

# Channel Estimation for RIS Assisted Wireless Communications: Stationary or Non-Stationary?

Yuhao Chen , Graduate Student Member, IEEE, Mengnan Jian , and Linglong Dai , Fellow, IEEE

**Abstract**—Reconfigurable intelligent surface (RIS) is considered as a promising technology for future 6G communications. In RIS assisted communication systems, precise channel state information (CSI) is the prerequisite of efficient beamforming. Most existing channel estimation schemes rely on the spatial stationarity assumption. However, with the large array aperture of RIS, spatial non-stationarity effect naturally exists, which cannot be captured by most existing channel estimation schemes due to the spatial ambiguity effect of RIS and causes a significant decrease in the estimation accuracy. To tackle this problem, in this paper, we study the non-stationary channel estimation for RIS assisted communication systems. Specifically, we propose a two-dimensional group time block code (GTBC) based RIS decoupling scheme to convert the entire spatial non-stationary channel into a series of spatial stationary sub-channels. The proposed scheme is realized by artificially creating the time-domain relevance and exploiting it to recognize the spatial non-stationarity effect in the space domain. Based on the decoupled signals, a GTBC based dynamic orthogonal matching pursuit (GD-OMP) scheme is then proposed to update the dynamic codebook and estimate each path iteratively. Moreover, the computational complexity of the proposed GD-OMP scheme is analyzed. Finally, simulation results verify that the proposed scheme can capture the spatial non-stationarity effect and reach a higher channel estimation accuracy than existing schemes.

**Index Terms**—Spatial non-stationary, RIS, channel estimation.

## I. INTRODUCTION

**R**ECONFIGURABLE intelligent surface (RIS) is considered as one of the key technologies for future 6G wireless communications [1]. With numerous low-cost reflecting

elements, RIS can control the electromagnetic environment intelligently with low power consumption [2], [3], [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]. RIS should be equipped with thousands of elements in order to reach a high beamforming gain [7], [8]. To realize effective beamforming so as to unleash the potential benefits of RIS, accurate channel state information (CSI) is essential [9], [10].

### A. Prior Works

In RIS assisted communication systems, due to the passive characteristic of RIS elements, only the cascaded channel (i.e., the composite of user-RIS channel and RIS-BS channel) can be estimated. Intuitively, the cascaded channel can be obtained by turning on the RIS elements one-by-one [11], while this scheme suffers from unacceptable pilot overhead considering the large element number of RIS. To address this issue, researchers have developed various low pilot overhead channel estimation schemes by making use of the underlying structure of RIS's cascaded channels. For single-user systems, several compressive sensing (CS) based schemes have been studied [12], [13], [14], [15], [16], [17]. For example, the authors in [12], [13], [14] applied the orthogonal matching pursuit (OMP) based schemes to exploit the spatial sparsity such as the angular-domain sparsity [12], [13] or the angular-distance domain (polar domain) sparsity [14] to estimate the RIS cascaded channel with low pilot overhead. To avoid the grid mismatch in OMP based schemes, [15] studied an off-grid sparse Bayesian learning (SBL) based scheme to carry out direct estimation on the continuous angular domain. Meanwhile, the message passing (MP) schemes were utilized in [16], [17] to effectively estimate cascaded channels by performing inference on graphical models in an iterative manner. Furthermore, for multi-user systems, since all users share the common BS-RIS channel, more structured characteristics could be utilized to improve the channel estimation accuracy [18], [19], [20], [21]. Specifically, in [18], the common column-block sparsity of the cascaded channel matrices of multiple users is exploited to iteratively estimate the multi-user channels. In addition, the authors in [19] revealed the double-structured sparsity of the cascaded channel, which could be utilized to further reduce the pilot overhead. To utilize the correlation between the cascaded channel matrices of different users, the authors in [20], [21] studied a two-stage

Manuscript received 6 February 2024; revised 12 June 2024 and 23 July 2024; accepted 5 August 2024. Date of publication 8 August 2024; date of current version 26 August 2024. This work was supported in part by the National Key Research and Development Program of China under Grant 2023YFB3811503, in part by the National Science Fund for Distinguished Young Scholars under Grant 62325106, and in part by the ZTE Industry-University-Institute Cooperation Funds under Grant HC-CN-20200923014. The associate editor coordinating the review of this article and approving it for publication was Dr. Foad Sahrabi. (Corresponding author: Linglong Dai.)

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

Mengnan Jian is with the State Key Laboratory of Mobile Network and Mobile Multimedia Technology, Shenzhen 518055, China, and also with the Wireless Research Institute, ZTE Corporation, Beijing 100029, China (e-mail: jian.mengnan@zte.com.cn).

Digital Object Identifier 10.1109/TSP.2024.3440326

channel estimation scheme to fully exploit the angle and gain information of the common BS-RIS channel. By conducting the channel estimation for a typical user in the first stage, only the low-dimensional user-RIS channels, rather than the high-dimensional cascaded channels, need to be updated in the second stage. As a result, the required overhead for estimating the cascaded channels of other users could be greatly reduced.

Despite the substantial studies conducted on the RIS channel estimation, the above schemes are all based on the spatial stationarity assumption, i.e., all RIS elements are visible to the same scatterers or users. However, due to the large aperture of RIS, spatial non-stationarity effect naturally exists in both far-field and near-field scenarios, which has been verified in [22], [23]. Specifically, spatial non-stationarity effect means that different part of the RIS may be visible to different scatterers or users, which is usually described by visibility regions (VR) [24]. With the presence of spatial non-stationarity, the basis of the channel matrices is no longer the array response vector as in the scheme mentioned above. As a consequence, traditional channel estimation schemes fail to capture the spatial non-stationarity in RIS systems, inevitably leading to a severe decrease in the channel estimation accuracy. Actually, though the spatial non-stationarity effect has not been considered in RIS systems in most studies, some studies have been conducted to diminish the drawbacks caused by this effect in extremely large-scale antenna array (ELAA) systems recently [25], [26], [27], which are based on the fact that though the entire channel is spatial non-stationary, the sub-channel corresponding to a partial of the array (i.e., a sub-array) can be treated as spatial stationary. For example, in [25], the antenna array was divided equally into several spatial stationary sub-arrays, and the OMP scheme was applied to each sub-array. The whole channel can thus be recovered by combining all the estimated sub-channels. Meanwhile, the authors in [26] applied the Bayesian based channel estimation scheme to exploit the sparsity in the antenna domain caused by spatial non-stationarity to realize an accurate channel estimation. To recognize the VR more effectively, a machine learning based scheme was studied in [27] to estimate the VR via a You Only Look Once (YOLO) network by converting the channel estimation problem to an image processing problem. These schemes are all based on fully-digital architectures, where the signals on all antennas can be extracted respectively for analyses.

Unfortunately, up to now, there have been no related works on non-stationary channel estimation for RIS assisted communication systems. A non-negligible difference between the ELAA systems considered in the above papers [25], [26], [27] and the RIS system in this paper is that, in ELAA systems, the BS antennas can directly receive the pilots transmitted by users. Thus, the BS antenna array can be easily divided into several independent sub-arrays for signal processing. Nevertheless, in RIS systems, the BS can only receive the mixed signals reflected by the whole RIS array, but the contributions from different sub-arrays on the RIS are ambiguous. We call this effect as the *spatial ambiguity effect of RIS*, which makes it challenging to decouple RIS into several sub-arrays for signal processing. Authors in [28] studied a similar localization problem in RIS

systems. By assuming that there is only one path between the user and the RIS, the VR could be recognized easily based on the amplitude of the estimated channel matrix. Since the sub-arrays invisible to the user cannot reflect the pilots, the amplitude of the estimated channel matrix corresponding to these sub-arrays approaches zero. However, in practical communication systems, there are usually multiple paths. As a result, the VR corresponding to each path cannot be trivially recognized by the amplitude of the estimated channel matrix and the above scheme cannot work anymore. To the best of our knowledge, existing schemes cannot solve the spatial ambiguity in RIS assisted communication systems. With the spatial non-stationarity, it is challenging to realize the accurate channel estimation in RIS assisted communication systems.

## B. Our Contributions

To deal with the spatial ambiguity effect of RIS, in this paper, we study the spatial non-stationary channel estimation for RIS assisted communication systems, which is realized by an extended group time block code (GTBC) based RIS decoupling scheme<sup>1</sup>. The specific contributions are listed as follows.

- First, to decouple the RIS into several sub-arrays in the presence of the spatial ambiguity effect, we extend the one-dimensional GTBC that we recently studied in [29] to the two-dimensional scenario in RIS assisted communication systems by mapping 1D to 2D according to the geometric structure of RIS and propose a GTBC-based RIS decoupling scheme. This scheme enhances the relevance of the received signals in the time domain, which can break the limits of traditional schemes, where the reflecting matrix of RIS is generated randomly in each time slot. Specifically, the proposed scheme is composed of an encoding stage and a decoding stage. At the encoding stage, the reflecting matrix of the RIS elements in the same group is changed as a whole based on the designed GTBC in each time slot. At the decoding stage, the received signals at the BS in different time slots are combined according to the designed GTBC to decouple the received signals corresponding to each group. By this means, the estimation problem of the entire non-stationary channel can be converted to a series of sub-problems of estimating the stationary sub-channel of each sub-array. Then, CS based channel estimation schemes can be utilized for each spatial stationary sub-channel. By combining all sub-channels, the spatial non-stationary channel can be accurately recovered.
- Since the array aperture of RIS is large in the considered scenario, in addition to the spatial non-stationarity effect, the users are also possibly be located in the near-field region of RIS. In this case, both the angle and the distance need to be utilized to describe the cascaded channel, which leads to a sharp increase in the codebook size, resulting in a large computational complexity and storage burden. To reduce the necessary codebook size for the near-field

<sup>1</sup>Simulation codes will be provided to reproduce the results in this article: <http://oa.ee.tsinghua.edu.cn/dailinglong/publications/publications.html>.

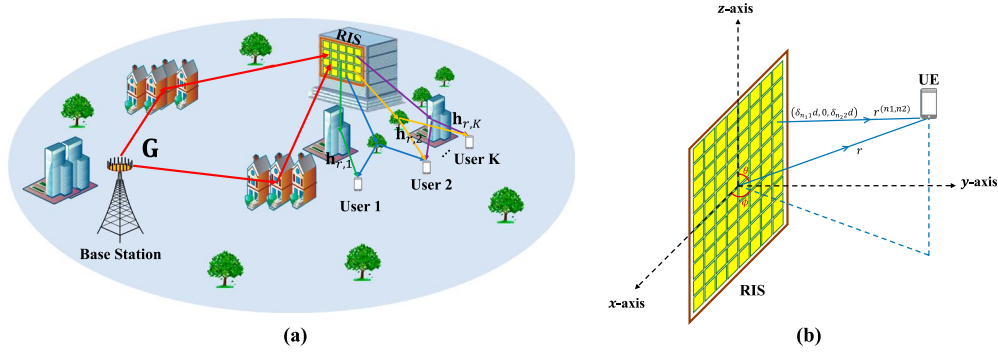


Fig. 1. (a) An example of RIS assisted wireless communication system and (b) the vertical view of the system.

cascaded channel and represent the RIS cascaded channel effectively, we abandon the traditional polar-domain codebook and propose a dynamic sparse codebook, which is applicable for both far-field channels and near-field channels. By regarding the distance as a parameter, the size of the proposed codebook is the same as that of the traditional far-field codebook, which is far smaller than that of traditional near-field codebook. Thus, the proposed codebook can significantly reduce the storage burden and the computational complexity in near-field scenarios.

- After converting the spatial non-stationary channel to several spatial stationary sub-channels by the proposed GTBC-based RIS decoupling scheme and representing them with the proposed dynamic codebooks, we further propose a GTBC-based dynamic orthogonal matching pursuit (GD-OMP) scheme to estimate the spatial non-stationary cascaded channel in RIS assisted communication systems. For each path, the cascaded angle is firstly estimated. Then, we exploit the geometric structure of adjacent sub-arrays to dynamically update the distance and the corresponding codebook. After a few rounds of iteration, the parameter of each path can thus be estimated. After estimating all paths sequentially, the spatial stationary sub-channel can be recovered, and the entire spatial non-stationary channel can be accurately estimated by combining all sub-channels. In addition, the computational complexity of the proposed GD-OMP scheme is analyzed. Finally, simulation results reveal that the proposed GD-OMP scheme can capture the spatial non-stationarity effect in RIS systems and realize a more accurate channel estimation than existing schemes.

### C. Organization and Notation

1) *Organization*: The remainder of this paper is organized as follows. In Section II, we introduce the system model, where the signal model of the RIS assisted communication system and the spatial non-stationary channel model are elaborated. In Section III, the two-dimensional GTBC-based RIS decoupling scheme is proposed. In Section IV, the proposed dynamic sparse codebook for the cascaded RIS channel and the corresponding GD-OMP scheme are elaborated, followed by the computational complexity analysis. Simulation results

are provided in Section V, and conclusions are finally drawn in Section VI.

2) *Notation*: Lower-case and upper-case boldface letters represent vectors and matrices, respectively;  $\mathbf{X}[i, j]$  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;  $\|\cdot\|_F$  denotes the Frobenius norm operator;  $\odot$  denotes the Hadamard product operator;  $\otimes$  denotes the Kronecker product operator;  $\bullet$  denotes the column-wise Khatri-Rao product operator;  $\lceil \cdot \rceil$  denotes the ceiling operator;  $\text{Tr}(\cdot)$  denotes the trace operator;  $\mathcal{CN}(\mu, \Sigma)$  and  $\mathcal{U}(a, b)$  denote the Gaussian distribution with mean  $\mu$  and covariance  $\Sigma$ , and the uniform distribution between  $a$  and  $b$ , respectively.

## II. SYSTEM MODEL

In this section, we will firstly introduce the signal model of the considered RIS assisted communication system. Then, the spatial non-stationary channel model of the RIS assisted communication system will be elaborated.

