# Hybrid Beamforming for Reconfigurable Intelligent Surface based Multi-User Communications: Achievable Rates With Limited Discrete Phase Shifts

Boya Di<sup>®</sup>, *Member, IEEE*, Hongliang Zhang<sup>®</sup>, *Member, IEEE*, Lingyang Song<sup>®</sup>, *Fellow, IEEE*, Yonghui Li<sup>®</sup>, *Fellow, IEEE*, Zhu Han<sup>®</sup>, *Fellow, IEEE*, and H. Vincent Poor, *Life Fellow, IEEE* 

Abstract-Reconfigurable intelligent surfaces (RISs) have drawn considerable attention from the research community recently. RISs create favorable propagation conditions by controlling the phase shifts of reflected waves at the surface, thereby enhancing wireless transmissions. In this paper, we study a downlink multi-user system where the transmission from a multiantenna base station (BS) to various users is achieved by an RIS reflecting the incident signals of the BS towards the users. Unlike most existing works, we consider the practical case where only a limited number of discrete phase shifts can be realized by a finite-sized RIS. A hybrid beamforming scheme is proposed and the sum-rate maximization problem is formulated. Specifically, continuous digital beamforming and discrete RIS-based analog beamforming are performed at the BS and the RIS, respectively, and an iterative algorithm is designed to solve this problem. Both theoretical analysis and numerical validations show that the RISbased system can achieve good sum-rate performance by setting a reasonable size of the RIS and a small number of discrete phase shifts.

Manuscript received October 1, 2019; revised January 15, 2020; accepted February 17, 2020. Date of publication June 8, 2020; date of current version August 20, 2020. This work was supported in part by the National Natural Science Foundation of China under Grant 61625101, Grant 61829101, Grant 61941101, in part by the U.S. Air Force Office of Scientific Research under MURI Grant 18RT0073, and in part by the U.S. National Science Foundation under Grant EARS-1839818, Grant CNS1717454, Grant CNS-1731424, Grant CNS-1702850, Grant CNS-1646607, and Grant CCF-1908308. The work of Yonghui Li was supported by ARC under Grant DP190101988. (*Corresponding author: Lingyang Song.*)

Boya Di is with the Department of Electronics Engineering, Peking University, Beijing 100871, China, and also with the Department of Computing, Imperial College London, London SW7 2AZ, U.K. (e-mail: diboya92@gmail.com).

Hongliang Zhang is with the Department of Electronics Engineering, Peking University, Beijing 100871, China, and also with the Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77004 USA (e-mail: hongliang.zhang92@gmail.com).

Lingyang Song is with the Department of Electronics Engineering, Peking University, Beijing 100871, China (e-mail: lingyang.song@pku.edu.cn).

Yonghui Li is with the School of Electrical and Information Engineering, The University of Sydney, Sydney, NSW 2006, Australia (e-mail: yonghui.li@sydney.edu.au).

Zhu Han is with the Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77004 USA, and also with the Department of Computer Science and Engineering, Kyung Hee University, Seoul 130-701, South Korea (e-mail: zhan2@uh.edu).

H. Vincent Poor is with the School of Engineering and Applied Science, Princeton University, Princeton, NY 08544 USA (e-mail: poor@princeton.edu).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSAC.2020.3000813

*Index Terms*— Reconfigurable intelligent surface, hybrid beamforming, multi-user communications, limited discrete phase shifts, non-convex optimization.

#### I. INTRODUCTION

THE past decade has witnessed an enormous increase in the number of mobile devices [1], triggering urgent needs for high-speed and seamless data services in future wireless systems. To meet such demands, one fundamental issue is how to improve the link quality in the complicated time-varying wireless environments involving unpredictable fading and strong shadowing effects. Various technologies have been developed such as relaying [2] and massive multiple input and multiple output (MIMO) systems [3], aiming to actively strengthen the target signals by forwarding and taking advantage of multi-path effects, respectively. However, these techniques require extra hardware implementation with inevitable power consumption and high complexity for signal processing, and the quality of service is also not always guaranteed in harsh propagation environments.

A more recently proposed technology involves reconfigurable intelligent surfaces (RISs), which are an ultrathin surfaces inlaid with multiple sub-wavelength scatters, i.e., RIS elements, whose electromagnetic response (such as phase shifts) can be controlled by simple programmable PIN diodes [6]. Based on the ON/OFF functions of PIN diodes, only a limited number of discrete phases shifts can be achieved by an RIS [7]. Instead of scattered waves emanated from traditional antennas, the sub-wavelength separation between adjacent RIS elements enables the refracted and reflected waves to be generated via superposition of incident waves at the surface [8]. Benefitting from such a programmable characteristic of molding the wavefronts into desired shapes, the RIS serves as a part of reconfigurable propagation environment such that the received signals are directly reflected towards the receivers without any extra cost of power sources or hardware [9], thereby improving the link quality and coverage.

To exploit the potential of RIS techniques, many existing works have considered the RIS as a reflection-type surface deployed between sources and destinations in either point-to-point communications [5], [10]–[12] or multi-user (MU) systems [13]–[17]. In [11], a point-to-point RIS-assisted

0733-8716 © 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information. multi-input single-out system has been investigated where the beamformer at the transmitter and continuous phase shifts of the RIS are jointly optimized to maximize the sum rate. In [15], a channel estimation protocol has been proposed for a multi-user RIS-assisted system and the continuous phase shifts have been designed to maximize the minimum user date. In [16], the authors have minimized the transmit power of the access point by optimizing the continuous digital beamforming and discrete phase shifts. An algorithm has been designed for the single-user case and extended to the multiuser case. In [17], the authors have considered a downlink RISbased multi-user MISO network and studied the joint active and passive beamforming problem to maximize the weighted sum rate. An iterative algorithm has been designed for the continuous phase shift case, which can also be extended to the discrete case.

However, two major issues still remain to be further discussed in the open literature.

- For the multi-user case, how can the limited discrete phase shifts be determined directly such that the interuser interference can be eliminated? How does the quantization level influence the sum rate of the system?
- Considering the strengthened coupling between propagation and discrete phase shifts brought by the dominant reflection ray via the RIS, how do we design the size of RIS and perform beamforming in a multi-antenna system to achieve the maximum sum rate?

To address the above issues, in this paper, we consider a downlink multi-user multi-antenna system where the direct links between the multi-antenna base station (BS) and users suffer from deep shadowing. To provide high-quality data services, an RIS with limited discrete phase shifts is deployed between the BS and users such that the signals sent by the BS are reflected by the RIS towards the users. To achieve better directional reflection rays, it is vitally important to determine the phase shifts of all RIS elements, the process of which is also known as RIS configuration. Such a builtin programmable configuration [5] is actually equivalent to analog beamforming, realized by the RIS inherently. Since the RIS elements do not have any digital processing capability, we consider hybrid beamforming (HBF) [19] consisting of digital beamforming at the BS and RIS configuration based analog beamforming. A novel HBF scheme for RIS-based communications with discrete phase shifts is thus required for better shaping the propagation environment and sum rate maximization.

