditional exhaustive beam training framework, however, can barely detect the best beam tuple since the pilot number is severely insufficient.

To demonstrate the performance improvement of the traditional exhaustive beam training framework more clearly, in Fig. 9, the pilot overhead is increasing from 4 to 20000. In our considered scenario, the necessary beam training overhead for the traditional exhaustive beam training framework should be $N_t N_r = 16384$. From Fig. 9, we can see that when the pilot number is below 16000, the achievable rate of the exhaustive framework improves gradually. When the pilot number is around 16000, the achievable rate is nearly the same as the hierarchical framework. When the pilot number is sufficient, the achievable rate is the same as the proposed coded beam training framework. From Fig. 8 and Fig. 9, we can see that due to the error correction capability brought by the encoding-decoding process, the proposed coded beam training framework can outperform existing frameworks under different pilot numbers, which further verified the advantage of the proposed framework.

VI. CONCLUSIONS

In this paper, we exploited the error correction capability of channel coding to realize accurate beam training under low SNR. By mapping the angles in space to a bitstream, we enabled the encoding-decoding procedure during beam training. Then, considering the constant modulus constraints of RIS elements, we adopted a new codeword design criterion and proposed a relaxed GS-based codeword design scheme. Furthermore, we proposed a dimension reduced encoder design scheme to improve the quality of the beam shape and the capability of error correction simultaneously. Simulation results verified the effectiveness of the proposed scheme. The proposed framework revealed the similarity of intrinsic mathematical structures between channel coding and beam training, which enabled the error correction during beam training and provided a promising solution for accurate and reliable beam training in RIS systems. For future works, this coded beam training framework can be extended to more scenarios such as near-field scenarios. In addition, various channel coding methods can be applied to the proposed framework to enable reliable beam training under low SNR.

REFERENCES

[1] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M.-S. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116 753–116 773, Jul. 2019.

[2] M. Di Renzo, A. Zappone, M. Debbah, M.-S. Alouini, C. Yuen, J. De Rosny, and S. Tretyakov, "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, pp. 2450–2525, Jul. 2020.

[3] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, Nov. 2019.

[4] C. Pan, H. Ren, K. Wang, J. F. Kolb, M. Elkashlan, M. Chen, M. Di Renzo, Y. Hao, J. Wang, A. L. Swindlehurst *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.

[5] G. Zhou, C. Pan, H. Ren, K. Wang, and A. Nallanathan, "A framework of robust transmission design for IRS-aided MISO communications with imperfect cascaded channels," *IEEE Tran. Signal Process.*, vol. 68, pp. 5092–5106, Aug. 2020.

[6] K. Zhi, C. Pan, H. Ren, and K. Wang, "Power scaling law analysis and phase shift optimization of RIS-aided massive MIMO systems with statistical CSI," *IEEE Trans. Commun.*, vol. 70, no. 5, pp. 3558–3574, Mar. 2022.

[7] B. Zheng, C. You, W. Mei, and R. Zhang, "A survey on channel estimation and practical passive beamforming design for intelligent reflecting surface aided wireless communications," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 2, pp. 1035–1071, Feb. 2022.

[8] H. Alwazani, A. Kammoun, A. Chaaban, M. Debbah, M.-S. Alouini *et al.*, "Intelligent reflecting surface-assisted multi-user MISO communication: Channel estimation and beamforming design," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 661–680, May 2020.

[9] Z. Wang, L. Liu, and S. Cui, "Channel estimation for intelligent reflecting surface assisted multiuser communications: Framework, algorithms, and analysis," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6607–6620, Jun. 2020.

[10] L. Wei, C. Huang, G. C. Alexandropoulos, C. Yuen, Z. Zhang, and M. Debbah, "Channel estimation for RIS-empowered multi-user MISO wireless communications," *IEEE Trans. Commun.*, vol. 69, no. 6, pp. 4144–4157, Mar. 2021.

[11] J. Suh, C. Kim, W. Sung, J. So, and S. W. Heo, "Construction of a generalized DFT codebook using channel-adaptive parameters," *IEEE Commun. Lett.*, vol. 21, no. 1, pp. 196–199, Jan. 2016.

[12] Y. Chen, J. Tan, M. Hao, R. MacKenzie, and L. Dai, "Accurate beam training for RIS-assisted wideband terahertz communication," *IEEE Trans. Commun.*, vol. 71, no. 12, pp. 7425–7440, Dec. 2023.

[13] 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.

[14] P. Wang, J. Fang, W. Zhang, and H. Li, "Fast beam training and alignment for IRS-assisted millimeter wave/terahertz systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 4, pp. 2710–2724, Apr. 2021.

[15] 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.

[16] Y. Xu, C. Huang, L. Wei, Z. Yang, X. Chen, Z. Zhang, C. Yuen, and M. Debbah, "Low-complexity beam training for multi-RIS-assisted multi-user communications," *IEEE Wireless Commun. Lett.*, 2024.

[17] J. Wang, W. Tang, S. Jin, C.-K. Wen, X. Li, and X. Hou, "Hierarchical codebook-based beam training for RIS-assisted mmWave communication systems," *IEEE Trans. Commun.*, vol. 71, no. 6, pp. 3650–3662, Mar. 2023.

[18] Z. Zhang, L. Dai, X. Chen, C. Liu, F. Yang, R. Schober, and H. V. Poor, "Active RIS vs. passive RIS: Which will prevail in 6G?" *IEEE Trans. Commun.*, vol. 71, no. 3, pp. 1707–1725, Dec. 2022.

[19] T. Zheng, J. Zhu, Q. Yu, Y. Yan, and L. Dai, "Coded beam training," *arXiv preprint arXiv:2401.01673*, 2024.

[20] Y. Lu, Z. Zhang, and L. Dai, "Hierarchical beam training for extremely large-scale MIMO: From far-field to near-field," *IEEE Trans. Commun.*, vol. 72, no. 4, pp. 2247–2259, Apr. 2024.

[21] L. Dai, B. Wang, M. Wang, X. Yang, J. Tan, S. Bi, S. Xu, F. Yang, Z. Chen, M. Di Renzo *et al.*, "Reconfigurable intelligent surface-based wireless communications: Antenna design, prototyping, and experimental results," *IEEE Access*, vol. 8, pp. 45 913–45 923, Mar. 2020.

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

[23] C. Han, L. Yan, and J. Yuan, "Hybrid beamforming for terahertz wireless communications: Challenges, architectures, and open problems," *IEEE Wireless Commun.*, vol. 28, no. 4, pp. 198–204, Aug. 2021.

[24] C. Qi, K. Chen, O. A. Dobre, and G. Y. Li, "Hierarchical codebook-based multiuser beam training for millimeter wave massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 8142–8152, Dec. 2020.

[25] R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of plane from image and diffraction pictures," *Optik*, vol. 35, no. 2, pp. 237–246, Sep. 1972.

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

[27] S. Lin and D. Costello, *Error Control Coding*. Pearson Education, 2004.