### A. Signal Model

An uplink time division duplexing (TDD) system is considered in this paper. As illustrated in Fig. 1(a), a RIS is deployed between the BS equipped with an  $N_t = N_{t1} \times N_{t2}$ -antenna uniform planar array (UPA) and  $K$  single-antenna users. The RIS is composed of  $M = M_1 \times M_2$  elements. The direct links between the BS and the users are assumed to be blocked by obstacles such as trees or buildings [30]. When conducting channel estimation, orthogonal pilot sequences are widely utilized [31], so we only consider an arbitrary user for the uplink channel estimation. Then the received pilot  $\mathbf{y}_p \in \mathbb{C}^{N_t \times 1}$  in the  $p$ -th time slot at the BS can be represented as

$$\mathbf{y}_p = \mathbf{G} \text{diag}(\mathbf{v}_p) \mathbf{h}_r s_p + \mathbf{n} = \mathbf{G} \text{diag}(\mathbf{h}_r) \mathbf{v}_p s_p + \mathbf{n}_p, \quad (1)$$

where  $\mathbf{G} \in \mathbb{C}^{N_t \times M}$  denotes the channel between the RIS and the BS;  $\mathbf{v}_p \in \mathbb{C}^{M \times 1}$  denotes the reflecting vector of the RIS at the  $p$ -th time slot;  $\mathbf{h}_r \in \mathbb{C}^{M \times 1}$  denotes the channel between the user and the RIS;  $s_p$  denotes the pilot sent by the user at the  $p$ -th time slot; and  $\mathbf{n}_p \in \mathbb{C}^{N_t \times 1}$  denotes the additive white Gaussian complex noise at the  $p$ -th time slot satisfying  $\mathcal{CN}(0, \sigma^2 \mathbf{I}_{N_t})$ .



with  $\sigma^2$  being the noise power, respectively. The reflecting vector of the RIS  $\mathbf{v}_p = [\beta_1 e^{j\vartheta_1}, \beta_2 e^{j\vartheta_2}, \dots, \beta_M e^{j\vartheta_M}]$ , where  $\beta_n \in [0, 1]$ ,  $n = 1, 2, \dots, N$  represents the amplitude and  $\vartheta_n \in [0, 2\pi]$ ,  $n = 1, 2, \dots, N$  represents the phase shift of the  $n$ -th element, respectively. Let  $P$  be the pilot length during the channel estimation procedure. Thus, by assuming that  $s_p = 1$ ,  $p = 1, 2, \dots, P$ , the overall received pilots  $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_P] \in \mathbb{C}^{N_t \times P}$  can be represented as

$$\mathbf{Y} = \mathbf{G} \text{diag}(\mathbf{h}_r) \mathbf{V} + \mathbf{n}, \quad (2)$$

where  $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_P] \in \mathbb{C}^{M \times P}$  denotes the overall reflecting matrix of the RIS, and  $\mathbf{n} = [\mathbf{n}_1, \mathbf{n}_2, \dots, \mathbf{n}_P] \in \mathbb{C}^{N_t \times P}$  denotes the overall noise, respectively. In high-frequency communication systems, the path loss of the signal is very high, so the number of paths is small, which enables the effective channel estimation. Specifically, we can re-write the cascaded channel  $\mathbf{H}_c$  as

$$\mathbf{H}_c = \mathbf{G} \text{diag}(\mathbf{h}_r) = \mathbf{U}_{\text{BS}} \tilde{\mathbf{G}} \mathbf{U}_{\text{RIS}_G}^T \text{diag}(\mathbf{U}_{\text{RIS}_r} \tilde{\mathbf{h}}_r), \quad (3)$$

where  $\mathbf{U}_{\text{BS}}$ ,  $\mathbf{U}_{\text{RIS}_G}$ ,  $\mathbf{U}_{\text{RIS}_r}$  denote the codebook at the BS, the codebook at RIS of BS-RIS channel and the codebook at RIS of RIS-UE channel, respectively. Thanks to the small number of paths, the  $\tilde{\mathbf{G}}$  and  $\tilde{\mathbf{h}}_r$  only have a few elements that do not equal zero. By this means, we can utilize several compressed sensing based schemes to recover  $\tilde{\mathbf{G}}$  and  $\tilde{\mathbf{h}}_r$  effectively. To properly extract the sparsity in the channel, we need to design the codebook  $\mathbf{U}_{\text{BS}}$ ,  $\mathbf{U}_{\text{RIS}_G}$  and  $\mathbf{U}_{\text{RIS}_r}$  according to the channel characteristic, which is elaborated in Section IV-A.

### B. Spatial Non-Stationary Channel Model

Due to the “multiplicative fading” effect introduced by RIS, the equivalent path loss of the BS-RIS-UE link is the product (rather than the sum) of the path losses of the BS-RIS and RIS-UE links [32]. To leverage the benefits brought by RIS, we prefer to deploy RIS either near the BS side or near the UE side. Considering the relatively high altitude of BS and RIS, the propagation environment of the BS-RIS channel is simpler than that of RIS-UE channel. Therefore, in this paper, we consider the scenario where the RIS is deployed near the UE so as to provide stronger ability to manipulate the RIS-UE channel. Based on the above assumption, the electromagnetic wave between the BS and the RIS experiences a planar wavefront since the distance between the BS and the RIS is large. Then, the channel  $\mathbf{G}$  between the RIS and the BS can be represented as

$$\mathbf{G} = \sqrt{\frac{N_t M}{L_G}} \sum_{\ell=1}^{L_G} \alpha_\ell^G \mathbf{a}(\phi_\ell^{G_r}, \theta_\ell^{G_r}) \left( \mathbf{a}(\phi_\ell^{G_t}, \theta_\ell^{G_t}) \odot \mathbf{q}(\Upsilon_\ell^G) \right)^T, \quad (4)$$

where  $L_G$  represents the number of paths between the RIS and the BS;  $\alpha_\ell^G$  represents the path gain of the  $\ell$ -th path;  $\phi_\ell^{G_r}, \theta_\ell^{G_r}, \phi_\ell^{G_t}, \theta_\ell^{G_t}$  represent the azimuth angle at the BS, the elevation angle at the BS, the azimuth angle at the RIS, the elevation angle at the RIS for the  $\ell$ -th path, respectively. Besides we use  $\mathbf{q}(\Upsilon_\ell^G) \in \{0, 1\}^{M \times 1}$  to represent the mask of the VR, the function of which is to select the RIS elements visible to the

$\ell$ -th scatterer. The steering vector  $\mathbf{a}(\phi, \theta)$  for a  $N = N_1 \times N_2$ -antenna UPA can be elaborated as

$$\mathbf{a}(\phi, \theta) = \frac{1}{\sqrt{N}} \left[ e^{-j2\pi d \sin(\phi) \sin(\theta) \delta_1 / \lambda} \right] \otimes \left[ e^{-j2\pi d \cos(\theta) \delta_2 / \lambda} \right], \quad (5)$$

where  $\lambda = c/f_c$  denotes the wavelength of the electromagnetic wave with  $f_c$  being the central frequency and  $c$  being the speed of light, and  $d$  denotes the antenna spacing satisfying  $d = \lambda/2$ , respectively. Here, we denote the antenna index as  $\delta_1$  and  $\delta_2$ , which is presented as

$$\begin{aligned} \delta_1 &= [\delta_{11}, \delta_{21}, \dots, \delta_{N_1 1}]^T = \left[ \frac{1-N_1}{2}, \frac{3-N_1}{2}, \dots, \frac{N_1-1}{2} \right]^T \\ \delta_2 &= [\delta_{12}, \delta_{22}, \dots, \delta_{N_2 2}]^T = \left[ \frac{1-N_2}{2}, \frac{3-N_2}{2}, \dots, \frac{N_2-1}{2} \right]^T. \end{aligned} \quad (6)$$

For the mask vector  $\mathbf{q}(\Upsilon_\ell^G)$ ,  $\Upsilon_\ell^G$  is the set of the RIS elements which can see the  $\ell$ -th scatterer. Accordingly,  $\mathbf{q}(\Upsilon_\ell^G)$  can be elaborated as

$$[\mathbf{q}(\Upsilon_\ell^G)]_n = \begin{cases} 1, & n \in \Upsilon_\ell^G \\ 0, & n \notin \Upsilon_\ell^G \end{cases}. \quad (7)$$

According to [22], [23], though the entire channel corresponding to the RIS is spatial non-stationary, it is reasonable to suppose that the sub-channels associated with a part of the array (i.e., a sub-array) are spatial stationary. To be specific, we divide the RIS into  $M_s = M_{s,1} \times M_{s,2}$  sub-arrays, then (7) can be further simplified as

$$[\mathbf{q}(\Upsilon_\ell)]_n = \begin{cases} 1, & f(n) \in \tilde{\Upsilon}_\ell^G \\ 0, & f(n) \notin \tilde{\Upsilon}_\ell^G \end{cases}, \quad (8)$$

where the sub-arrays visible to the  $\ell$ -th scatterer are indexed by  $\tilde{\Upsilon}_\ell^G = \{m_{s,1}, \dots, m_{s,S}\}$ , with  $1 \leq m_{s,i} \leq M_s$  being the index of the sub-array. The function  $f(n)$  represents the mapping function from the index of element to the index of sub-array, which can be elaborated as

$$f(n) = \left( \left\lceil \frac{n M_{s,1}}{M} \right\rceil - 1 \right) M_{s,2} + \left\lceil \frac{\text{mod}(n, M_2)}{M_2 / M_{s,2}} \right\rceil. \quad (9)$$

For the channel between the RIS and the user, since the RIS is deployed near the user, spherical wavefront should be considered to formulate  $\mathbf{h}_r$  as

$$\mathbf{h}_r = \sqrt{\frac{M}{L_r}} \sum_{\ell=1}^{L_r} \alpha_\ell^r \mathbf{b}(r_\ell^r, \phi_\ell^r, \theta_\ell^r) \odot \mathbf{q}(\Upsilon_\ell^r), \quad (10)$$

where  $L_r$  represents the number of paths between the user and the RIS;  $\alpha_\ell^r$  represents the path gain of the  $\ell$ -th path;  $r_\ell^r, \phi_\ell^r, \theta_\ell^r$  represent the distance between the user and the RIS; the azimuth angle at the RIS; and the elevation angle at the RIS, respectively, and  $\mathbf{q}(\Upsilon_\ell^r) \in \{0, 1\}^{M \times 1}$  represents the mask of the VR which selects the RIS elements visible to the  $\ell$ -th scatterer. The near-field steering vector  $\mathbf{b}(r, \phi, \theta)$  of an  $N = N_1 \times N_2$ -antenna UPA can be formulated as

$$\mathbf{b}(r, \phi, \theta) = \frac{1}{\sqrt{N}} \left[ e^{-j2\pi r^{(1,1)} / \lambda}, \dots, e^{-j2\pi r^{(1, N_2)} / \lambda}, \dots, e^{-j2\pi r^{(N_1, 1)} / \lambda}, \dots, e^{-j2\pi r^{(N_1, N_2)} / \lambda} \right]^T, \quad (11)$$

where  $r^{(n_1, n_2)}$  represents the distance between the user and the  $(n_1, n_2)$ -th antenna, which can be elaborated as [33]

$$\begin{aligned}
 r^{(n_1, n_2)} &= ((r \sin \theta \cos \phi)^2 + (\delta_{n_1 1} d - r \sin \theta \sin \phi)^2 \\
 &\quad + (\delta_{n_2 2} d - r \cos \theta)^2)^{1/2} \\
 &\stackrel{(a)}{\approx} r - \delta_{n_1 1} d \sin \theta \sin \phi - \delta_{n_2 2} d \cos \theta + \frac{\delta_{n_2 2}^2 d^2 \sin^2 \theta}{2r} \\
 &\quad + \frac{\delta_{n_1 1}^2 d^2 (1 - \sin^2 \theta \sin^2 \phi)}{2r} - \frac{\delta_{n_1 1} \delta_{n_2 2} d^2 \sin \theta \cos \theta \sin \phi}{r},
 \end{aligned} \quad (12)$$

where  $\delta_{n_1 1} = \frac{2n_1 - N_1 - 1}{2}$  and  $\delta_{n_2 2} = \frac{2n_2 - N_2 - 1}{2}$  denote the antenna indices and (a) is derived by keeping the first two-order of the Taylor series expansion. The mask vector  $\mathbf{q}(\Upsilon_\ell^r)$  can be obtained similarly as (8). Although some existing studies have developed the channel estimation scheme in near-field scenarios [34], the non-stationarity effect has not been considered.

According to (4) and (10), the spatial non-stationary mask vector  $\mathbf{q}(\cdot)$  brings new challenge for the RIS assisted communication system. Since the RIS can only reflect the incident signals, the BS can merely receive the mixed signals reflected by the whole RIS array, but the contributions from different sub-arrays are ambiguous. This phenomenon, also called the spatial ambiguity effect of RIS in this paper, leads to the mismatch between the traditional channel estimation schemes and the actual non-stationary channel of RIS systems. The passive characteristic of RIS further makes it challenging to decouple the reflected signals from different sub-arrays. Therefore, addressing the spatial ambiguity effect of RIS is necessary to realize accurate non-stationary channel estimation.

### III. PROPOSED GTBC-BASED RIS DECOUPLING SCHEME

To effectively deal with the spatial ambiguity effect and recognize the spatial non-stationarity effect on RIS, the sub-channels associated with different spatial stationary sub-arrays should be decoupled. In this section, we extend the one-dimensional group time block code (GTBC) studied in [29] to the two-dimensional scenario in RIS systems and propose a two-dimensional GTBC based RIS decoupling scheme to extract the reflected signal of different sub-arrays in RIS assisted communication systems.

#### A. Encoding Stage

Due to the passive characteristic of RIS, the received signal at the BS is the mixture of all  $M_s = M_{s,1} \times M_{s,2}$  sub-arrays (groups), which is hard to decouple. The main reason for this is that the reflecting vector of the RIS is randomly and independently generated at different time slots in traditional schemes. This independence at the time domain makes it impossible to extract the impact of a certain sub-array to the received signal. In order to obtain the ability to decouple different sub-arrays, it is necessary to design a consistent behavior of the reflecting vector belonging to each sub-array among different time slots.

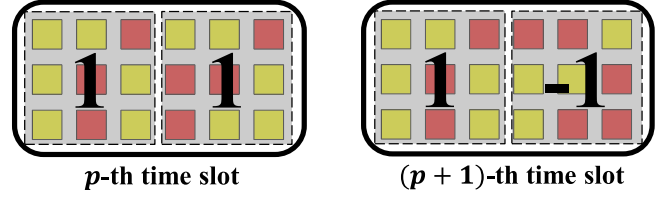


Fig. 2. A simple example of the encoding stage.