Designing an HBF scheme presents several major challenges. *First*, the reflection-dominated propagation [20] and the RIS configuration based analog beamforming are coupled, rendering the optimal scheme very hard to obtain. The traditional beamforming schemes with separate channel matrix and analog beamformer do not work any more. *Second*, requiring the RIS to have discrete phase shifts renders the sum rate maximization as a mixed integer programming problem, which is non-trivial to solve especially in the complex domain. *Third*, given the dense placing of the RIS elements, the correlation between elements may degrade the data rate performance. Thus, it is necessary to explore how the achievable rate is influenced by the size of the RIS, which is challenging due to the complicated propagation environments especially in a multi-user case.

Through solving the above challenges, we aim to design an HBF scheme for the RIS-based multi-user system with limited discrete phase shifts to maximize the sum rate. Our main contributions can be summarized as follows.

- We consider a downlink RIS-based multi-user system where an RIS with limited discrete phase shifts reflects signals from the BS towards various users. We design an HBF scheme where digital beamforming is performed at the BS, and analog beamforming is conducted at the RIS.
- A mixed-integer sum rate maximization problem for RIS-based HBF is formulated and decomposed into two subproblems. We propose an iterative algorithm in which the digital beamforming subproblem is solved by zeroforcing (ZF) beamforming with power allocation and the RIS-based analog beamforming is solved by the outer approximation.
- We prove that the proposed RIS-based HBF scheme can save as many as half of the radio frequency (RF) chains compared to traditional HBF schemes. Extending from our theoretical analysis on the pure Line-of-Sight (LoS) case, we reveal the influence of the size of the RIS and the number of discrete phase shifts on the sum rate both theoretically and numerically.

The rest of this paper is organized as follows. In Section II, we introduce the system model of the downlink RIS-based MU multi-antenna system. The frequency-response model of the RIS and the channel model are derived. In Section III, the HBF scheme for the RIS-based system is proposed. A sum rate maximization problem is formulated and decomposed into two subproblems: digital beamforming and RIS configuration based analog beamforming. An iterative algorithm is developed in Section IV to solve the above two subproblems and a sub-optimal solution is obtained. In Section V, we compare the RIS-based HBF scheme with the traditional one theoretically, and discuss how to achieve the maximum sum rate in the pure LoS case. The complexity and convergence of the proposed algorithm are also analyzed. Numerical results in Section VI evaluate the performance of the proposed algorithm and validate our analysis. Finally, conclusions are drawn in Section VII.

Notation: Scalars are denoted by italic letters, vectors and matrices are denoted by bold-face lower-case and uppercase letters, respectively. For a complex-valued vector x, ||x|| denotes its Euclidean norm, and diag(x) denotes a diagonal matrix whose diagonal element is the corresponding element in x. For a square matrix S, Tr(S) denotes its trace. For any general matrix M,  $M^H$  and  $M^T$  denote its conjugate transpose and transpose, respectively. I, 0, and 1 denote an identity matrix, all-zero and all-one vectors, respectively.

# II. SYSTEM MODEL

In this section, we first introduce the RIS-based downlink multi-user multi-antenna system in which a BS with multiple



Fig. 1. System model of the RIS-based downlink multi-user communication system.

antennas serves various single-antenna users via the RIS such that the propagation environment can be pre-designed and configured to optimize the system performance. The discrete phase shift model of the RIS and the channel model are then constructed, respectively.

## A. Scenario Description

Consider a downlink multi-user communication system as shown in Fig. 1 where a BS equipped with  $N_t$  antennas transmits to K single-antenna users. Due to the complicated and dynamic wireless environment involving unexpected fading and potential obstacles, the BS - user link may not be stable enough or even be in outage. To alleviate this issue, we consider to deploy an RIS between the BS and users, which reflects the signals from the BS and directly projects to the users by actively shaping the propagation environment into a desirable form.

An RIS consists of  $N_R \times N_R$  electrically controlled RIS elements as shown in Fig. 1, each of which is a sub-wavelength meta-material particle with very small features. An RIS controller, equipped with several PIN diodes, can control the ON/OFF state of the connection between adjacent metal plates where the RIS elements are laid, thereby manipulating the electromagnetic response of the RIS elements towards incident waves. Due to these built-in programmable elements, the RIS requires no extra active power sources nor do not have any signal processing capability such as decoding [5]. In other words, it serves as a low-cost reconfigurable phased array that only relies on the combination of multiple programmable radiating elements to realize a desired transformation on the transmitted, received, or reflected waves [9].

# B. Reconfigurable Intelligent Surface With Limited Discrete Phase Shifts

As shown in Fig. 2, the RIS is achieved by the *b*-bit re-programmable meta-material, which has been implemented as a set of radiative elements layered on a guiding structure following the wave guide techniques, forming a 2-dimensional (2D) planar antenna array [7]. Being a miniature radiative element, the field radiated from the RIS element has a *phase* and *amplitude* determined by this element's polarizability, which



Fig. 2. Schematic structure of *b*-bit encoded RIS.

can be tuned by the RIS controller via multiple PIN diodes (ON/OFF) [5]. However, the phase and amplitude introduced by an RIS element are not generated randomly; instead they are constrained by the Lorentzian resonance response [4], which greatly limits the range of phase values. Based on such constraints, one common manner of implementation is to constrain the amplitude and sample the phase values from the finite feasible set [4] such that the voltage-controlled diodes can easily manipulate a *discrete* set of phase values at a very low cost.

Specifically, we assume that each RIS element is encoded by the controller (e.g., via PIN diodes) to conduct  $2^b$  possible phase shifts to reflect the radio wave. Due to the frequencyselective nature of the meta-materials, these elements only vibrate in resonance with the incoming waves over a narrow band centering at the resonance frequency. Without loss of generality, we denote the frequency response of each element  $(l_1, l_2)$  at the  $l_1$ -th row and  $l_2$ -th column of the 2D RIS within the considered frequency range as  $q_{l_1,l_2}$ ,  $0 \le l_1, l_2 \le N_R - 1$ . Since the RIS is *b*-bit controllable,  $2^b$  possible configuration modes (i.e., phases) of each  $q_{l_1,l_2}$  can be defined according to the Lorentzian resonance response [4].

$$q_{l_1,l_2} = \frac{j + e^{j\theta_{l_1,l_2}}}{2}, \theta_{l_1,l_2} = \frac{m_{l_1,l_2}\pi}{2^{b-1}},$$
  
$$m_{l_1,l_2} \in \left\{0, 1, \dots, 2^b - 1\right\}, 0 \le l_1, l_2 \le N_R - 1, \quad (1)$$

where  $\theta_{l_1,l_2}$  denotes the phase shift of RIS element  $(l_1, l_2)$ . For convenience, we refer to b as the number of *quantization bits*.