We consider the received signal at the  $p$ -th time slot, given by

$$\mathbf{y}_p = \mathbf{G} \text{diag}(\mathbf{h}_r) \mathbf{v}_p + \mathbf{n}_p = \mathbf{H}_c \mathbf{v}_p + \mathbf{n}_p, \quad (13)$$

where  $\mathbf{H}_c = \mathbf{G} \text{diag}(\mathbf{h}_r) \in \mathbb{C}^{N_t \times M}$  is the BS-RIS-UE cascaded channel, which can be further rewritten as  $\mathbf{H}_c = [\mathbf{H}_c^{(1)}, \mathbf{H}_c^{(2)}, \dots, \mathbf{H}_c^{(M_s)}]$ , where  $\mathbf{H}_c^{(m)} \in \mathbb{C}^{N_t \times M/M_s}$ ,  $m = 1, 2, \dots, M_s$  denotes the sub-channel associated with the  $m$ -th sub-array. To reveal the key idea of the encoding stage of the proposed scheme, we first discuss a simplified example that  $M_{s,1} = 2, M_{s,2} = 1$ . As illustrated in Fig. 2, in the  $p$ -th time slot, the received signal  $\mathbf{y}_p$  can be rewritten as

$$\begin{aligned}
 \mathbf{y}_p &= \begin{bmatrix} \mathbf{H}_c^{(1)} & \mathbf{H}_c^{(2)} \end{bmatrix} \begin{bmatrix} \mathbf{v}_p^{(1)} \\ \mathbf{v}_p^{(2)} \end{bmatrix} + \mathbf{n}_p \\
 &= \mathbf{H}_c^{(1)} \mathbf{v}_p^{(1)} + \mathbf{H}_c^{(2)} \mathbf{v}_p^{(2)} + \mathbf{n}_p,
 \end{aligned} \quad (14)$$

where  $\mathbf{v}_p^{(1)}$  and  $\mathbf{v}_p^{(2)} \in \mathbb{C}^{M/2 \times 1}$  denote the reflecting vector corresponding to the first and second sub-array of RIS, respectively. In order to discriminate the two items in (14), in the  $(p+1)$ -th time slot, the reflecting vector  $\mathbf{v}_{p+1}$  should have a similar form as  $\mathbf{v}_p$ . For example, let  $\mathbf{v}_{p+1}^{(1)} = \mathbf{v}_p^{(1)}$ ,  $\mathbf{v}_{p+1}^{(2)} = -\mathbf{v}_p^{(2)}$ , and the received signal at the  $(p+1)$ -th time slot can be represented as

$$\begin{aligned}
 \mathbf{y}_{p+1} &= \begin{bmatrix} \mathbf{H}_c^{(1)} & \mathbf{H}_c^{(2)} \end{bmatrix} \begin{bmatrix} \mathbf{v}_{p+1}^{(1)} \\ \mathbf{v}_{p+1}^{(2)} \end{bmatrix} + \mathbf{n}_{p+1} \\
 &= \mathbf{H}_c^{(1)} \mathbf{v}_p^{(1)} - \mathbf{H}_c^{(2)} \mathbf{v}_p^{(2)} + \mathbf{n}_{p+1}.
 \end{aligned} \quad (15)$$

The above encoding operation can be abstracted as

$$\begin{bmatrix} \mathbf{v}_p & \mathbf{v}_{p+1} \end{bmatrix} = \begin{bmatrix} \mathbf{v}_p^{(1)} & \mathbf{v}_p^{(1)} \\ \mathbf{v}_p^{(2)} & -\mathbf{v}_p^{(2)} \end{bmatrix} \underbrace{\begin{bmatrix} \mathbf{1} & \mathbf{1} \\ \mathbf{1} & -\mathbf{1} \end{bmatrix}}_{\mathbf{P}} = \begin{bmatrix} \mathbf{v}_p^{(1)} & \mathbf{v}_p^{(1)} \\ \mathbf{v}_p^{(2)} & -\mathbf{v}_p^{(2)} \end{bmatrix}. \quad (16)$$

Here, the matrix  $\mathbf{P}$  is termed as a GTBC matrix in this paper, which plays an important role in generating the reflecting vectors of RIS.

The above case is similar to the one-dimensional case studied in [29]. For a non-trivial two-dimensional case where the RIS is divided into arbitrary  $M_s = M_{s,1} \times M_{s,2}$  sub-arrays, the GTBC matrix  $\mathbf{P}$  should be properly designed to decouple the reflecting signals of  $M_s$  sub-arrays based on every  $M_s$  adjacent time slots. To effectively extract the reflecting signals,  $\mathbf{P}$  should satisfy: 1) the elements in  $\mathbf{P}$  should not require an accurate phase shift since the phase shifters on RIS might only be able

to provide low resolution phase states; 2)  $\mathbf{P} \in \mathbb{C}^{M_s \times M_s}$  and  $\text{rank}(\mathbf{P}) = M_s$  so as to allow the matrix inversion in decoding stage; 3) each column of  $\mathbf{P}$  should be orthogonal to each other to eliminate the noise. To meet the above requirements, Hadamard matrix is selected. Specifically, Hadamard matrix is composed of 1 and  $-1$ , which meets condition 1). Denote  $n$ -order Hadamard matrix as  $\mathbf{\Lambda}_n$ . Then, the Hadamard matrix with order  $2^k$  can be derived by

$$\mathbf{\Lambda}_{2^k} = \begin{bmatrix} \mathbf{\Lambda}_{2^{k-1}} & \mathbf{\Lambda}_{2^{k-1}} \\ \mathbf{\Lambda}_{2^{k-1}} & -\mathbf{\Lambda}_{2^{k-1}} \end{bmatrix}, \quad (17)$$

and the initial condition is  $\mathbf{\Lambda}_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ . From the above recursion, we can observe that  $\frac{1}{n} \mathbf{\Lambda}_n^H \mathbf{\Lambda}_n = \mathbf{I}_{n \times n}$ . That is,  $\frac{1}{\sqrt{n}} \mathbf{\Lambda}_n$  is an orthogonal matrix, which naturally satisfy the above condition 2) and 3).

Considering the reflecting vectors from the  $p$ -th time slot to the  $(p + M_s - 1)$ -th time slot,  $M_s$  reflecting vectors need to be generated based on the GTBC matrix  $\mathbf{P} = \mathbf{\Lambda}_{M_s}$  and the basic reflecting vector  $\mathbf{v}_p$ . To describe the encoding process more clearly, we introduce an auxiliary variable  $\tilde{\mathbf{v}}_{p+i} \in \mathbb{C}^{M_2 \times M_1}$  which share the same geometric structure as the RIS array and satisfies  $\text{vec}(\tilde{\mathbf{v}}_{p+i}) = \mathbf{v}_{p+i}$ ,  $i = 0, 1, \dots, M_s - 1$ . Then,  $\tilde{\mathbf{v}}_{p+i}$  can be divided as

$$\tilde{\mathbf{v}}_{p+i} = \begin{bmatrix} \tilde{\mathbf{v}}_{p+i}^{(1,1)} & \tilde{\mathbf{v}}_{p+i}^{(1,2)} & \dots & \tilde{\mathbf{v}}_{p+i}^{(1,M_{s,1})} \\ \tilde{\mathbf{v}}_{p+i}^{(2,1)} & \tilde{\mathbf{v}}_{p+i}^{(2,2)} & \dots & \tilde{\mathbf{v}}_{p+i}^{(2,M_{s,1})} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\mathbf{v}}_{p+i}^{(M_{s,2},1)} & \tilde{\mathbf{v}}_{p+i}^{(M_{s,2},2)} & \dots & \tilde{\mathbf{v}}_{p+i}^{(M_{s,2},M_{s,1})} \end{bmatrix} \quad (18)$$

with  $\tilde{\mathbf{v}}_{p+i}^{(j,k)} \in \mathbb{C}^{\frac{M_2}{M_{s,2}} \times \frac{M_1}{M_{s,1}}}$  being the reflecting vector of the  $((j-1)M_{s,1} + k)$ -th sub-array. Accordingly, the reflecting vector  $\tilde{\mathbf{v}}_{p+i}$  can be obtained by

$$\tilde{\mathbf{v}}_{p+i} = \tilde{\mathbf{v}}_p \odot \begin{bmatrix} \tilde{\mathbf{\Lambda}}_{M_s}^{(i+1,1)} & \dots & \tilde{\mathbf{\Lambda}}_{M_s}^{(i+1,M_{s,1})} \\ \tilde{\mathbf{\Lambda}}_{M_s}^{(i+1,M_{s,1}+1)} & \dots & \tilde{\mathbf{\Lambda}}_{M_s}^{(i+1,2M_{s,1})} \\ \vdots & \ddots & \vdots \\ \tilde{\mathbf{\Lambda}}_{M_s}^{(i+1,M_s-M_{s,1}+1)} & \dots & \tilde{\mathbf{\Lambda}}_{M_s}^{(i+1,M_s)} \end{bmatrix}, \quad (19)$$

where  $\tilde{\mathbf{\Lambda}}_{M_s}^{(i+1,k)}$  denotes the reflecting vector corresponding to the  $k$ -th sub-array, which can be generated by

$$\tilde{\mathbf{\Lambda}}_{M_s}^{(i+1,k)} = \mathbf{\Lambda}_{M_s}^{(i+1,k)} \mathbf{1}_{\frac{M_2}{M_{s,2}} \times \frac{M_1}{M_{s,1}}}. \quad (20)$$

The reflecting vectors from the  $p$ -th time slot to the  $(p + M_s - 1)$ -th time slot can therefore be acquired by

$$\tilde{\mathbf{V}}_p = [\text{vec}(\tilde{\mathbf{v}}_p) \quad \text{vec}(\tilde{\mathbf{v}}_{p+1}) \quad \dots \quad \text{vec}(\tilde{\mathbf{v}}_{p+M_s-1})]. \quad (21)$$

By repeating the above process  $P/M_s$  times, the reflecting matrix  $\mathbf{V}$  in the channel estimation process can be generated by

$$\mathbf{V} = [\tilde{\mathbf{V}}_1 \quad \tilde{\mathbf{V}}_{M_s+1} \quad \dots \quad \tilde{\mathbf{V}}_{P-M_s+1}], \quad (22)$$

with each item being the reflecting vectors in one pilot transmission sub-frame. Based on the designed reflecting vectors, the RIS shifts the configuration in each time slot during the pilot transmission stage. After receiving all the pilot signals, the BS then conducts the decoding stage to decouple each sub-array of the RIS, which is elaborated in the following subsection.

### B. Decoding Stage

During the decoding stage, the received pilots are combined also based on the above GTBC matrix  $\mathbf{P}$  to decouple the reflecting signals of different sub-arrays of the RIS. We also start from the simple case where  $M_{s,1} = 2, M_{s,2} = 1$ . After receiving  $\mathbf{y}_p, \mathbf{y}_{p+1}$ , from (14) and (15), the reflecting signals of the two sub-array can be derived by

$$\begin{cases} \tilde{\mathbf{y}}_p^{(1)} = \frac{\mathbf{y}_p + \mathbf{y}_{p+1}}{2} = \mathbf{H}_c^{(1)} \mathbf{v}_p^{(1)} + \frac{\mathbf{n}_p + \mathbf{n}_{p+1}}{2} \\ \tilde{\mathbf{y}}_p^{(2)} = \frac{\mathbf{y}_p - \mathbf{y}_{p+1}}{2} = \mathbf{H}_c^{(2)} \mathbf{v}_p^{(2)} + \frac{\mathbf{n}_p - \mathbf{n}_{p+1}}{2} \end{cases} \quad (23)$$

Then, the reflecting signals of the two sub-arrays are decoupled without interference. Similar to (16), the above operation can be abstracted as

$$\begin{bmatrix} \tilde{\mathbf{y}}_p^{(1)} & \tilde{\mathbf{y}}_p^{(2)} \end{bmatrix} = [\mathbf{y}_p \quad \mathbf{y}_{p+1}] \underbrace{\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}}_{\mathbf{P}^{-1}}. \quad (24)$$

Extending to a more general two-dimensional case where the RIS is divided into  $M_s = M_{s,1} \times M_{s,2}$  sub-arrays, after receiving  $P$  pilots, the reflecting signals of each sub-array in the  $m$ -th sub-frame can be decoupled by

$$\begin{bmatrix} \tilde{\mathbf{y}}_m^{(1)} & \dots & \tilde{\mathbf{y}}_m^{(M_s)} \end{bmatrix} = [\mathbf{y}_{(m-1)M_s+1} \quad \dots \quad \mathbf{y}_{mM_s}] \underbrace{\frac{1}{M_s} \mathbf{\Lambda}_{M_s}}_{\mathbf{P}^{-1}}. \quad (25)$$

Conducting (25) for all  $P/M_s$  sub-frames, the reflecting signals of the  $k$ -th sub-array can be represented as  $\tilde{\mathbf{Y}}^{(k)} = [\tilde{\mathbf{y}}_1^{(k)}, \tilde{\mathbf{y}}_2^{(k)}, \dots, \tilde{\mathbf{y}}_{P/M_s}^{(k)}] \in \mathbb{C}^{N_t \times P/M_s}$ ,  $k = 1, 2, \dots, M_s$ . Accordingly, the effective reflecting vector  $\tilde{\mathbf{V}}_e^{(k)} \in \mathbb{C}^{M/M_s \times P/M_s}$  corresponding to the  $k$ -th sub-array can be extracted as

$$\tilde{\mathbf{V}}_e^{(k)} = \begin{bmatrix} \text{vec}(\tilde{\mathbf{v}}_1^{(\ell,\iota)}) & \text{vec}(\tilde{\mathbf{v}}_{1+M_s}^{(\ell,\iota)}) & \dots & \text{vec}(\tilde{\mathbf{v}}_{P-M_s+1}^{(\ell,\iota)}) \end{bmatrix}, \quad (26)$$

where  $\tilde{\mathbf{v}}_i^{(\ell,\iota)}$  is defined in (18) and  $(\ell, \iota)$  satisfy  $\ell = \lceil \frac{k}{M_{s,1}} \rceil$ ,  $\iota = \text{mod}(k, M_{s,1})$ . Denoting the sub-channel corresponding to the  $k$ -th sub-array as  $\mathbf{H}_c^{(k)}$ , the received signal after decoding can be rewritten as

$$\tilde{\mathbf{Y}}^{(k)} = \mathbf{H}_c^{(k)} \tilde{\mathbf{V}}_e^{(k)} + \tilde{\mathbf{n}}^{(k)}, \quad (27)$$

where  $\tilde{\mathbf{n}}$  is the effective noise, which can be derived as  $\tilde{\mathbf{n}}^{(k)} = [\tilde{\mathbf{n}}_1^{(k)}, \tilde{\mathbf{n}}_2^{(k)}, \dots, \tilde{\mathbf{n}}_{P/M_s}^{(k)}] \in \mathbb{C}^{N_t \times P/M_s}$ , with  $\tilde{\mathbf{n}}_m^{(k)}$  being the effective noise of the  $k$ -th sub-array in the  $m$ -th sub-frame satisfying

$$\tilde{\mathbf{n}}_m^{(k)} = [\mathbf{n}_{(m-1)M_s+1} \quad \dots \quad \mathbf{n}_{mM_s}] \frac{1}{M_s} \mathbf{\Lambda}_{M_s}[:, k]. \quad (28)$$

After decoupling the received signals of all spatial stationary sub-arrays, the channel estimation can be carried out respectively for each sub-array, and the entire spatial non-stationary can thus be recovered based on the estimated sub-channels of all sub-arrays.

#### IV. PROPOSED NON-STATIONARY CHANNEL ESTIMATION FOR RIS SYSTEMS

To realize an accurate channel estimation for RIS assisted communication systems, in this section, we first propose a low complexity dynamic sparse codebook for the RIS cascaded channel. Then, based on the proposed dynamic codebook, we further propose a GTBC-based dynamic orthogonal matching pursuit (GD-OMP) scheme to estimate the non-stationary channel of RIS assisted communication systems. Finally, the computational complexity of the proposed scheme is analyzed.

##### A. Proposed Dynamic Sparse Codebook for RIS Cascaded Channel

Due to the assumption that the channel corresponding to an arbitrary sub-array is spatial stationary, the problem of estimating a large size spatial non-stationary cascaded channel can be converted to a series of problems of estimating a relatively small size spatial stationary cascaded channel. Therefore, the analysis in this subsection is aimed at the spatial stationary channel for brevity, and the mask vector  $\mathbf{q}(\cdot)$  in (4) and (10) can be omitted. In order to estimate the cascaded channel with a low pilot overhead, the sparsity of the high frequency channel could be fully utilized. Thus, we rewrite  $\mathbf{H}_c$  in a sparse way as

$$\mathbf{H}_c = \mathbf{G} \text{diag}(\mathbf{h}_r) = \mathbf{U}_{\text{BS}} \tilde{\mathbf{G}} \mathbf{U}_{\text{RIS}}^T \text{diag}(\mathbf{D} \tilde{\mathbf{h}}_r), \quad (29)$$

where  $\mathbf{U}_{\text{BS}}$  and  $\mathbf{U}_{\text{RIS}}$  are the codebooks for the BS and RIS in the channel between the RIS and the BS, respectively. Since the distance between RIS and UE is small, the codebook at RIS should consider the near-field effect and need to be designed carefully. For clarity, we denote the codebook for the RIS in RIS-UE channel as  $\mathbf{D}$ . According to the planar wave assumption, the different paths is only related to the azimuth and elevation angle, and the DFT codebook is suitable to extract the sparsity in  $\mathbf{G}$ , which is defined as

$$\mathbf{U}_{\text{BS}} = [\mathbf{a}(\phi_1^B, \theta_1^B), \dots, \mathbf{a}(\phi_{N_{t2}}^B, \theta_{N_{t2}}^B), \dots, \mathbf{a}(\phi_{N_{t1}}^B, \theta_1^B), \dots, \mathbf{a}(\phi_{N_{t1}}^B, \theta_{N_{t2}}^B)], \quad (30)$$

$$\mathbf{U}_{\text{RIS}} = [\mathbf{a}(\phi_1^R, \theta_1^R), \dots, \mathbf{a}(\phi_{N_{t2}}^R, \theta_{N_{t2}}^R), \dots, \mathbf{a}(\phi_{M_1}^R, \theta_1^R), \dots, \mathbf{a}(\phi_{M_1}^R, \theta_{M_2}^R)], \quad (31)$$

where  $\theta_i^B$  and  $\phi_j^B$  satisfy  $\cos(\theta_i^B) = -1 + \frac{2i}{N_{t2}}, i = 1, 2, \dots, N_{t2}$  and  $\sin(\phi_j^B) \sin(\theta_i^B) = -1 + \frac{2j}{N_{t1}}, j = 1, 2, \dots, N_{t1}$ , and  $\theta_i^R$  and  $\phi_j^R$  satisfy  $\cos(\theta_i^R) = -1 + \frac{2i}{M_2}, i = 1, 2, \dots, M_2$  and  $\sin(\phi_j^R) \sin(\theta_i^R) = -1 + \frac{2j}{M_1}, j = 1, 2, \dots, M_1$ , respectively.  $\mathbf{D}$  is the sparse codebook for  $\mathbf{h}_r$ . Due to the spherical wave propagation, there exists energy spread effect in the angular domain as illustrated in [34] and the channel can no longer be represented sparsely by the DFT codebook.

To exploit the sparsity in the near-field channel, researchers in [34] and [35] proposed the polar-domain codebooks for the ULA and UPA respectively, which accounted for both the angle and the distance of the scatterers/users. However, ascribed to the multiple sampling in the distance domain, the size of the polar-domain codebook is several times larger than the traditional DFT codebook, which imported a large storage burden and computational complexity, especially when the number of antenna elements is large. Moreover, the orthogonality of the polar-domain codebook is not as good as the DFT codebook, which may lead to higher channel estimation error.

To reduce the size of the codebook and improve the orthogonality, we propose a dynamic sparse codebook for the RIS cascaded channel, which regards the distance as a parameter. Specifically, the proposed dynamic codebook can be represented as

$$\mathbf{D} = \mathbf{D}(r) = [\mathbf{b}(r, \phi_1, \theta_1), \dots, \mathbf{b}(r, \phi_1, \theta_{M_2}), \dots, \mathbf{b}(r, \phi_{M_1}, \theta_1), \dots, \mathbf{b}(r, \phi_{M_1}, \theta_{M_2})], \quad (32)$$

where  $r$  denotes the distance parameter and  $\mathbf{b}(\cdot)$  denotes the array response vector in (11). The proposed dynamic codebook  $\mathbf{D} \in \mathbb{C}^{M \times M}$  has the same size as the DFT codebook, which is several times smaller than the existing near-field polar-domain codebooks. To ensure the orthogonality of different codewords in the proposed codebook, like in DFT codebook,  $\theta_i$  and  $\phi_j$  should satisfy  $\cos(\theta_i) = -1 + \frac{2i}{M_2}, i = 1, 2, \dots, M_2$  and  $\sin(\phi_j) \sin(\theta_i) = -1 + \frac{2j}{M_1}, j = 1, 2, \dots, M_1$ , respectively. Then, the column coherence of  $\mathbf{D}$  can be derived by the following **Lemma 1**.

**Lemma 1:** For the codewords with the same distance  $r$  in  $\mathbf{D}$ , the column coherence  $f(\phi_p, \phi_q, \theta_p, \theta_q)$  can be approximated by

$$f(\phi_p, \phi_q, \theta_p, \theta_q) = |\mathbf{b}^T(r, \phi_p, \theta_p) \mathbf{b}(r, \phi_q, \theta_q)| \approx |G(\alpha)G(\beta)|, \quad (33)$$

where  $\alpha = \sqrt{\frac{f_c r \sin \theta_q \sin \phi_q - \sin \theta_p \sin \phi_p}{2c \sin \theta_q \sin \phi_q + \sin \theta_p \sin \phi_p}}, \beta = \sqrt{\frac{f_c r \cos \theta_q - \cos \theta_p}{2c \cos \theta_q + \cos \theta_p}},$  and  $G(\mu) = \frac{C(\mu) + jS(\mu)}{\mu}$  with  $C(\alpha) = \int_0^\alpha \cos(\frac{\pi}{2} t^2) dt$  and  $S(\alpha) = \int_0^\alpha \sin(\frac{\pi}{2} t^2) dt$  being the Fresnel functions [36].

*Proof:* See Appendix A.  $\square$

In high frequency bands,  $f_c r$  is far larger than  $c$ , so  $\alpha$  and  $\beta$  are both large, which means  $G(\alpha)G(\beta)$  is sufficiently small. To reveal this proposition, we give an example when  $f_c = 30$ ,  $M_2 = 64$ ,  $r = 10$  m. The value of  $G(\beta)$  versus  $(\cos \theta_q + \cos \theta_p)$  and  $(\cos \theta_q - \cos \theta_p)$  is depicted in Fig. 3. It can be seen that the value of  $G(\beta)$  approaches 1 only when  $\cos \theta_p = \cos \theta_q$ . Similarly, the value of  $G(\alpha)$  approaches 1 only when  $\sin \theta_p \sin \phi_p = \sin \theta_q \sin \phi_q$ . Thus, the coherence of any two different codewords in the proposed dynamic codebook can be neglected. The proposed dynamic codebook can capture the distance information in the near-field channel, so it can better represent the near-field channel than the traditional DFT codebook. In addition, since it views the distance as a parameter, it can reduce



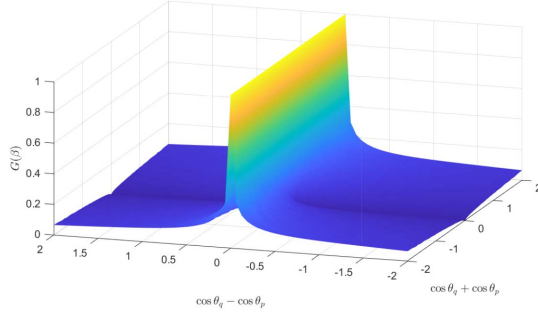


Fig. 3.  $G(\beta)$  vs.  $(\cos \theta_q + \cos \theta_p)$  and  $(\cos \theta_q - \cos \theta_p)$ .

the computational complexity to a large extent compared to the traditional polar-domain codebook.

Based on the proposed dynamic codebook, the cascaded RIS channel can be sparsely represented as

$$\mathbf{H}_c = \mathbf{U}_{\text{BS}} \tilde{\mathbf{G}} \mathbf{U}_{\text{RIS}}^T \text{diag}(\mathbf{D}(r) \tilde{\mathbf{h}}_r). \quad (34)$$

In order to determine the parameter  $r$  in  $\mathbf{D}(r)$ , in the following subsection, we propose a GTBC-based dynamic orthogonal matching pursuit (GD-OMP) scheme to alternatively update the parameter  $r$  and estimate each path of the cascaded channel.

### B. Proposed GTBC-Based Dynamic Orthogonal Matching Pursuit Scheme

While estimating the spatial non-stationary channel  $\mathbf{H}_c$ , the RIS is first decoupled into  $M_s = M_{s,1} \times M_{s,2}$  sub-arrays by the proposed GTBC-based RIS decoupling scheme in Section III, with each sub-array regarded as spatial stationary. In the following analysis, the index of sub-array  $k$  in (27) is omitted for brevity.

Due to the angular sparsity at the BS,  $\mathbf{H}_c$  can be converted to a row-sparse matrix by

$$\mathbf{U}_{\text{BS}}^H \mathbf{H}_c = \tilde{\mathbf{G}} \mathbf{U}_{\text{RIS}}^T \text{diag}(\mathbf{D}(r) \tilde{\mathbf{h}}_r), \quad (35)$$

where only a few (i.e.,  $L_G$ ) rows of  $\tilde{\mathbf{G}}$  are not equal to zero. Then, the row support  $\Omega_r$  for  $\tilde{\mathbf{G}}$  can be obtained by selecting the rows with maximum received powers. Specifically, the received powers of the  $n$ -th angle can be denoted as  $\rho(n) = \|(\mathbf{U}_{\text{BS}}^H \mathbf{Y})[n, :]\|_F^2$ , and the estimated row support for  $\mathbf{H}_c$  can be derived as

$$\hat{\Omega}_r = \Gamma_{\mathcal{T}}(\rho, L_G), \quad (36)$$

where  $\mathcal{T}(\rho, L_G)$  denotes the prune operator, which sets all but  $L_G$  elements with the largest values in  $\rho$  to zero, and  $\Gamma(\rho)$

denotes the support of  $\rho$  which contains all the indices of non-zero elements.

After the row-support of  $\tilde{\mathbf{G}}$  is estimated, the column support related to the RIS should be determined. We take an arbitrary non-zero row in  $\mathbf{U}_{\text{BS}}^H \mathbf{Y}$  as an example for analysis, i.e.,  $\check{\mathbf{y}} = (\mathbf{U}_{\text{BS}}^H \mathbf{Y})[n, :] \in \mathbb{C}^{1 \times P/M_s}$ ,  $n \in \hat{\Omega}_r$ . In order to acquire the capability of updating the distance  $r$  dynamically, each sub-array is further divided into  $2 \times 2 = 4$  parts with  $M_1/M_{s,1}/2 \times M_2/M_{s,2}/2$  elements. It is worth noting that this division is also realized by the proposed RIS decoupling scheme in Section III. Then, the corresponding row-channel of the  $i$ -th part can be represented as

$$\begin{aligned} \check{\mathbf{y}}^{(i)} &= \underbrace{\sqrt{\frac{N_t M}{L_G}} \alpha_{\ell}^G \tilde{\mathbf{G}}^{(i)}[n, :]}_{\tilde{\mathbf{h}}^{(i)}} \mathbf{U}_{\text{RIS}}^T \text{diag}(\mathbf{D}(r) \tilde{\mathbf{h}}_r^{(i)}) \check{\mathbf{v}}^{(i)} \\ &= \sqrt{\frac{N_t M}{L_G}} \alpha_{\ell}^G \underbrace{(\tilde{\mathbf{G}}^{(i)}[n, :] \otimes \tilde{\mathbf{h}}_r^{(i)T})}_{\tilde{\mathbf{G}}_e^{(i)}} \underbrace{(\mathbf{U}_{\text{RIS}}^T \mathbf{D}(r)^T)}_{\tilde{\mathbf{D}}_e^{(i)}} \check{\mathbf{v}}^{(i)}, \end{aligned} \quad (37)$$

where  $\tilde{\mathbf{G}}_e^{(i)} \in \mathbb{C}^{1 \times (M/M_s/4)^2}$  denotes the expanded column support related to RIS,  $\check{\mathbf{v}}^{(i)}$  denotes the decoupled reflecting matrix and  $\tilde{\mathbf{D}}_e^{(i)} \in \mathbb{C}^{(M/M_s/4)^2 \times M/M_s/4}$  denotes the expanded codebook. Compared to the sparse codebook in [19],  $\tilde{\mathbf{G}}_e$  and  $\tilde{\mathbf{D}}_e$  both have a much larger size. This is because the spherical wave propagation imposes the non-linear term related to the distance that cannot be simply summed. To avoid the matrix operations related to the large-size  $\tilde{\mathbf{G}}_e^{(i)}$  and  $\tilde{\mathbf{D}}_e^{(i)}$ , we further divided the estimation of column support into two stages: Estimating the cascaded angles and estimating the exact angles and the distance dynamically, which is elaborated in the following part.

1) *Stage 1: Estimating the cascaded angles.* Selecting an arbitrary path in the channel between the user and the RIS, we get

$$\begin{aligned} \check{\mathbf{h}}_{\ell_2}^{(i)} &= \sqrt{\frac{N_t M}{L_G L_r}} \alpha_{\ell_1}^G \alpha_{\ell_2}^r \mathbf{c}(\phi_{\ell_1}, \theta_{\ell_1}, r_{\ell_2}^{(i)}, \phi_{\ell_2}^{(i)}, \theta_{\ell_2}^{(i)})^T \\ &= \sqrt{\frac{N_t}{L_G L_r}} \alpha_{\ell_1}^G \alpha_{\ell_2}^r \begin{bmatrix} e^{-j2\pi \tilde{r}^{(1,1)}/\lambda}, \dots, e^{-j2\pi \tilde{r}^{(1, \frac{M_2}{2M_{s,2}})}/\lambda}, \\ \dots, e^{-j2\pi \tilde{r}^{(\frac{M_1}{2M_{s,1}}, 1)}/\lambda}, \dots, e^{-j2\pi \tilde{r}^{(\frac{M_1}{2M_{s,1}}, \frac{M_2}{2M_{s,2}})}/\lambda} \end{bmatrix}, \end{aligned} \quad (38)$$

where  $\tilde{r}^{(m_1, m_2)}$  is formulated as (39). Add up all paths, we get  $\check{\mathbf{h}}^{(i)} = \sum_{\ell_2=1}^{L_r} \check{\mathbf{h}}_{\ell_2}^{(i)}$ . In Stage 1, the cascaded azimuth angle  $\varphi_c$

$$\begin{aligned} \tilde{r}^{(m_1, m_2)} &\approx r_{\ell_2}^{(i)} - \delta_{m_1 1} d \underbrace{(\sin \theta_{\ell_1} \sin \phi_{\ell_1} + \sin \theta_{\ell_2}^{(i)} \sin \phi_{\ell_2}^{(i)})}_{\sin \varphi_c \sin \vartheta_c} - \delta_{m_2 2} d \underbrace{(\cos \theta_{\ell_1} + \cos \theta_{\ell_2}^{(i)})}_{\cos \vartheta_c} \\ &\quad + \frac{\delta_{m_1 1}^2 d^2 (1 - \sin^2 \theta_{\ell_2}^{(i)} \sin^2 \phi_{\ell_2}^{(i)})}{2r_{\ell_2}^{(i)}} + \frac{\delta_{m_2 2}^2 d^2 (1 - \cos^2 \phi_{\ell_2}^{(i)})}{2r_{\ell_2}^{(i)}} - \frac{\delta_{m_1 1} \delta_{m_2 2} d^2 \sin \theta_{\ell_2}^{(i)} \cos \theta_{\ell_2}^{(i)} \sin \phi_{\ell_2}^{(i)}}{r_{\ell_2}^{(i)}}. \end{aligned} \quad (39)$$



and the cascaded elevation angle  $\vartheta_c$  can be estimated by the traditional DFT codebook based on the following **Lemma 2**.