## C. Reflection-Dominated Channel Model

In this subsection, we model the channel between antenna  $0 \le n \le N_t - 1$  of the BS and user k. Benefited from the directional reflections of the RIS, the BS - RIS - user link is usually stronger than other multipaths as well as the degraded direct link between the BS and the user [18]. Therefore, we model the channel between the BS and each user k as a Rician model such that the BS - RIS - user link acts as the dominant "LoS" component and all the other paths

together form the "non-LoS (NLoS)" component. We denote the channel between BS antenna  $1 \le n \le N_t$  and user k via RIS element  $(l_1, l_2)$  as

$$\tilde{h}_{l_1,l_2}^{(k,n)} = \sqrt{\frac{\kappa}{1+\kappa}} h_{l_1,l_2}^{(k,n)} + \sqrt{\frac{1}{1+\kappa}} h_{NLOS,(l_1,l_2)}^{(k,n)}, \quad (2)$$

where  $\kappa$  is the Rician factor,  $h_{l_1,l_2}^{(k,n)}$  is the "LoS" component, and  $h_{NLOS,(l_1,l_2)}^{(k,n)}$  is the "NLoS" component.

Let  $D_{l_1,l_2}^{(n)}$  and  $d_{l_1,l_2}^{(k)}$  denote the distance between antenna n and RIS element  $(l_1, l_2)$ , and that between user k and RIS element  $(l_1, l_2)$ , respectively. The channel  $h_{l_1,l_2}^{(k,n)}$  between each BS antenna  $1 \le n \le N_t$  and user k via RIS element  $(l_1, l_2)$  is given by [20]

$$h_{l_1,l_2}^{(k,n)} = \left( D_{l_1,l_2}^{(n)} \cdot d_{l_1,l_2}^{(k)} \right)^{-\alpha} \cdot e^{-j\frac{2\pi}{\lambda} \left( D_{l_1,l_2}^{(n)} + d_{l_1,l_2}^{(k)} \right)}$$
(3a)

$$= \left[ \left( D_{l_{1},l_{2}}^{(n)} \right)^{-\alpha} \cdot e^{-j\frac{2\pi}{\lambda}D_{l_{1},l_{2}}^{(n)}} \right] \\ \cdot \left[ \left( d_{l_{1},l_{2}}^{(k)} \right)^{-\alpha} \cdot e^{-j\frac{2\pi}{\lambda}d_{l_{1},l_{2}}^{(k)}} \right]$$
(3b)

$$=h_{l_1,l_2}^{(n)} \cdot h_{l_1,l_2}^{(k)}$$
(3c)

where  $\alpha$  is the path loss parameter. We use  $h_{l_1,l_2}^{(n)}$  and  $h_{l_1,l_2}^{(k)}$  to denote the channel between BS antenna n and RIS element  $(l_1, l_2)$ , and that between RIS element  $(l_1, l_2)$  and user k, respectively. Similarly, we have

$$h_{NLOS,(l_1,l_2)}^{(k,n)} = PL\left(D_{l_1,l_2}^{(n)}, d_{l_1,l_2}^{(k)}\right)\hat{h}_{l_1,l_2}^{(k,n)}$$
(4)

where  $PL(\cdot)$  is the path loss model for NLoS transmissions, and  $\hat{h}_{l_1,l_2}^{(k,n)} \sim \mathcal{CN}(0,1)$  denotes the small-scale NLoS component. Here we assume that the perfect channel state information is known to the BS via communicating with the RIS controller over a dedicated wireless link. A number of channel estimation methods can be found in [15], [16], which is out of the scope of this paper. For the case where channel information is partially known to the BS, i.e., only large-scale fading <sup>1</sup> is available, we will consider the pure LoS transmission and discuss it in detail in Section V.

Geometric Model of the Propagation: We now derive the above distances  $D_{l_1,l_2}^{(n)}$  and  $d_{l_1,l_2}^{(k)}$  by using the geometry information [21], [22]. Without loss of generality, we assume that the uniform planar arrays (UPA) and uniform linear arrays (ULA) are deployed at the RIS and the BS, respectively. The UPA is placed in a way that the surface is perpendicular to the ground. To describe the geometry of the above mentioned UPA and ULA, we employ the spherical coordinates as shown in Fig. 3. For the RIS, we define the local origin as the higher corner of the surface; while for the BS, its local origin is set as one end of the linear array. The y axis is set along the direction of BS antenna 0 - element (0,0) link, and the z axis is perpendicular to the ground. For convenience, the directions



Fig. 3. Placement of the antenna arrays at the BS and the RIS.

of the ULA and RIS on the *x*-*y* plane can be depicted by the angles  $\theta_B \in [0, 2\pi]$  and  $\theta_R \in [0, 2\pi]$ , respectively, as shown in Fig. 3. We can then define the principle directions of the linear antenna arrays at the BS and the RIS, denoted by  $\mathbf{n}_B$  and  $\left\{\mathbf{n}_R^{(1)}, \mathbf{n}_R^{(2)}\right\}$ , as

$$\mathbf{n}_B = \mathbf{n}_x \cos \theta_B + \mathbf{n}_y \sin \theta_B, \tag{5a}$$

$$\mathbf{n}_{R}^{(1)} = \mathbf{n}_{x}\cos\theta_{R} + \mathbf{n}_{y}\sin\theta_{R}, \tag{5b}$$

$$\mathbf{n}_R^{(2)} = \mathbf{n}_z,\tag{5c}$$

where  $\mathbf{n}_x$ ,  $\mathbf{n}_y$ , and  $\mathbf{n}_z$  are directions of the x, y, and z axis, respectively.