**Lemma 2:** For a certain codeword  $\mathbf{a}(\varphi, \vartheta)$  in the DFT codebook  $\mathbf{U} \in \mathbb{C}^{M \times M}$ , only when  $\sin \varphi \sin \vartheta = \sin \varphi_c \sin \vartheta_c$  and  $\cos \vartheta = \cos \vartheta_c$ ,  $|\mathbf{a}(\varphi, \vartheta)^H \mathbf{c}(\phi_{\ell_1}, \theta_{\ell_1}, r_{\ell_2}, \phi_{\ell_2}, \theta_{\ell_2})| \approx 1$ . In other cases,  $|\mathbf{a}(\varphi, \vartheta)^H \mathbf{c}(\phi_{\ell_1}, \theta_{\ell_1}, r_{\ell_2}, \phi_{\ell_2}, \theta_{\ell_2})| \rightarrow 0$ .

*Proof:* See Appendix B.  $\square$

Specifically, the cascaded angles of the path with the maximum power can be obtained according to the coherence between  $\check{\mathbf{y}}^{(i)}$  and each codeword in  $\mathbf{U}$  as

$$\{\hat{\varphi}_c, \hat{\vartheta}_c\} = \arg \max_{\varphi, \vartheta} \left\| \check{\mathbf{y}}^{(i)} \check{\mathbf{V}}^{(i)H} \text{conj}(\mathbf{a}(\varphi, \vartheta)) \right\|_F^2, \quad (40)$$

with  $\mathbf{a}(\varphi, \vartheta)$  being a certain codeword in  $\mathbf{U}$ . Based on the estimated cascaded azimuth angle and elevation angle, the exact angles and the distance of the users can be estimated dynamically in Stage 2.

2) *Stage 2: Estimating the exact angles and the distances dynamically.* The cascaded angles are composed of two parts: the exact angles of the BS-RIS channel and the exact angles of the RIS-UE channel. Based on the estimated cascaded angles, the size of  $\tilde{\mathbf{D}}_e^{(i)}$  in (37) can be dramatically reduced by selecting the rows satisfying the estimated cascaded angles. Specifically, given a certain codeword  $\mathbf{b}(r_{(m_1-1)M_1+m_2}^{(i)}, \phi_{m_1}, \theta_{m_2})$  in  $\mathbf{D}(r^{(i)})$ , the selected codeword in  $\tilde{\mathbf{D}}_e^{(i)}$  should be  $\mathbf{a}(\phi'_{m_1}, \theta'_{m_2})$ , where  $\phi'_{m_1}, \theta'_{m_2}$  satisfy<sup>2</sup>

$$\begin{cases} \sin \phi_{m_1} \sin \theta_{m_2} + \sin \phi'_{m_1} \sin \theta'_{m_2} = \sin \hat{\varphi}_c \sin \hat{\vartheta}_c \\ \cos \theta_{m_2} + \cos \theta'_{m_2} = \cos \hat{\vartheta}_c \end{cases}. \quad (41)$$

Conducting the above selection for all codewords in  $\mathbf{D}(r^{(i)})$ , the initial codebook for Stage 2  $\tilde{\mathbf{D}}_0^{(i)}$  can then be generated by

$$\begin{aligned} \tilde{\mathbf{D}}_0^{(i)} [:(m_1-1)M_1+m_2] \\ = \mathbf{a}(\phi'_{m_1}, \theta'_{m_2}) \odot \mathbf{b}(r_{(m_1-1)M_1+m_2}^{(i)}, \phi_{m_1}, \theta_{m_2}), \\ m_1 = 1, 2, \dots, M_1/M_{s,1}/2, m_2 = 1, 2, \dots, M_2/M_{s,2}/2 \end{aligned} \quad (42)$$

Similarly, the initial estimated exact angles for the BS-RIS channel and RIS-UE channel can be derived by

$$\begin{aligned} \{\hat{\phi}'_0, \hat{\theta}'_0, \hat{\phi}_0^{(i)}, \hat{\theta}_0^{(i)}\} \\ = \arg \max_{\phi', \theta', \phi, \theta} \left\| \check{\mathbf{y}}^{(i)} \check{\mathbf{V}}^{(i)H} \text{conj}(\mathbf{a}(\phi', \theta') \odot \mathbf{b}(r, \phi, \theta)) \right\|_F^2, \end{aligned} \quad (43)$$

with  $\mathbf{a}(\phi', \theta') \odot \mathbf{b}(r, \phi, \theta)$  being a certain codeword in  $\tilde{\mathbf{D}}_0^{(i)}$ . Due to the planar wave propagation between the BS and the RIS, the exact angles  $\hat{\phi}'_0, \hat{\theta}'_0$  would be the same for all different parts in the sub-array. However, the exact angles  $\hat{\phi}_0^{(i)}, \hat{\theta}_0^{(i)}$  would be different for different parts. As illustrated in Fig. 4, different parts might have different angles  $\phi^{(i)}$  and  $\theta^{(i)}$

<sup>2</sup>Here, for each  $(\phi_{m_1}, \theta_{m_2})$ , we need to choose  $(\phi'_{m_1}, \theta'_{m_2})$  that satisfy (41), so the codebook size remains the same as that of the angular-domain codebook.

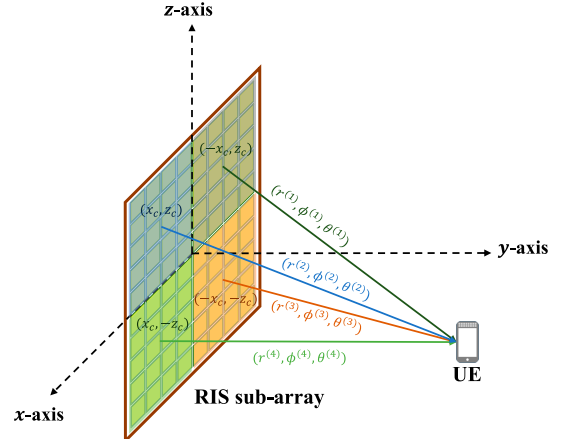


Fig. 4. The geometric structure of the four sub-array parts.

and thus different distances  $r^{(i)}$ , which satisfies the following geometric constraints

$$\underbrace{\begin{bmatrix} 1 & -\tan \phi^{(i)} & 0 \\ 0 & -\frac{1}{\tan \theta^{(i)} \cos \phi^{(i)}} & 1 \end{bmatrix}}_{\mathbf{A}^{(i)}} \underbrace{\begin{bmatrix} x \\ y \\ z \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} (-1)^i x_c \\ (-1)^{[i/2]+1} z_c \end{bmatrix}}_{\boldsymbol{\xi}^{(i)}}, \quad (44)$$

where  $i = 1, 2, 3, 4$  and  $x_c, z_c$  denote the absolute coordinate value of the center of each part as illustrated in Fig. 4. Lying in the intersection of four parts, the coordinate of the user should satisfy

$$\underbrace{\begin{bmatrix} \mathbf{A}^{(1)} \\ \mathbf{A}^{(2)} \\ \mathbf{A}^{(3)} \\ \mathbf{A}^{(4)} \end{bmatrix}}_{\mathbf{A}} \mathbf{x} = \underbrace{\begin{bmatrix} \boldsymbol{\xi}^{(1)} \\ \boldsymbol{\xi}^{(2)} \\ \boldsymbol{\xi}^{(3)} \\ \boldsymbol{\xi}^{(4)} \end{bmatrix}}_{\boldsymbol{\xi}}. \quad (45)$$

Based on the condition in (45), the location of the user and the codebook for Stage 2 can be updated iteratively. Specifically, in the  $k$ -th iteration, the estimated angle  $\{\hat{\phi}'_k, \hat{\theta}'_k, \hat{\phi}_k^{(i)}, \hat{\theta}_k^{(i)}\}$  can be acquired by (43). Then, the coordinate of the user can be estimated by  $\hat{\mathbf{x}}_k = \hat{\mathbf{A}}_k^\dagger \hat{\boldsymbol{\xi}}_k$  according to (45). With the help of the user location, the corresponding distance of the user to each part can be updated as  $\hat{r}_{k+1}^{(i)} = \left\| \hat{\mathbf{x}}_k - [(-1)^i x_c, 0, (-1)^{[i/2]+1} z_c]^T \right\|_F$ . Based on the updated distances, the dynamic codebook in the  $(k+1)$ -th iteration  $\tilde{\mathbf{D}}_{k+1}^{(i)}$  can be updated by

$$\begin{aligned} \tilde{\mathbf{D}}_{k+1}^{(i)} [:(m_1-1)M_1+m_2] \\ = \mathbf{a}(\phi'_{m_1}, \theta'_{m_2}) \odot \mathbf{b}(r_{k+1}^{(i)}, \phi_{m_1}, \theta_{m_2}), \\ m_1 = 1, 2, \dots, M_1/M_{s,1}/2, m_2 = 1, 2, \dots, M_2/M_{s,2}/2 \end{aligned} \quad (46)$$

and the iteration can be continued. The procedure of Stage 2 are summarized in **Algorithm 1**. It is worth noting that when the UE is located in the far field, all angles will be the same and the distance is determined as infinite by the initialization,

**Algorithm 1** Estimating the exact angles and the distances dynamically

**Input:** Cascaded angles  $\hat{\varphi}_c, \hat{\vartheta}_c$ ; initialized distances  $r^{(i)}$ ; codebook update iterations  $K_{\text{iter}}$ .

**Output:** The estimation of the exact angles and distances  $\hat{\phi}_{\ell_1}, \hat{\theta}_{\ell_1}, \hat{\phi}_{\ell_2}^{(i)}, \hat{\theta}_{\ell_2}^{(i)}, \hat{r}_{\ell_2}^{(i)}$ .

- 1: Initialize the dynamic codebook  $\tilde{\mathbf{D}}_0^{(i)}$  by (42).
- 2: **for**  $k = 0$  to  $K_{\text{iter}}$  **do**
- 3:   Estimated the exact angles  $\hat{\phi}'_k, \hat{\theta}'_k, \hat{\phi}_k^{(i)}, \hat{\theta}_k^{(i)}$  by (43).
- 4:   Update the location of the user by  $\hat{\mathbf{x}}_k = \mathbf{A}_k^\dagger \hat{\xi}_k$ .
- 5:   Update the distances to each part by  $\hat{r}_{k+1}^{(i)} = \left\| \hat{\mathbf{x}}_k - [(-1)^i x_c, 0, (-1)^{\lceil i/2 \rceil + 1} z_c]^T \right\|_F$ .
- 6:   Update the dynamic codebook  $\tilde{\mathbf{D}}_{k+1}^{(i)}$  by (46).
- 7: **end for**
- 8: **return**  $\hat{\phi}_{\ell_1} = \hat{\phi}'_{K_{\text{iter}}}, \hat{\theta}_{\ell_1} = \hat{\theta}'_{K_{\text{iter}}}, \hat{\phi}_{\ell_2}^{(i)} = \hat{\phi}_{K_{\text{iter}}}^{(i)}, \hat{\theta}_{\ell_2}^{(i)} = \hat{\theta}_{K_{\text{iter}}}^{(i)}, \hat{r}_{\ell_2}^{(i)} = \hat{r}_{K_{\text{iter}}}^{(i)}$ .

which means this scheme is applicable in both near-field and far-field scenarios.

After obtaining all the exact angles of different paths, the path gains  $\alpha_{\ell_1}^G \alpha_{\ell_2}^r$  of the  $i$ -th part can be obtained by

$$\hat{\alpha}^{(i)} = \mathbf{z}^{(i)} \Psi^{(i)\dagger}, \quad (47)$$

with the  $\ell_2$ -th row of  $\Psi^{(i)}$  being  $(\mathbf{a}(\hat{\phi}_{\ell_1}, \hat{\theta}_{\ell_1}) \odot \mathbf{b}(\hat{r}_{\ell_2}^{(i)}, \hat{\phi}_{\ell_2}^{(i)}, \hat{\theta}_{\ell_2}^{(i)}))^H \check{\mathbf{V}}^{(i)H}$  and  $\mathbf{z}^{(i)}$  being the residual signal with the initial value  $\check{\mathbf{y}}^{(i)}$ . Based on the estimated path gains, the residual signal can thus be updated to estimate the next path. After all paths are estimated, the channel corresponding to the  $\ell_1$ -th path between BS and RIS is estimated. Repeating the above procedure for  $L_G$  times, we can recover the channel corresponding to an arbitrary sub-array, and the whole spatial non-stationary channel can thus be recovered by combining all the spatial stationary sub-channels. The whole procedure of the proposed GD-OMP scheme is summarized in **Algorithm 2**.

### C. Computational Complexity Analysis

For the proposed GD-OMP scheme, the OMP procedure is the dominant factor of the overall computational complexity. In step 5, the power of each row of  $\mathbf{U}_{\text{BS}}^T \mathbf{Y}$  are calculated, and the corresponding computational complexity is  $\mathcal{O}(N_t^2 P)$ . In step 13, the correlation between the received signal and the codewords in  $\mathbf{U}$  are calculated, so the corresponding computational complexity is  $\mathcal{O}(PM^2/(16M_s^3))$ . Considering all the paths belonging to all sub-arrays, the total computational complexity is  $\mathcal{O}(\hat{L}_G \hat{L}_r PM^2/(4M_s^2))$ . In step 15, the computational complexity of dynamically estimating the exact angles and the distances is  $\mathcal{O}(K_{\text{iter}} PM^2/(4M_s^3))$ , and the total computational complexity is  $\mathcal{O}(\hat{L}_G \hat{L}_r K_{\text{iter}} PM^2/(4M_s^3))$ . In step 17 and 18, the computational complexity is  $\mathcal{O}(\hat{L}_r P/4M_s)$ , and the total computational complexity is  $\mathcal{O}(\hat{L}_G \hat{L}_r^2 P)$ . Due to the large RIS element number, the overhead  $P$  is relatively small and can be neglected. Therefore, the overall computational complexity of the proposed GD-OMP scheme is

**Algorithm 2** The proposed GD-OMP scheme

**Input:** Received signal  $\mathbf{Y}$ ; reflecting matrix  $\mathbf{V}$ ; path number  $\hat{L}_G$  between BS and RIS; path number  $\hat{L}_r$  between RIS and UE; sub-array number  $M_s = M_{s,1} \times M_{s,2}$ ; pilot number  $P$ .

**Output:** The estimation of the non-stationary cascaded channel  $\hat{\mathbf{H}}_c$ .

- 1: **for**  $m = 1$  to  $P/M_s$  **do**
- 2:   Decouple the received signals corresponding to each spatial stationary sub-array at the  $m$ -th sub-frame by (25).
- 3: **end for**
- 4: Combine the results in step 2 to acquire the received signals corresponding to each sub-array  $\hat{\mathbf{Y}}^{(k)}$ .
- 5: Return the estimated common row support  $\hat{\Omega}_r$  by (36)
- 6: **for**  $k = 1$  to  $M_s$  **do**
- 7:   **for**  $\ell_1 = 1$  to  $\hat{L}_G$  **do**
- 8:     Decouple the received signals of one non-zero row into  $2 \times 2 = 4$  parts by (25).
- 9:     Initialize  $\mathbf{z}^{(1)} = \check{\mathbf{y}}^{(1)}, \mathbf{z}^{(2)} = \check{\mathbf{y}}^{(2)}, \mathbf{z}^{(3)} = \check{\mathbf{y}}^{(3)}, \mathbf{z}^{(4)} = \check{\mathbf{y}}^{(4)}$
- 10:     **for**  $\ell_2 = 1$  to  $\hat{L}_r$  **do**
- 11:       **for**  $i = 1$  to 4 **do**
- 12:         Obtain the reflecting matrix corresponding to the  $i$ -th part  $\check{\mathbf{V}}^{(i)}$  by (26).
- 13:         Estimate the cascaded angles by (40).
- 14:       **end for**
- 15:       Estimate the exact angles and the distances dynamically by **Algorithm 1**.
- 16:       **for**  $i = 1$  to 4 **do**
- 17:         Estimate the path gains  $\hat{\alpha}^{(i)}$  by (47).
- 18:          $\mathbf{z}^{(i)} = \check{\mathbf{y}}^{(i)} - \hat{\alpha}^{(i)} \Psi^{(i)}$ .
- 19:       **end for**
- 20:     **end for**
- 21:     **for**  $i = 1$  to 4 **do**
- 22:       Recover the channel corresponding to each part by  $\hat{\alpha} \check{\mathbf{V}}^{(i)H}$ .
- 23:     **end for**
- 24:     Combine the four parts to recover the channel corresponding to the  $\ell_1$ -th path  $\hat{\mathbf{H}}_c^{(k)} [\hat{\Omega}_r(\ell_1), :] = \hat{\mathbf{h}}_{\ell_1}$ .
- 25:   **end for**
- 26:    $\hat{\mathbf{H}}_c^{(k)} = \mathbf{U} \hat{\mathbf{H}}_c^{(k)}$ .
- 27: **end for**
- 28: Combine all sub-channels  $\hat{\mathbf{H}}_c^{(k)}$  to recover  $\hat{\mathbf{H}}_c$ .

$\mathcal{O}(N_t^2 P + \hat{L}_G \hat{L}_r K_{\text{iter}} PM^2/(4M_s^2))$ . For the traditional channel estimation scheme utilizing the polar-domain codebook, we need to substitute the codebook  $\mathbf{D}_r \in \mathbb{C}^{M/M_s \times M/M_s}$  into the polar-domain codebook  $\mathbf{U}_{\text{polar}} \in \mathbb{C}^{M/M_s \times MS/M_s}$ , where  $S$  is the number of distance sampling. For the power of each row at the BS, the computational complexity is still  $\mathcal{O}(N_t^2 P)$ . For the estimation of the support set at RIS, the complexity of each path is  $\mathcal{O}(PM^3 S/M_s^4)$ . Considering all paths and all sub-arrays, the computational complexity at RIS is  $\mathcal{O}(\hat{L}_G \hat{L}_r PM^3 S/M_s^3)$ . Therefore, the overall channel estimation complexity of the polar-domain codebook based schemes is

TABLE I  
SIMULATION PARAMETERS

|   |   |
|---|---|
| BS antenna number $N_t$                     | $8 \times 8 = 64$   |
| User number $K$                             | 4   |
| Central frequency $f_c$                     | 30 GHz  |
| The distribution of $\theta, \phi$          | $\mathcal{U}\left(-\frac{\sqrt{3}}{2}, \frac{\sqrt{3}}{2}\right)$ |
| The distribution of $r$                     | $\mathcal{U}(10, 20)$ m   |
| Path number $L_G$ between BS and RIS        | 2   |
| Path number $L_r$ between RIS and each user | 3   |

$\mathcal{O}(N_t^2 P + \hat{L}_G \hat{L}_r P M^3 S / M_s^3)$ . During simulation, the typical value of the number of iterations  $K_{\text{iter}}$  is 2 or 3, which is small. As a result, when the number of elements in each sub-array  $M/M_s$  is large, the computational complexity of polar-domain based channel estimation is much larger than that of our proposed scheme.

## V. SIMULATION RESULTS

In this section, numerical experiments are conducted to verify the effectiveness of the proposed GD-OMP scheme. To evaluate the accuracy of channel estimation, the normalized mean square error (NMSE) is utilized in this section, which is defined as  $\text{NMSE} = \mathbb{E} \left( \frac{\|\mathbf{H} - \hat{\mathbf{H}}\|_2^2}{\|\mathbf{H}\|_2^2} \right)$ . The system parameters are listed in Table I. The path loss between BS and RIS is  $|\alpha_\ell^G| = 10^{-3} d_{\text{BR}}^{-2.2}$  and the path loss between RIS and UE is  $|\alpha_\ell^r| = 10^{-3} d_{\text{RU}}^{-2.8}$ , where  $d_{\text{BR}}$  and  $d_{\text{RU}}$  represent the distance between BS and RIS and the distance between RIS and UE, respectively. The antenna spacing at both BS and RIS is set to  $d = \frac{\lambda}{2} = \frac{c}{2f_c}$ , with  $f_c$  being the central frequency. For each path, the VR is chosen from all sub-arrays randomly. We compare the NMSE performance of the proposed GD-OMP scheme with that of the existing schemes, including the traditional angular-domain OMP scheme in [12], the row-structured OMP scheme in [18], which exploits the sparsity in the rows of the channel matrices to effectively estimate the cascaded channel. These two schemes both utilize the far-field codebook and assume that the channel is spatial stationary. We also apply the polar-domain OMP (P-OMP) channel estimation scheme in [14] and combine the polar-domain OMP scheme with the row-structure sparsity in [18] as two baselines. These two schemes utilize the polar-domain near-field codebook and assume that the channel is spatial stationary. In addition, we have compared the NMSE performance of our proposed GD-OMP scheme with that of [28], which considers the spatial non-stationarity effect and studies a joint localization and channel reconstruction (JLCR) scheme in RIS systems.

Fig. 5 demonstrates the NMSE performance against the pilot length  $P$ . The SNR of the system is set to 10 dB. For Fig. 5(a), the number of RIS elements is set to  $128 \times 1$ . For Fig. 5(b), the number of RIS elements is set to  $64 \times 32$ . The pilot length is increasing from 32 to 256. It is illustrated in Fig. 5 that the NMSE of all schemes decreases with the increment of pilot length. In Fig. 5(a), the number of sub-arrays is set to  $4 \times 1$ . It can be observed that due to the spherical propagation of

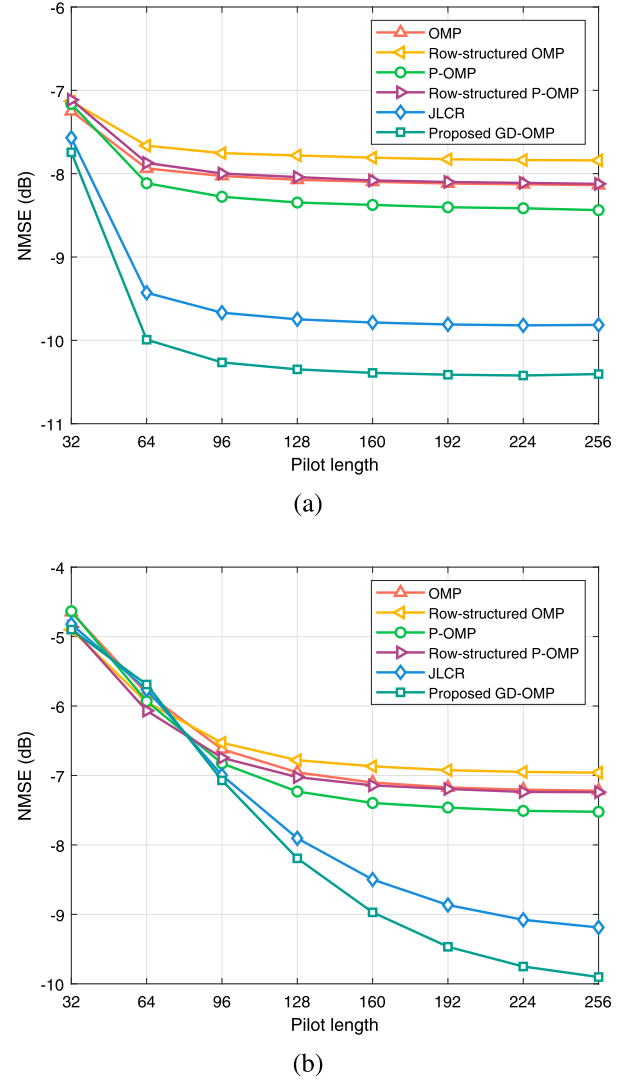


Fig. 5. The NMSE performance v.s. pilot length. The number of elements of RIS (a)  $128 \times 1$  uniform linear array (ULA), (b)  $64 \times 32$  uniform planar array (UPA).

near-field communications, the channel estimation accuracy of polar-domain based schemes (i.e., P-OMP and row structured P-OMP) is better than that of angular-domain based schemes (i.e., OMP and row structured OMP). It is also worth noting that the accuracy of row-structured sparsity based schemes (i.e., row structured OMP and row structured P-OMP) is worse than traditional OMP schemes (i.e., OMP and P-OMP), which seems contradictory to the common sense. The reason for this phenomenon is that due to the spatial non-stationarity effect, the paths between the BS and the RIS may not be the same for different sub-arrays, which leads to the differences of the row-structured sparsity support sets for different sub-arrays. In traditional channel estimation schemes, the spatial non-stationarity effect is neglected and the row-structured sparsity support sets are treated as the same, which will lead to further accuracy loss. For the JLCR scheme, since it considers the spatial non-stationarity effect and recognizes the VRs based on the amplitude of the estimated channel, the channel estimation accuracy

is better than above schemes. However, it cannot distinguish the VRs corresponding to each path clearly, so the NMSE performance is not as good as the proposed GD-OMP scheme. As for the proposed GD-OMP scheme, it is illustrated that thanks to the decoupling of the RIS, all spatial stationary sub-arrays are estimated respectively, which improves the channel estimation accuracy in spatial non-stationary systems. In Fig. 5(b), the number of sub-arrays is set to  $4 \times 2$ . Since the number of elements is larger than that in the  $128 \times 1$  ULA, the needed pilot length also increases. Similar to the ULA case in Fig. 5(a), the polar-domain based schemes outperform the angular-domain based schemes due to the spherical propagation. The traditional OMP schemes outperform the row-structured sparsity based schemes due to the spatial non-stationarity effect. The JLCR scheme can further improve the channel estimation accuracy by considering the spatial non-stationarity effect on RIS. Thanks to the capability of capturing the spatial non-stationarity effect on the RIS array, the proposed GD-OMP scheme realizes a more accurate channel estimation than all the above traditional schemes.

In addition, Fig. 6 demonstrates the NMSE performance against the SNR. The pilot length  $P$  is set to 256. For Fig. 6(a), the number of RIS elements is set to  $128 \times 1$  and the number of sub-arrays is set to  $4 \times 1$ . The SNR is increasing from -14 dB to 1 dB. It is illustrated that the accuracy of all schemes increases with the SNR. Similarly, the NMSE performance of polar-domain based schemes are better than the angular-domain based schemes since the user may be located in the near-field region of the RIS. The traditional OMP schemes has a better NMSE performance than row-structured sparsity based schemes at all considered SNR since the spatial non-stationarity in the system may make the row-structured sparsity support sets at sub-arrays differ from each other. The JLCR scheme can realize more accurate channel estimation than the above schemes since it considers the spatial non-stationarity effect on RIS. By decoupling the RIS into several spatial stationary sub-arrays, the proposed GD-OMP scheme has a better NMSE performance than all considered schemes. For Fig. 6(b), the number of RIS elements is set to  $64 \times 32$  and the number of sub-arrays is set to  $4 \times 2$ . The SNR is increasing from 3 dB to 18 dB. It can also be observed that higher SNR leads to a better NMSE performance. In addition, similar to the ULA case, the polar-domain based schemes can catch the spherical propagation characteristic in the near-field and outperforms the angular-domain based schemes, while the spatial non-stationarity effect ruins the row-structured sparsity to some extent and decreases the NMSE performance of the row-structured sparsity based schemes due to the spatial non-stationarity effect. By recognizing the VR based on the amplitude of the estimated channel, the JLCR scheme can improve the channel estimation accuracy. The proposed GD-OMP scheme, with the ability to capture the spatial non-stationarity effect in RIS systems, can realize a more accurate channel estimation at all considered SNR.

We also verify the scalability of the proposed GD-OMP scheme by increasing the path number in the system. As demonstrated in Fig. 7, the number of RIS elements is set to  $64 \times 32$ , the number of sub-arrays is set to  $4 \times 2$ , the SNR of the system is set to 10 dB, the path number is set to  $L_G = 8, L_r = 10$ ,

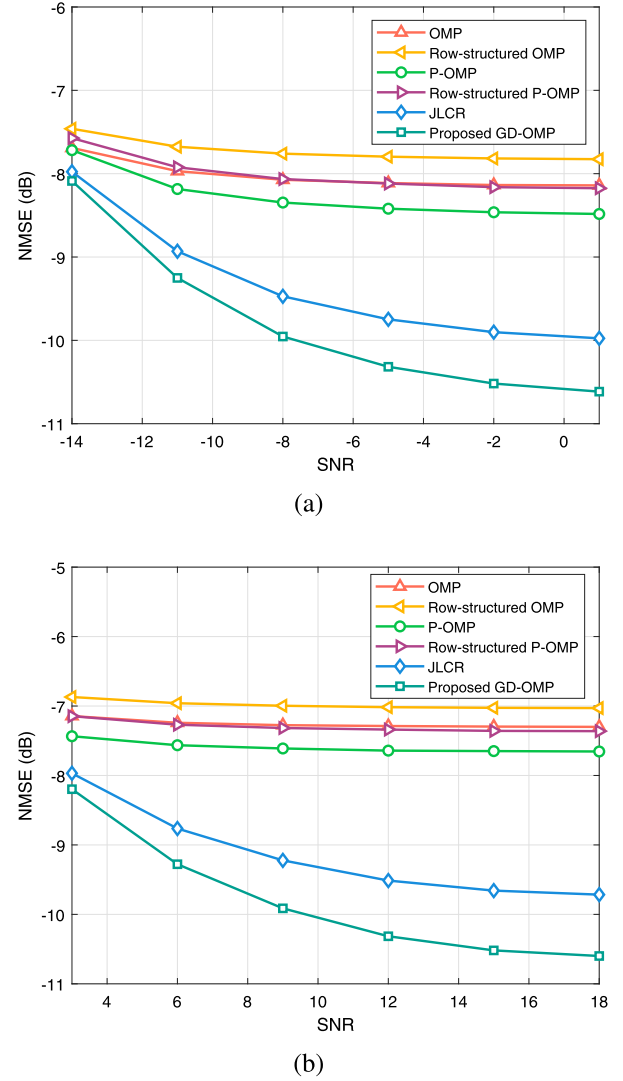


Fig. 6. The NMSE performance v.s. the SNR. The number of elements of RIS (a)  $128 \times 1$  uniform linear array (ULA), (b)  $64 \times 32$  uniform planar array (UPA).

and the pilot length is increasing from 32 to 256. The four schemes that do not consider the spatial non-stationarity effect cannot realize accurate channel estimation due to the mismatch between the spatial stationary assumption and the spatial non-stationary channel. For the JLCR scheme, larger path number makes it harder to recognize the VRs corresponding to different paths, so the NMSE performance only improves a little. For the proposed GD-OMP scheme, since it decouples the channel into several stationary sub-channels, it can still realize accurate channel estimation. The above results reveal that the proposed scheme can still work well in a system with a large number of paths. In practical systems, we can exploit the statistical CSI in a long period of time to determine the proper number of paths so as to realize accurate channel estimation.