We denote the uniform separation between any two adjacent elements in the above defined ULA and UPA as  $d_B$  and  $\left\{d_R^{(1)}, d_R^{(2)}\right\}$ , respectively. The position of any antenna n at the BS can be represented by

$$\mathbf{c}_B^{(n)} = n d_B \mathbf{n}_B,\tag{6}$$

and the position of any RIS element  $(l_1, l_2)$  can be given by

$$\mathbf{c}_{R}^{(l_{1},l_{2})} = l_{1}d_{R}^{(1)}\mathbf{n}_{R}^{(1)} + l_{2}d_{R}^{(2)}\mathbf{n}_{R}^{(2)} + D_{0,0}^{(0)}\mathbf{n}_{y}.$$
 (7)

Based on (5) - (7), the distance between antenna n and RIS element  $(l_1, l_2)$ , i.e.,  $D_{l_1, l_2}^{(n)}$ , can be calculated as, (8a)–(8c) shown at the bottom of the next page, where (8c), as shown at the bottom of the next page, is obtained by adopting  $\sqrt{1 + a} \approx 1 + a/2$  when  $a \ll 1$ . Denote the position of each user k as  $\mathbf{c}_k = (d_{x,k}, d_{y,k}, d_{z,k})$ . The distance between the RIS element and any user k can be expressed by

$$d_{l_1,l_2}^{(k)} = \left\| \mathbf{c}_R^{(l_1,l_2)} - \mathbf{c}_k \right\|_2.$$
(9)

Since the distance between any two antennas or RIS elements is much smaller than the distance between the BS and the user, i.e.,  $\min \left\{ D_{l_1,l_2}^{(n)}, d_{l_1,l_2}^{(k)} \right\} \gg \max \left\{ d_R^{(1)}, d_R^{(2)}, d_B \right\}$ , we assume that the path loss of each BS-user link via different RIS elements and BS antennas is the same. Therefore, the channel propagation  $h_{l_1,l_2}^{(n)}$  and  $h_{l_1,l_2}^{(k)}$  can be rewritten as

$$h_{l_1,l_2}^{(n)} = \left(D_{0,0}^{(0)}\right)^{-\alpha} \cdot e^{-j\frac{2\pi}{\lambda}D_{l_1,l_2}^{(n)}},$$
 (10a)

$$h_{l_1,l_2}^{(k)} = \left(d_{0,0}^{(k)}\right)^{-\alpha} \cdot e^{-j\frac{2\pi}{\lambda}d_{l_1,l_2}^{(k)}},\tag{10b}$$

where  $D_{l_1,l_2}^{(n)}$  and  $d_{l_1,l_2}^{(k)}$  are given in (8) and (9). Based on the propagation characteristics introduced above,

Based on the propagation characteristics introduced above, we will investigate how RIS can be utilized to assist multi-user transmissions in the following section.

<sup>&</sup>lt;sup>1</sup>The large-scale fading information is related to positions of users (assuming that the position of the RIS is known to the BS). Various estimation techniques can be utilized such as GPS and multi-BS triangular positioning. For example, the BS can utilize the received pilot signals from the users to locate them based on the signal strength, the power of which is usually larger than that of data transmissions.



Fig. 4. Block diagram of the RIS-based transmission between the BS and user k.

# III. RIS-BASED HYBRID BEAMFORMING AND PROBLEM FORMULATION FOR MULTI-USER COMMUNICATIONS

Note that the RIS usually consists of a large number of RIS elements, which can be viewed as antenna elements fay away from the BS, inherently capable of realizing analog beamforming via RIS configuration. However, these RIS elements do not have any digital processing capacity, requiring signal processing to be carried out at the BS.

In this section, to realize reflected waves towards preferable directions, we present an HBF scheme for RIS-based multiuser communications given the phase shift model and the channel model of RIS-based transmissions in Section II. As shown in Fig. 4, the digital beamforming is performed at the BS while the analog beamforming is achieved by the RIS with discrete phase shifts. Based on the considered HBF scheme, we formulate a sum rate maximization problem, and then decompose it into the digital beamforming subproblem and the RIS configuration based analog beamforming subproblem.

#### A. Hybrid Beamforming Scheme

1) Digital Beamforming at the BS: The BS first encodes K different data steams via a digital beamformer,  $\mathbf{V}_D$ , of size  $N_t \times K$ , satisfying  $N_t \geq K$ . After up-converting the encoded signal over the carrier frequency and allocating the transmit powers, the BS sends users' signals directly through  $N_t$  antennas. Denote the intended signal vector for K users as  $s \in \mathbb{C}^{K \times 1}$ . The transmitted signals of the BS can be given by

$$\boldsymbol{x} = \mathbf{V}_D \boldsymbol{s}. \tag{11}$$

2) *RIS Configuration Based Analog Beamforming:* After travelling through the reflection-dominated channel introduced

in Section II-C, the received signal at the antenna of user k can then be expressed as

$$z_{k} = \sum_{n} \sum_{l_{1},l_{2}} \phi_{l_{1},l_{2}}^{(k,n)} h_{l_{1},l_{2}}^{(k,n)} q_{l_{1},l_{2}} V_{D_{k,n}} s_{k} + \sum_{k \neq k'} \sum_{n} \sum_{l_{1},l_{2}} \phi_{l_{1},l_{2}}^{(k,n)} h_{l_{1},l_{2}}^{(k,n)} q_{l_{1},l_{2}} V_{D_{k',n}} s_{k'} + w_{k},$$
(12)

where  $w_k \sim C\mathcal{N}(0, \sigma^2)$  is the additive white Gaussian noise and  $V_{D_{k,n}}$  denotes the k-th element in row n of matrix  $\mathbf{V}_D$ . In (12),  $\phi_{l_1,l_2}^{(k,n)}$  denotes the reflection co-efficient of the RIS element  $(l_1, l_2)$  with respect to the transmitting antenna n and user k. In practice, it is a function of the incidence and reflection angels, but here without loss of generality we assume that  $\phi_{l_1,l_2}^{(k,n)} = \phi^{(k)}, \forall n, l_1, l_2$ . We ignore the coupling between any two RIS elements here for simplicity, and thus the received signal of each user k comes from the accumulated radiations of all RIS elements, as shown in (12). This is a common assumption widely used in the literature on both meta-surfaces [4] and traditional antenna arrays [23].