To reveal the effectiveness of the proposed codebook, we compare the NMSE performance of the proposed dynamic codebook with the angular-domain DFT codebook and the polar-domain codebook in Fig. 8. Here, we assume the channel is spatial stationary in order to evaluate the performance of



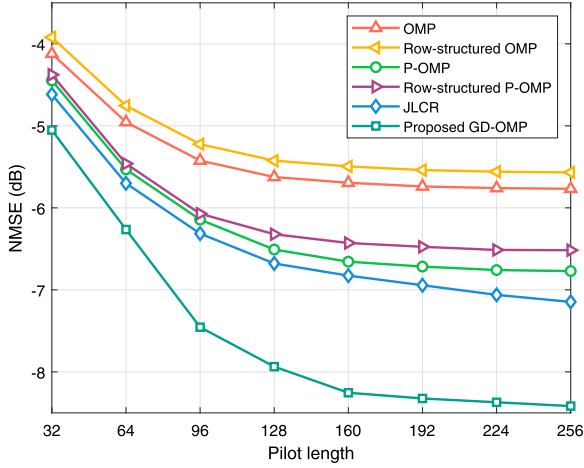


Fig. 7. The NMSE performance v.s. pilot length. The number of paths is  $L_G = 8$ ,  $L_r = 10$ .

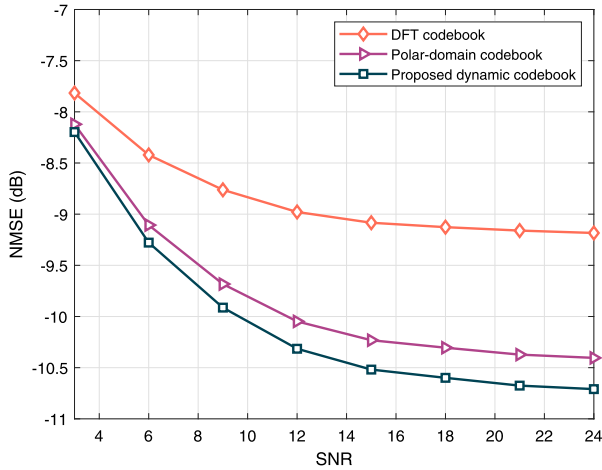


Fig. 8. The NMSE performance v.s. the SNR with different codebooks.

different codebooks. We can see that due to the existence of the near-field effect, the NMSE performance of the polar-domain codebook and the proposed codebook is better than that of the DFT codebook because these two codebooks take the distances into consideration. In addition, the proposed dynamic codebook can realize more accurate channel estimation since the distance is determined by calculating rather than simple sampling.

Furthermore, we evaluate the scalability of the proposed scheme by considering a more random VR. The number of RIS elements is set to  $64 \times 32$  and the SNR of the system is set to 10 dB. In the above scenarios, the VRs are generated according to the sub-arrays. By dividing the RIS into  $4 \times 2$  sub-arrays, the VRs corresponding to all paths are generated in the unit of  $16 \times 16$  element sub-array. However, in practical systems, the VRs may not follow the sub-array structure. In order to evaluate the performance of the proposed scheme in such scenarios, we set a minimum area of the VR, which is  $256 d^2$  with  $d$  being the antenna spacing. For each path, the specific size of the VR and the central location of the VR are both generated randomly.

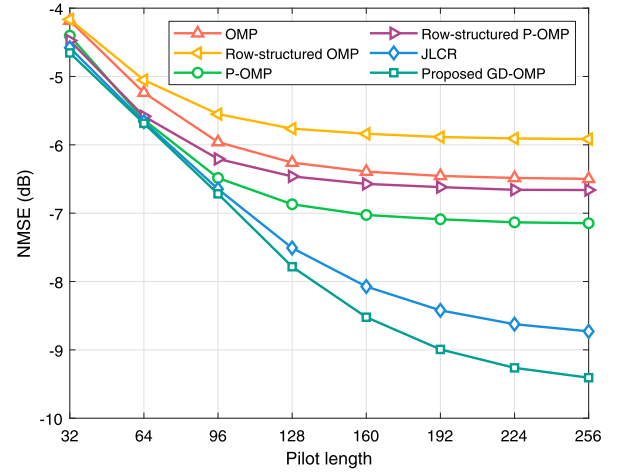


Fig. 9. The NMSE performance v.s. pilot length with random VRs.

The result is demonstrated in Fig. 9. Similar to Fig. 5, the schemes that do not consider the spatial non-stationarity effect cannot realize accurate channel estimation. The JLCR scheme can improve the channel estimation accuracy by recognizing the VRs according to the amplitude of the estimated channel. For our proposed scheme, by dividing the RIS into several sub-arrays, the negative influence of the spatial non-stationarity effect can be reduced and the channel estimation accuracy is further improved.

In order to evaluate the effect of the number of sub-arrays  $M_s$  on the channel estimation accuracy, we compare the NMSE performance of different sub-array configurations in Fig. 10. The number of RIS elements is set to  $64 \times 32$ , the SNR of the system is set to 10 dB and the pilot length ranges from 64 to 512. In Fig. 10(a), the RIS is divided into  $4 \times 2$  sub-arrays and VRs are generated based on these sub-arrays. We can observe that if the number of divided sub-arrays is smaller than the practical parameter (e.g.,  $M_s = 1 \times 1$ ,  $M_s = 1 \times 2$ ,  $M_s = 2 \times 2$ ), the NMSE performance is worse than that of the configuration  $M_s = 4 \times 2$ . In this case, the less sub-arrays we apply, the worse the NMSE performance is. On the contrary, when the number of sub-arrays is larger than the practical parameter (e.g.,  $M_s = 4 \times 4$ ), the NMSE performance will also decrease slightly since the effective number of pilots is smaller and the angular sampling in the codebook is more sparse. In Fig. 10(b), the VRs are generated randomly as in Fig. 9, where the minimum VR area is  $256 d^2$ . It is worth noting that when the RIS is divided into  $4 \times 2$  sub-arrays, the area of each sub-array is also  $256 d^2$ , which is in line with the minimum VR area. In this case, when the number of sub-arrays is smaller than  $4 \times 2$  (e.g.,  $M_s = 1 \times 1$ ,  $M_s = 1 \times 2$ ,  $M_s = 2 \times 2$ ), the NMSE performance is also worse than that of the configuration  $M_s = 4 \times 2$ . Nevertheless, when the number of sub-arrays is a little bit larger than the practical parameter, the NMSE performance is even better. This is because in random VR case, the VR may not locate exactly in a certain sub-array, so when the RIS array is divided more finely, the spatial non-stationary effect can be better distinguished. In practical systems, we should determine

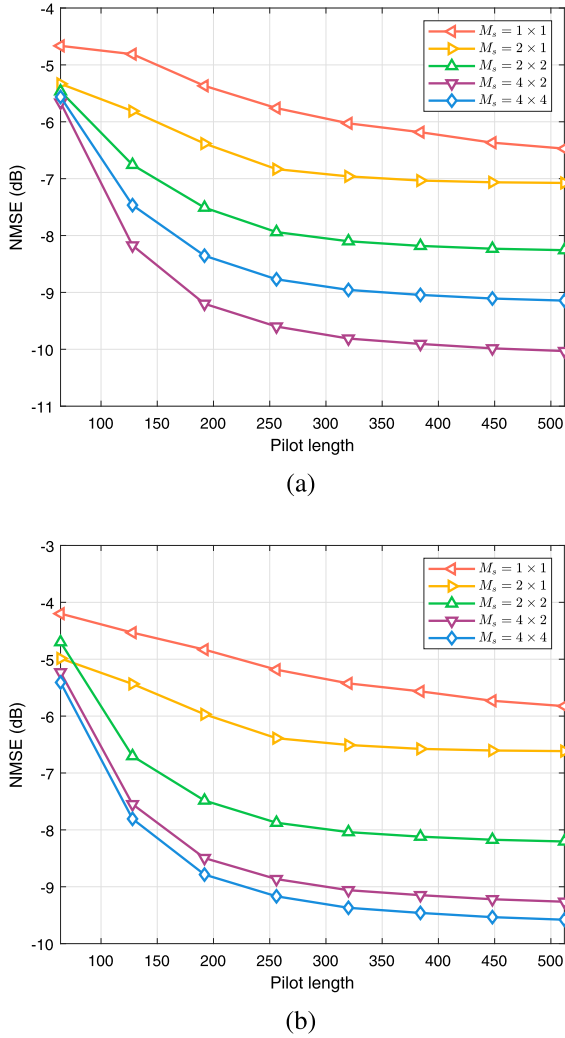


Fig. 10. The NMSE performance v.s. pilot length under different  $M_s$ . (a) sub-array based VR, (b) random VR.

the proper sub-array configuration according to the statistical CSI in a long period of time.

To sum up, under the spatial non-stationarity effect, traditional schemes suffer from a severe performance loss. In addition, the performance of some structured sparsity based schemes is even worse due to the mismatch with the spatial non-stationarity effect. For the JLCR scheme, although it considers the spatial non-stationarity effect, it cannot accurately recognize the VRs corresponding to different paths, resulting in an inevitable accuracy loss. By decoupling the spatial non-stationary RIS into several stationary sub-arrays, the proposed GD-OMP scheme is capable of capturing the spatial non-stationarity effect in RIS assisted communication systems and acquires a better NMSE performance than traditional schemes.

## VI. CONCLUSION

In this paper, we have investigated the channel estimation for non-stationary RIS assisted communication systems. To capture the spatial non-stationarity effect in RIS systems in the presence of the spatial ambiguity effect of RIS, we proposed

a two-dimensional group time block code (GTBC) based RIS decoupling scheme. By artificially create the relevance in the time-domain, the entire spatial non-stationary channel can be converted into a series of spatial stationary sub-channels and acquired the ability to recognize the spatial non-stationarity effect in the space domain. Based on the decoupled signals, we proposed a GTBC-based dynamic orthogonal matching pursuit (GD-OMP) scheme. By exploiting the geometric structure of adjacent sub-arrays, the proposed GD-OMP scheme could dynamically update the distance and sequentially estimate each path of the sub-channels. After estimating the sub-channels of all sub-arrays, the entire spatial non-stationary channel could then be accurately recovered. Analysis of the computational complexity of the proposed GD-OMP scheme was also carried out. Simulation results illustrated that the proposed GD-OMP scheme could capture the spatial non-stationarity effect in RIS systems and realize a more accurate channel estimation than existing schemes. Our proposed scheme provides a proper solution for accurate channel estimation for future RIS systems. For future works, one may consider to recognize the VR more accurately to improve the channel estimation accuracy.

## APPENDIX A

### PROOF OF LEMMA 1

The column coherence of  $\mathbf{D}$  can be presented as

$$\begin{aligned}
 & f(\phi_p, \phi_q, \theta_p, \theta_q) \\
 & \stackrel{(a)}{\approx} \frac{1}{M} \left| \sum_{\delta_{m_1 1}} \sum_{\delta_{m_2 2}} e^{jk_c \delta_{m_1 1} d(\sin \theta_p \sin \phi_p - \sin \theta_q \sin \phi_q)} \right. \\
 & \quad \cdot e^{jk_c \left( \frac{\delta_{m_1 1}^2 d^2 (\sin^2 \theta_q \sin^2 \phi_q - \sin^2 \theta_p \sin^2 \phi_p)}{2r} + \delta_{m_2 2} d(\cos \theta_p - \cos \theta_q) \right)} \\
 & \quad \cdot e^{jk_c \left( \frac{\delta_{m_2 2}^2 d^2 (\cos^2 \theta_q - \cos^2 \theta_p)}{2r} \right)} \left. \right| \\
 & \stackrel{(b)}{=} \frac{1}{M} \left| \sum_{m_1 = -\frac{M_1-1}{2}}^{\frac{M_1-1}{2}} e^{j(\mathcal{A}m_1^2 + \mathcal{B}m_1)} \sum_{m_2 = -\frac{M_2-1}{2}}^{\frac{M_2-1}{2}} e^{j(\mathcal{C}m_2^2 + \mathcal{D}m_2)} \right|, \quad (48)
 \end{aligned}$$

where  $k_c = 2\pi f_c/c$  is the wave number of the central frequency. (a) is obtained by applying (12) and Lemma 3 in [35], and (b) is obtained by substituting  $\delta_{m_1 1} = \frac{2m_1 - M_1 + 1}{2}$ ,  $\delta_{m_2 2} = \frac{2m_2 - M_2 + 1}{2}$ .  $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}$  are four constants satisfying  $\mathcal{A} = k_c \frac{d^2 (\sin^2 \theta_q \sin^2 \phi_q - \sin^2 \theta_p \sin^2 \phi_p)}{2r}$ ,  $\mathcal{B} = k_c d(\sin \theta_p \sin \phi_p - \sin \theta_q \sin \phi_q)$ ,  $\mathcal{C} = k_c \frac{d^2 (\cos^2 \theta_q - \cos^2 \theta_p)}{2r}$  and  $\mathcal{D} = k_c d(\cos \theta_p - \cos \theta_q)$ . By approximating the summation by integral, (48) can be further derived as

$$\begin{aligned}
 f(\phi_p, \phi_q, \theta_p, \theta_q) & \approx \frac{1}{M} \left| \int_{-\frac{M_1-1}{2}}^{\frac{M_1-1}{2}} e^{j(\mathcal{A}m_1^2 + \mathcal{B}m_1)} dm_1 \right. \\
 & \quad \left. \int_{-\frac{M_2-1}{2}}^{\frac{M_2-1}{2}} e^{j(\mathcal{C}m_2^2 + \mathcal{D}m_2)} dm_2 \right|. \quad (49)
 \end{aligned}$$