3) Received Signal at the User: For each user k, after it receives the signal  $z_k$ , it down converts the signal to the baseband and then recovers the final signal. The whole transmission model of K users can be formulated by

$$\boldsymbol{z} = \mathbf{F} \mathbf{V}_D \boldsymbol{s} + \boldsymbol{w},\tag{13}$$

where  $\mathbf{w} = [w_1, \cdots, w_K]^T$  is the noise vector. The transmission matrix  $\mathbf{F}$  in (13) is of size  $K \times N_t$ , in which each element  $f_{k,n}$  is defined as

$$f_{k,n} = \operatorname{Tr}\left(\phi^{(k)} \mathbf{Q}^T \mathbf{H}^{(k,n)}\right),\tag{14}$$

where **Q** is a  $N_R \times N_R$  matrix consisting of the phase shifts  $\{q_{l_1,l_2}\}$ , and  $\mathbf{H}^{(k,n)}$  is the  $N_R \times N_R$  channel matrix between each BS antenna n and user k, consisting of elements  $\{h_{l_1,l_2}^{(k,n)}\}$ .

$$D_{l_{1},l_{2}}^{(n)} = \left\| \mathbf{c}_{R}^{(l_{1},l_{2})} - \mathbf{c}_{B}^{(n)} \right\|_{2}$$

$$= \left[ \left( l_{1}d_{R}^{(1)}\cos\theta_{R} - nd_{B}\cos\theta_{B} \right)^{2} + \left( l_{1}d_{R}^{(1)}\sin\theta_{R} + D_{0,0}^{(0)} - nd_{B}\sin\theta_{B} \right)^{2} + \left( l_{2}d_{R}^{(2)} \right)^{2} \right]^{\frac{1}{2}}$$
(8a)
$$\left( l_{1}d_{R}^{(1)}\cos\theta_{R} - nd_{B}\cos\theta_{B} \right)^{2} + \left( l_{2}d_{R}^{(2)} \right)^{2} \right]^{\frac{1}{2}}$$
(8b)

$$\approx \left( l_1 d_R^{(1)} \sin \theta_R + D_{0,0}^{(0)} - n d_B \sin \theta_B \right) + \frac{\left( l_1 u_R^2 \cos \theta_R - n u_B \cos \theta_B \right) + \left( l_2 u_R^2 \right)}{2 D_{0,0}^0}, \tag{8c}$$

#### B. Sum Rate Maximization Problem Formulation

To explore how the HBF design influences the sum-rate performance, we evaluate the achievable data rates of all users in the RIS-based system. Based on (12) and (13), we first rewrite the received signal of user k in matrix form as

$$z_{k} = \mathbf{F}_{k}^{H} \mathbf{V}_{D,k} s_{k} + \underbrace{\sum_{k' \neq k} \mathbf{F}_{k}^{H} \mathbf{V}_{D,k'} s_{k'}}_{\text{inter-user interference}} + w_{k}, \quad (15)$$

where  $\mathbf{F}_k$  and  $\mathbf{V}_{D,k}$  denote the k-th columns of matrices  $\mathbf{F}$  and  $\mathbf{V}_D$ , respectively. The achievable rate of user k can then be given by

$$R_{k} = \log_{2} \left( 1 + \frac{\left| \mathbf{F}_{k}^{H} \mathbf{V}_{D,k} \right|^{2}}{\sum_{k' \neq k} \left| \mathbf{F}_{k}^{H} \mathbf{V}_{D,k'} \right|^{2} + \sigma^{2}} \right).$$
(16)

We aim to maximize the achievable rates of all users by optimizing the digital beamformer  $V_D$  and the RIS configuration  $\{q_{l_1,l_2}\}$ , as formulated below:

$$\underset{\mathbf{V}_{D}, \{q_{l_{1}, l_{2}}\}}{\operatorname{maximize}} \sum_{1 \le k \le K} R_{k}$$
(17a)

subject to  $\operatorname{Tr}\left(\mathbf{V}_{D}^{H}\mathbf{V}_{D}\right) \leq P_{T},$  (17b)

$$q_{l_1,l_2} = \frac{j + e^{j\theta_{l_1,l_2}}}{2}, 0 \le l_1, l_2 \le N_R - 1,$$

$$(17c)$$

$$\theta_{l_1,l_2} = \frac{m_{l_1,l_2}\pi}{2^{b-1}}, m_{l_1,l_2} \in \left\{0, 1, \dots, 2^b - 1\right\},$$

$$(17d)$$

where  $P_T$  is the total transmit power of the BS.

#### C. Problem Decomposition

Note that problem (17) is a mixed integer non-convex optimization problem which is very challenging due to the large number of discrete variables  $\{q_{l_1,l_2}\}$  as well as the coupling between propagation and RIS configuration based analog beamforming. Traditional analog beamforming design methods with finite resolution phase shifters [19] may not fit well since it is non-trivial to decouple the transmission matrix **F** into the product of a channel matrix and a beamformer matrix in our case. To solve this problem efficiently, we decouple it into two subproblems as shown below.

1) Digital Beamforming: Given RIS configuration  $\{q_{l_1,l_2}\}$ , the digital beamforming subproblem can be written by

$$\underset{\mathbf{V}_{D}}{\operatorname{maximize}} \quad \sum_{1 \le k \le K} R_{k},$$
 (18a)

subject to 
$$\operatorname{Tr}\left(\mathbf{V}_{D}^{H}\mathbf{V}_{D}\right) \leq P_{T},$$
 (18b)

where F is fixed.

2) RIS Configuration Based Analog Beamforming: Based on constraint (17c), the RIS configuration subproblem with fixed beamformer  $V_D$  is equivalent to

$$\underset{\{\theta_{l_1, l_2}\}}{\text{maximize}} \quad \sum_{1 \le k \le K} R_k, \tag{19a}$$

subject to  $\theta_{l_1,l_2} = \frac{m_{l_1,l_2}\pi}{2^{b-1}}, m_{l_1,l_2} \in \{0, 1, \dots, 2^b - 1\}.$ (19b)

In the next section, we will design two algorithms to solve these subproblems, respectively.

#### IV. SUM RATE MAXIMIZATION ALGORITHM DESIGN

In this section, we will develop a sum rate maximization (SRM) algorithm to obtain a suboptimal solution of problem (17) in Section III. Specially, we iteratively solve subproblem (18) given RIS configuration  $\{q_{l_1,l_2}\}$ , and solve subproblem (19) given beamformer  $V_D$ . Finally, we will summarize the overall algorithm and provide convergence and complexity analysis.

## A. Digital Beamforming Algorithm

Subproblem (18) is a well-known digital beamforming problem. According to the results in [24], the ZF digital beamformer can obtain a near optimal solution. Therefore, we consider ZF beamforming together with power allocation as the beamformer at the BS to alleviate the interference among users. <sup>2</sup> Based on the results in [25], the beamformer can be given by

$$V_D = F^H (FF^H)^{-1} P^{\frac{1}{2}} = \tilde{V}_D P^{\frac{1}{2}}, \qquad (20)$$

where  $\tilde{V}_D = F^H (FF^H)^{-1}$  and P is a diagonal matrix whose k-th diagonal element is the received power at the k-th user, i.e.,  $p_k$ .

In the ZF beamforming, we have the following constraints:

$$|\mathbf{F}_{k}^{H}(\mathbf{V}_{D})_{k}| = \sqrt{p_{k}},$$
  
$$|\mathbf{F}_{k}^{H}(\mathbf{V}_{D})_{k'}| = 0, \forall k' \neq k.$$
 (21)

With these constraints, subproblem (18) can be reduced to the following power allocation problem:

$$\max_{\{p_k \ge 0\}} \quad \sum_{1 \le k \le K} \log_2\left(1 + \frac{p_k}{\sigma^2}\right), \tag{22a}$$

subject to 
$$\operatorname{Tr}\left(\boldsymbol{P}^{\frac{1}{2}}\tilde{\boldsymbol{V}}_{D}^{H}\tilde{\boldsymbol{V}}_{D}\boldsymbol{P}^{\frac{1}{2}}\right) \leq P_{T}.$$
 (22b)

The optimal solution of this problem can be obtained by water-filling [26] as

$$p_k = \frac{1}{\nu_k} \max\left\{\frac{1}{\mu} - \nu_k \sigma^2, 0\right\},$$
 (23)

where  $\nu_k$  is the *k*-th diagonal element of  $\tilde{V}_D^H \tilde{V}_D$  and  $\mu$  is a normalized factor which is selected such that  $\sum_{1 \le k \le K} \max\{\frac{1}{\mu} - \nu_k \sigma^2, 0\} = P_T$ . The algorithm can be summarized in Algorithm 1.

<sup>2</sup>Here we consider ZF as a feasible solution with very low complexity, which is suitable for large-scale antenna systems in practice. Moreover, the outcome (i.e., the obtained data rate) of the ZF method is only related to the RISbased analog beamforming via the power constraint in the original problem, which efficiently eliminates the interference. Other techniques for digital beamforming can also be utilized and do not influence the insights obtained in our work.

Algorithm	I Digital	Beamforming	Algorithm

- 1: Solve power allocation problem (22);
- 2: Obtain the optimal power allocation result (23);
- 3: Derive the beamformer matrix from the optimal power allocation based on (20);

#### B. RIS Configuration Based Analog Beamforming Algorithm

Since we iterate between the digital beamforming and RIS configuration based analog beamforming, the latter can be optimized assuming ZF precoding as shown in (20). Since the data rate with ZF precoding in (22) only depends on the RIS configuration through the power constraint (22b), the RIS configuration based analog beamforming problem can be reformulated as a power minimization problem:

$$\min_{\theta_{l_1,l_2}} \quad f(F), \tag{24a}$$

subject to  $\theta_{l_1,l_2} = \frac{m_{l_1,l_2}\pi}{2^{b-1}}, m_{l_1,l_2} \in \{0, 1, \dots, 2^b - 1\},$ (24b)

where

$$f(\boldsymbol{F}) = \operatorname{Tr}(\tilde{\boldsymbol{V}}_D \boldsymbol{P} \tilde{\boldsymbol{V}}_D^H) = \operatorname{Tr}(\boldsymbol{P}^{\frac{1}{2}} \tilde{\boldsymbol{V}}_D^H \tilde{\boldsymbol{V}}_D \boldsymbol{P}^{\frac{1}{2}})$$
  
= Tr(( $\tilde{\boldsymbol{F}} \tilde{\boldsymbol{F}}^H$ )<sup>-1</sup>). (25)

Here,  $\tilde{F} = P^{-\frac{1}{2}}F$ .

Since  $\tilde{F}\tilde{F}^{H}$  is a symmetric, positive semi-definite matrix, we can transform this problem into a semi-definite programming (SDP) problem. Let  $\text{Tr}((\tilde{F}\tilde{F}^{H})^{-1}) = \text{Tr}(\frac{w}{K}I_{K})$ . According to Schur complement [28], the problem can be rewritten by

$$\min_{\theta_{l_1,l_2,w}} w, \tag{26a}$$

subject to 
$$\boldsymbol{Z} = \begin{bmatrix} \frac{w}{K} \boldsymbol{I}_{K} & \boldsymbol{I}_{K} \\ \boldsymbol{I}_{K} & \tilde{\boldsymbol{F}} \tilde{\boldsymbol{F}}^{H} \end{bmatrix} \succeq 0,$$
 (26b)  
 $\theta_{l_{1},l_{2}} = \frac{m_{l_{1},l_{2}}\pi}{2^{b-1}}, m_{l_{1},l_{2}} \in \{0,1,\ldots,2^{b}-1\},$ 
(26c)

where  $X \succeq 0$  means that matrix X is a symmetric and positive semi-definite matrix. The equivalence between problems (24) and (26) is proved in Appendix A.

Remark on the tractability of the formulated problem: This problem is a mix-integer SDP, which is generally NP-hard. Moreover, any two discrete variables  $\{\theta_{l_1,l_2}\}$  are coupled with each other via constraint (26*b*), which makes the problem even more complicated. One commonly used solution is to first relax the discrete variables into continuous ones and then round the obtained solution to satisfy the discrete constraints. However, for the RIS-based systems, the typical value of the number of quantization bits is usually very small (e.g., 2 or 3) such that the round-off methods will lead to inevitable performance degrade.

To avoid the above issue, we consider to solve the SDP discretely. In the following, we first present the following Proposition 1 to transform the nonlinear functions in (26) with respect to  $\theta_{l_1,l_2}$  into linear ones, as proved in Appendix B.

We then use the outer approximation method [29] to solve this problem.

Proposition 1: Let

$$a = \left[ -\frac{(2^{b}-1)\pi}{2^{b-1}}, \dots, -\frac{m\pi}{2^{b-1}}, \dots, \frac{m\pi}{2^{B-1}}, \dots, \frac{(2^{b}-1)\pi}{2^{b-1}} \right],$$

$$c = \left[ \cos\left( -\frac{(2^{b}-1)\pi}{2^{b-1}} \right), \dots, \cos\left( -\frac{m\pi}{2^{b-1}} \right), \dots, \\ \cos\left( \frac{m\pi}{2^{b-1}} \right), \dots, \cos\left( \frac{(2^{b}-1)\pi}{2^{b-1}} \right) \right],$$

$$s = \left[ \sin\left( -\frac{(2^{b}-1)\pi}{2^{b-1}} \right), \dots, \sin\left( -\frac{m\pi}{2^{b-1}} \right), \dots, \\ \sin\left( \frac{m\pi}{2^{b-1}} \right), \dots, \sin\left( \frac{(2^{b}-1)\pi}{2^{b-1}} \right) \right].$$