Considering the same structure of the two integrals, we only consider the first integral in the following derivation. Denoting  $m'_1$  as  $m_1 + \frac{\mathcal{B}}{2\mathcal{A}}$ , we can re-write the first integral as

$$\left| \int_{-\frac{M_1-1}{2}}^{\frac{M_1-1}{2}} e^{j(\mathcal{A}m_1^2 + \mathcal{B}m_1)} dm_1 \right| = \left| \int_{-\frac{M_1-1}{2} + \frac{\mathcal{B}}{4\mathcal{A}}}^{\frac{M_1-1}{2} + \frac{\mathcal{B}}{4\mathcal{A}}} e^{j\mathcal{A}m_1'^2} dm_1' \right|$$

$$\stackrel{(a)}{=} \sqrt{\frac{\pi}{2\mathcal{A}}} \left| \int_0^{\sqrt{\frac{2\mathcal{A}}{\pi} \left[ \frac{M_1-1}{2} + \frac{\mathcal{B}}{4\mathcal{A}} \right]}} e^{j\frac{\pi}{2}t^2} dt + \int_0^{\sqrt{\frac{2\mathcal{A}}{\pi} \left[ \frac{M_1-1}{2} - \frac{\mathcal{B}}{4\mathcal{A}} \right]}} e^{j\frac{\pi}{2}t^2} dt \right|, \quad (50)$$

where (a) is obtained by letting  $\mathcal{A}m_1'^2 = \frac{\pi}{2}t^2$ . Note that (50) can be divided into the summation of four Fresnel functions by Euler's formula [36], which can be further derived as

$$\left| \int_{-\frac{M_1-1}{2}}^{\frac{M_1-1}{2}} e^{j(\mathcal{A}m_1^2 + \mathcal{B}m_1)} dm_1 \right| = \left| \left( \frac{M_1-1}{2} + \frac{\mathcal{B}}{4\mathcal{A}} \right) G(\alpha_1) + \left( \frac{M_1-1}{2} - \frac{\mathcal{B}}{4\mathcal{A}} \right) G(\alpha_2) \right|, \quad (51)$$

with  $\alpha_1$  being  $\sqrt{\frac{2\mathcal{A}}{\pi} \left[ \frac{M_1-1}{2} + \frac{\mathcal{B}}{4\mathcal{A}} \right]}$  and  $\alpha_2$  being  $\sqrt{\frac{2\mathcal{A}}{\pi} \left[ \frac{M_1-1}{2} - \frac{\mathcal{B}}{4\mathcal{A}} \right]}$ .  $G(\alpha) = \frac{C(\alpha) + jS(\alpha)}{\alpha}$  is composed of two Fresnel functions, which can be elaborated as

$$\begin{cases} C(\alpha) = \int_0^\alpha \cos\left(\frac{\pi}{2}t^2\right) dt \\ S(\alpha) = \int_0^\alpha \sin\left(\frac{\pi}{2}t^2\right) dt \end{cases}. \quad (52)$$

Since  $\frac{\mathcal{B}}{4\mathcal{A}} = \frac{rf_c}{c(\sin\theta_q \sin\phi_q + \sin\theta_p \sin\phi_p)}$ , which is far larger than the number of antennas  $M_1$ , by letting

$$\alpha_1 \approx \alpha_2 \approx \alpha = \sqrt{\frac{2\mathcal{A}}{\pi} \frac{\mathcal{B}}{4\mathcal{A}}} = \sqrt{\frac{f_c r \sin\theta_q \sin\phi_q - \sin\theta_p \sin\phi_p}{2c \sin\theta_q \sin\phi_q + \sin\theta_p \sin\phi_p}}, \quad (53)$$

(51) can be further written as

$$\left| \int_{-\frac{M_1-1}{2}}^{\frac{M_1-1}{2}} e^{j(\mathcal{A}m_1^2 + \mathcal{B}m_1)} dm_1 \right| \approx |(M_1 - 1)G(\alpha)|. \quad (54)$$

Similarly, the second integral in (49) can be represented as

$$\left| \int_{-\frac{M_2-1}{2}}^{\frac{M_2-1}{2}} e^{j(\mathcal{C}m_2^2 + \mathcal{D}m_2)} dm_2 \right| \approx |(M_2 - 1)G(\beta)|. \quad (55)$$

where

$$\beta = \sqrt{\frac{2\mathcal{C}}{\pi} \frac{\mathcal{D}}{4\mathcal{C}}} = \sqrt{\frac{f_c r \cos\theta_q - \cos\theta_p}{2c \cos\theta_q + \cos\theta_p}}, \quad (56)$$

Finally, the column coherence of  $\mathbf{D}$  can be represented as

$$f(\phi_p, \phi_q, \theta_p, \theta_q) \approx \frac{1}{M} |(M_1 - 1)G(\alpha)(M_2 - 1)G(\beta)| \approx |G(\alpha)G(\beta)|, \quad (57)$$

which completes the proof.

## APPENDIX B PROOF OF LEMMA 2

According to the **Lemma 3** in [35], the last intersection term in (39) can be omitted. Then,  $|\mathbf{a}(\varphi, \vartheta)^H \mathbf{c}(\phi_{\ell_1}, \theta_{\ell_1}, r_{\ell_2}, \phi_{\ell_2}, \theta_{\ell_2})|$  has the similar form as (48), and can be represented as

$$|\mathbf{a}(\varphi, \vartheta)^H \mathbf{c}(\phi_{\ell_1}, \theta_{\ell_1}, r_{\ell_2}, \phi_{\ell_2}, \theta_{\ell_2})| \approx |G(\varrho)G(\varsigma)|, \quad (58)$$

where  $\varrho$  and  $\varsigma$  satisfy

$$\begin{cases} \varrho = \sqrt{\frac{f_c r_{\ell_2}}{2c} \frac{(\sin\varphi \sin\vartheta - \sin\varphi_c \sin\vartheta_c)^2}{1 - \sin^2\theta_{\ell_2} \sin^2\phi_{\ell_2}}} \\ \varsigma = \sqrt{\frac{f_c r_{\ell_2}}{2c} \frac{(\cos\vartheta - \cos\vartheta_c)^2}{1 - \cos^2\theta_{\ell_2}}} \end{cases}. \quad (59)$$

When  $\sin\varphi \sin\vartheta = \sin\varphi_c \sin\vartheta_c$  and  $\cos\vartheta = \cos\vartheta_c$ ,  $\varrho = \varsigma = 0$ , so  $|\mathbf{a}(\varphi, \vartheta)^H \mathbf{c}(\phi_{\ell_1}, \theta_{\ell_1}, r_{\ell_2}, \phi_{\ell_2}, \theta_{\ell_2})| \approx |G(0)G(0)| = 1$ . Otherwise, according to the sampling criterion of the azimuth angle and elevation angle,  $\varrho \geq \sqrt{\frac{2f_c r_{\ell_2} M_{s,1}^2}{cM_1^2}}$ , and  $\varsigma \geq \sqrt{\frac{2f_c r_{\ell_2} M_{s,2}^2}{cM_2^2}}$ , and thus we have  $|\mathbf{a}(\varphi, \vartheta)^H \mathbf{c}(\phi_{\ell_1}, \theta_{\ell_1}, r_{\ell_2}, \phi_{\ell_2}, \theta_{\ell_2})| \rightarrow 0$  in this case, which finally completes the proof.

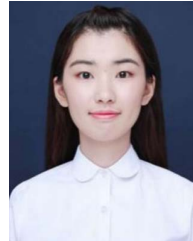
## REFERENCES

- [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. 116753–116773, 2019.
- [2] M. Di Renzo et al., "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2450–2525, Nov. 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, Jan. 2020.
- [4] 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.
- [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 Trans. Signal Process.*, vol. 68, pp. 5092–5106, 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] 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.
- [8] M. Cui, Z. Wu, Y. Chen, S. Xu, F. Yang, and L. Dai, "Low-power communications based on RIS and AI for 6G," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Jun. 2022, pp. 1–2.
- [9] 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, 2nd Quart. 2022.

- [10] Q.-U.-A. Nadeem, H. Alwazani, A. Kammoun, A. Chaaban, M. Debbah, 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.
- [11] 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.
- [12] 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.
- [13] Z. Wan, Z. Gao, and M.-S. Alouini, "Broadband channel estimation for intelligent reflecting surface aided mmWave massive MIMO systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Dublin, Ireland, Jul. 2020, pp. 1–6.
- [14] J. Wu, S. Kim, and B. Shim, "Near-field channel estimation for RIS-assisted wideband terahertz systems," in *Proc. IEEE Global Commun. Conf. (IEEE GLOBECOM)*, Dec. 2022, pp. 3893–3898.
- [15] M. Jian and Y. Zhao, "A modified off-grid SBL channel estimation and transmission strategy for RIS-assisted wireless communication systems," in *Proc. Int. Wireless Commun. Mobile Comput. (IWCMC)*, Limassol, Cyprus, Jun. 2020, pp. 1848–1853.
- [16] Z.-Q. He and X. Yuan, "Cascaded channel estimation for large intelligent metasurface assisted massive MIMO," *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 210–214, Feb. 2020.
- [17] J. Mirza and B. Ali, "Channel estimation method and phase shift design for reconfigurable intelligent surface assisted MIMO networks," *IEEE Trans. Cogn. Commun. Netw.*, vol. 7, no. 2, pp. 441–451, Jun. 2021.
- [18] J. Chen, Y.-C. Liang, H. V. Cheng, and W. Yu, "Channel estimation for reconfigurable intelligent surface aided multi-user MIMO systems," 2019, *arXiv:1912.03619*.
- [19] 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.
- [20] 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.
- [21] Z. Peng et al., "Channel estimation for RIS-aided multi-user mmWave systems with uniform planar arrays," *IEEE Trans. Commun.*, vol. 70, no. 12, pp. 8105–8122, Dec. 2022.
- [22] X. Gao, F. Tufvesson, and O. Edfors, "Massive MIMO channels—Measurements and models," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Nov. 2013, pp. 280–284.
- [23] Z. Yuan, J. Zhang, Y. Ji, G. F. Pedersen, and W. Fan, "Spatial non-stationary near-field channel modeling and validation for massive MIMO systems," *IEEE Trans. Antennas Propag.*, vol. 71, no. 1, pp. 921–933, Jan. 2023.
- [24] E. De Carvalho, A. Ali, A. Amiri, M. Angjelichinoski, and R. W. Heath, "Non-stationarities in extra-large-scale massive MIMO," *IEEE Wireless Commun.*, vol. 27, no. 4, pp. 74–80, Aug. 2020.
- [25] Y. Han, S. Jin, C.-K. 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.
- [26] Y. Zhu, H. Guo, and V. K. N. Lau, "Bayesian channel estimation in multi-user massive MIMO with extremely large antenna array," *IEEE Trans. Signal Process.*, vol. 69, pp. 5463–5478, 2021.
- [27] Y. Han, M. Li, S. Jin, C.-K. Wen, and X. Ma, "Deep learning-based FDD non-stationary massive MIMO downlink channel reconstruction," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 9, pp. 1980–1993, Sep. 2020.
- [28] Y. Han, S. Jin, C.-K. Wen, and T. Q. Quek, "Localization and channel reconstruction for extra large RIS-assisted massive MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 16, no. 5, pp. 1011–1025, May 2022.
- [29] Y. Chen and L. Dai, "Non-stationary channel estimation for extremely large-scale MIMO," *IEEE Trans. Wireless Commun.*, vol. 23, no. 7, pp. 7683–7697, Jul. 2024.
- [30] 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, Jun. 2021.
- [31] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
- [32] Z. Zhang et al., "Active RIS vs. passive RIS: Which will prevail in 6G?" *IEEE Trans. Commun.*, vol. 71, no. 3, pp. 1707–1725, Mar. 2023.
- [33] Y. Lu and L. Dai, "Near-field channel estimation in mixed LoS/NLoS environments for extremely large-scale MIMO systems," *IEEE Trans. Commun.*, vol. 71, no. 6, pp. 3694–3707, Jun. 2023.
- [34] 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.
- [35] Z. Wu and L. Dai, "Multiple access for near-field communications: SDMA or LDMA?" *IEEE J. Sel. Areas Commun.*, vol. 41, no. 6, pp. 1918–1935, Jun. 2023.
- [36] J. Sherman, "Properties of focused apertures in the Fresnel region," *IRE Trans. Antennas Propag.*, vol. 10, no. 4, pp. 399–408, Jul. 1962.



**Yuhao Chen** (Graduate Student Member, IEEE) received the B.E. degree in electronic engineering in 2021 from Tsinghua University, Beijing, China, where he is currently working toward the M.S. degree in electronic engineering. His research interests include massive MIMO, reconfigurable intelligent surface (RIS), and near-field communications. He received the IEEE ICC Outstanding Demo Award in 2022 and the National Scholarship in 2023.



**Mengnan Jian** received the M.S. degree from Tsinghua University, in 2019. Currently, she is a Master Researcher with the Wireless Research Institute, ZTE Corporation, responsible for long-term research and early standardization of candidate technologies for next generation wireless networks. Her research interests include ultra-massive MIMO, reconfigurable intelligent surface, electromagnetic information theory, and near-field communications. She has published more than 30 papers in the related fields.



**Linglong Dai** (Fellow, IEEE) received the B.S. degree from Zhejiang University, Hangzhou, China, in 2003, the M.S. (Hons.) degree from China Academy of Telecommunications Technology, Beijing, China, in 2006, and the Ph.D. (Hons.) degree from Tsinghua University, Beijing, China, in 2011. From 2011 to 2013, he was a Postdoctoral Research Fellow with the Department of Electronic Engineering, Tsinghua University, where he was an Assistant Professor, from 2013 to 2016, an Associate Professor, from 2016 to 2022, and has been a Professor since 2022. His research interests include massive MIMO, reconfigurable intelligent surface (RIS), millimeter-wave and Terahertz communications, near-field communications, machine learning for wireless communications, and electromagnetic information theory. He has co-authored the book titled *MmWave Massive MIMO: A Paradigm for 5G* (Academic Press, 2016). He has authored or co-authored over 100 IEEE journal papers and over 60 IEEE conference papers. He also holds over 20 granted patents. He has received five IEEE Best Paper Awards at the IEEE ICC 2013, the IEEE ICC 2014, the IEEE ICC 2017, the IEEE VTC 2017-Fall, the IEEE ICC 2018, and the IEEE GLOBECOM 2023. He has also received the Tsinghua University Outstanding Ph.D. Graduate Award in 2011, the Beijing Excellent Doctoral Dissertation Award in 2012, the China National Excellent Doctoral Dissertation Nomination Award in 2013, the URSI Young Scientist Award in 2014, the IEEE Transactions on Broadcasting Best Paper Award in 2015, the Electronics Letters Best Paper Award in 2016, the National Natural Science Foundation of China for Outstanding Young Scholars in 2017, the IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award in 2017, the IEEE ComSoc Asia-Pacific Outstanding Paper Award in 2018, the China Communications Best Paper Award in 2019, the IEEE Access Best Multimedia Award in 2020, the IEEE Communications Society Leonard G. Abraham Prize in 2020, the IEEE ComSoc Stephen O. Rice Prize in 2022, the IEEE ICC Outstanding Demo Award in 2022, and the National Science Foundation for Distinguished Young Scholars in 2023. He was listed as a highly cited researcher by Clarivate Analytics from 2020 to 2023.