We introduce a 0-1 binary vector  $\mathbf{x}^{l_1,l_2}$  of length  $2^{(b+1)} - 1$ , where  $x_i^{l_1,l_2}$  indicates whether  $\theta_{l_1,l_2} = a_i$ , and a binary vector  $\mathbf{y}^{l_1,l_2,l_1',l_2'}$  of length  $2^{(b+1)} - 1$  for phase difference  $\Delta \theta_{l_1,l_2,l_1',l_2'} = \theta_{l_1,l_2} - \theta_{l_1',l_2'}$ . Therefore, problem (26) can be rewritten by

$$\min_{\boldsymbol{x}^{l_1, l_2}, \boldsymbol{y}^{l_1, l_2, l_1', l_2'}, w} w, \tag{27a}$$

subject to 
$$\boldsymbol{Z} = \begin{bmatrix} \frac{w}{K} \boldsymbol{I}_{K} & \boldsymbol{I}_{K} \\ \boldsymbol{I}_{K} & \tilde{\boldsymbol{F}} \tilde{\boldsymbol{F}}^{H} \end{bmatrix} \succeq 0,$$
 (27b)

$$\begin{aligned} \| \boldsymbol{x}^{l_{1},l_{2}} \|_{1} &= 1, \boldsymbol{e}^{T} \boldsymbol{x}^{l_{1},l_{2}} = 0, \quad (27c) \\ \boldsymbol{a}^{T} (\boldsymbol{x}^{l_{1},l_{2}} - \boldsymbol{x}^{l_{1}',l_{2}'}) &= \boldsymbol{a}^{T} \boldsymbol{y}^{l_{1},l_{2},l_{1}',l_{2}'}. \end{aligned}$$
(27d)

Here, e is a constant vector whose first  $2^B - 1$  elements are 1 and others are 0.

Problem (27) is a mix-integer SDP with linear constraints, which can be solved by the outer approximation method. The basic idea of the outer approximation method is to enforce the SDP constraint via linear cuts and transform the original problem into a mix-integer linear programming one, which can be solved by the branch-and-bound (B&B) algorithm [30]. In the following, we will elaborate on how to enforce the SDP constraints via linear cuts.

Assume that a solution is  $\bar{w}, \bar{x}^{l_1, l_2}, \bar{y}^{l_1, l_2, l_1', l_2'}$ . In most mixinteger programming problems, it is very common to use the gradient cuts to approach the feasible set. However, the function of smallest eigenvalues is not always differentiable. Therefore, we use the characterization instead [29]. Note that  $Z \succeq 0$  is equivalent to  $u^T Z u \ge 0$  for arbitrary u. If Z with  $\bar{w}, \bar{x}^{l_1, l_2}, \bar{y}^{l_1, l_2, l_1', l_2'}$  is not positive semi-definite, we compute eigenvector u associated with the smallest eigenvalue. Then

$$\boldsymbol{u}^T \boldsymbol{Z} \boldsymbol{u} \ge 0 \tag{28}$$

is a valid cut that cuts off  $\bar{w}, \bar{x}^{l_1, l_2}, \bar{y}^{l_1, l_2, l_1', l_2'}$ . The RIS configuration based analog beamforming algorithm can be summarized in Algorithm 2.

Note that the outer approximation method might involve exponential complexity due to the B&B procedures. We also propose a sub-optimal algorithm to solve subproblem (24) based on discrete alternative optimization. Specifically, we select  $N_R \times N_R - 1$  phase shifts and fix them in each Algorithm 2 RIS Configuration Based Analog Beamforming Algorithm

1: Remove the semi-definite constraint of problem (27) and solve an initial solution;

- 3: Use the branch-and-bound method to solve problem (27) and obtain the optimal solution  $x^{l_1,l_2}$  for each RIS element;
- 4: Check the feasibility;
- 5: If the obtained solution is not feasible, add a cut according to (28);
- 6: until The obtained solution is feasible;
- 7: Derive the phase shits according to the obtained solution.

iteration such that the optimal solution of the remaining discrete phase shift  $q_{l_1,l_2}$  can be found by traversing constraint (26c), i.e.,

$$q_{l_1,l_2}^* = \arg\min_{q_{l_1,l_2}} f(\boldsymbol{F}).$$
<sup>(29)</sup>

We keep updating the phase shifts repeatedly until convergence.

## C. Overall Algorithm Description

Based on the results presented in the previous two subsections, we propose an overall iterative algorithm, i.e., the SRM algorithm, for solving the original problem in an iterative manner. Specially, the beamformer  $V_D$  is solved by Algorithm 1 while keeping the RIS configuration fixed. After obtaining the results, we will optimize the RIS configuration  $\theta^{l_1,l_2}$  by either Algorithm 2 or the ODS-based algorithm. Those obtained results are set as the initial solution for subsequent iterations. Define R as the value of the objective function. The two subproblems will be solved alternatively until in iteration t the value difference of the objective functions between two adjacent iterations is less than a predefined threshold  $\pi$ , i.e.,  $R^{(t+1)} - R^{(t)} \leq \pi$ .

## D. Convergence and Complexity Analysis

We now analyze the convergence and complexity of our proposed SRM algorithm (B&B adopted).

1) Convergence: First, according to Algorithm 1, in the digital beaforming subproblem, we can obtain a better result given RIS configuration  $\theta^{(t)}$  in the (t + 1)-th iteration. Therefore, we have

$$R(\boldsymbol{V}_D^{(t+1)}, \boldsymbol{\theta}^{(t)}) \ge R(\boldsymbol{V}_D^{(t)}, \boldsymbol{\theta}^{(t)}).$$
(30)

Second, given the beamforming result  $V_D^{(t+1)}$ , we maximize sum rate of all users, and thus, the following inequality holds:

$$R(V_D^{(t+1)}, Q^{(t+1)}) \ge R(V_D^{(t+1)}, \theta^{(t)}).$$
 (31)

Based the above inequalities, we can obtain

$$R(V_D^{(t+1)}, \theta^{(t+1)}) \ge R(V_D^{(t)}, Q^{(t)}).$$
(32)

which implies that the objective value of the original problem is non-decreasing after each iteration of the SRM algorithm. Since the objective value is upper bounded, the proposed SRM algorithm is guaranteed to converge.

2) *Complexity:* We consider the complexity of the proposed algorithms for two subproblems separately.

- In the digital beamforming subproblem, we need to optimize the received power for each user according to (23). Therefore, its computational complexity is O(K).
- In the RIS configuration based analog beamforming subproblem, we analyze the complexity of two algorithms. For the optimal Algorithm 2, we solve a series of linear programs by the branch-and-bound method. Since only one element in  $x^{l_1,l_2}$  can be 1,  $x^{l_1,l_2}$  can have at most  $2^b$  possible solutions. Thus, the scale of the computational complexity of each linear program is  $O(2^{bN_R^2})$ . For the sub-optimal discrete alternative optimization algorithm, a number of iterations is required for convergence, each of which has a complexity of  $O(2^bN_R^2)$ . Since the type value of b, i.e., the number of quantification bits, is usually within the range of  $2 \sim 5$ , the complexity of each iteration is low enough.

## V. PERFORMANCE ANALYSIS OF RIS-BASED Multi-User Communications

In this section, we compare the RIS-based HBF scheme with the traditional ones in terms of the minimum number of required RF chains. A special case, i.e., the pure LoS transmission, is also considered to explore theoretically how the size of the RIS and its placement influence the achievable rates.

#### A. Comparison With Traditional Hybrid Beamforming

We adopt the fully digital beamforming scheme as a benchmark to compare the traditional and RIS-based HBF schemes. In the traditional HBF, it has already been proved that when the number of RF chains is not smaller than twice the number of target data streams, any fully digital beamforming matrix can be realized [19]. However, in the RIS-based HBF, the inherent analog beamforming (i.e., the RIS configuration) is closely coupled with propagation, which offers more freedom for shaping the propagation environment than the traditional scheme. To capture this characteristic, we explore a new condition for the RIS-based system to achieve fully digital beamforming.

We start by describing the fully digital beamforming scheme in an RIS-based system. Consider an ideal case where each RIS element directly connects with an RF chain and ADC as if it is part of the BS. <sup>3</sup> The fully digital beamformer can then be denoted by  $V_{FD} \in N_R^2 \times K$ , based on which we present the following proposition, which will be proved in Appendix C.

Proposition 2: For the RIS-based HBF scheme with  $N_R^2 \ge KN_t$ , to achieve any fully digital beamforming scheme, the number of transmit antennas at the BS should not be smaller than the number of single-antenna users, i.e.,  $N_t \ge K$ .

<sup>2:</sup> repeat

<sup>&</sup>lt;sup>3</sup>In practice, the RIS does not connect to the RF chain directly unless it is installed at the BS, which is not the truth in our case where RIS actually only reflects signals. Therefore, we only consider such an ideal scheme as a benchmark to evaluate the effectiveness of RIS-based HBF.

This implies two conditions for the proposed RIS-based scheme to achieve the fully digial beamforming scheme. First, the size of the RIS should be no smaller than the product of the number of users and the size of the antenna array at the BS. Second, the number of transmit antennas at the BS should be no smaller than the number of single-antenna users.

Remark on dedicated hardware reduction for analog beamforming: In the traditional HBF, to offload part of the digital baseband processing to the analog domain, a number of RF chains are required at the BS to feed the analog beamformer, equipped with necessary hardwares such as mixers, filters, and phase shifters [31]. In contrast, as shown in Proposition 2, the number of minimum RF chains required by the RIS-based HBF to achieve the fully digial beamforming has already been reduced by half compared to the traditional scheme. Moreover, the phase shifters can be saved since the RIS inherently realizes analog beamforming owing to its flexible physical structure [32].

## B. Special Case: Pure Line-of-Sight Transmissions

We consider the data rate obtained by the pure LoS case as a lower bound of the achievable rate and analyze optimal RIS placement to provide orthogonal communication links. The LoS case also reveals insights on how the achievable data rate is influenced by RIS design and placement.

Since a multitude of RIS elements are placed in a subwavelength order, spacial correlation between these elements are inevitable. In this case, the channel matrix **F** approaches a low-rank matrix, leading to a degraded performance in terms of the achievable data rate. Traditionally, we can utilize the multi-path effect to decorrelate different channel links between the transceiver antennas. However, it is also important to understand how the system works in the pure LoS case, especially when it comes to the RIS-based systems where the reflection-based link between the BS and users acts as the dominated "LoS" component and is usually much stronger than other multi-paths as well as the degraded direct links.

In the pure LoS case, we aim to achieve a high-rank LoS channel matrix by designing the size of the RIS and the antenna array at the BS. Different from those existing works on antenna array design for LoS MIMO systems, the discrete phase shifts depicted by  $\{q_{l_1,l_2}\}$  needs to be considered in the RIS-based propagation. We thus present the results in Proposition 3 below, which is proved in Appendix D.

Proposition 3: In the RIS-based system, to make different links between the BS and user k via the RIS orthogonal to each other, the RIS design should satisfy the following conditions:

$$d_{R}^{(1)}d_{B} = \frac{\lambda D_{0,0}^{(0)}}{N_{R}\cos\theta_{R}\cos\theta_{B}},$$
(33a)
$$\sum_{l_{2}=0}^{N_{R}-1} (1+\sin\theta_{l_{1},l_{2}})$$

$$= \sum_{l_{2}=0}^{N_{R}-1} (1+\sin\theta_{l_{1}',l_{2}}), \forall l_{1}, l_{1}' \leq N_{R}-1.$$
(33b)

This proposition shows that the achievable data rate is highly related to the size of the RIS and its placement. For convenience, when all other parameters are fixed, we refer to the *threshold* of the RIS size as

$$N_{R}^{th} = \frac{\lambda D_{0,0}^{(0)}}{d_{R}^{(1)} d_{B} \cos \theta_{R} \cos \theta_{B}}.$$
 (34)

For the pure LoS case, the sum rate maximization problem can still be formulated as (17) with one extra constraint (33b). Our proposed SRM algorithm can be utilized to solve this problem after we convert the extra constraint (33b) into a linear one shown below by following the transformations in Proposition 1, i.e.,

$$\sum_{l_2=0}^{N_r-1} (\boldsymbol{x}^{l_1,l_2} - \boldsymbol{x}^{l_1',l_2}) \boldsymbol{s}^T = 0.$$
(35)

### VI. SIMULATION RESULTS

In this section, we evaluate the performance of our proposed algorithm for RIS-based HBF in terms of the sum rate. We show how the system performance is influenced by the SNR, number of users, the size of the RIS, and the number of quantization bits for discrete phase shifts. For comparison, the following algorithms are performed as well.

- *Simulated annealing*: We utilize the simulated annealing method [33] to approach the global optimal solution of the sum rate maximization problem with discrete phase shifts. The maximum number of iterations is set as 10<sup>7</sup>.
- *Pure LoS case*: We consider the pure LoS case as a lower bound to evaluate the performance of the HBF scheme. The proposed iterative SRM algorithm can still be utilized to solve the corresponding optimization problem .
- *Random phase shift*: Algorithm 1 is first performed, followed by a random algorithm to solve the RIS-based analog beamforming subproblem (19).
- *RIS-based HBF with continuous phase shifts*: This scheme only serves as a benchmark when we investigate how the discreteness level (i.e., the number of quantization bits, *b*) influences the sum rate. The discrete constraint in the original subproblem (19) is relaxed to  $\theta_{l_1,l_2} \in [0, 2\pi]$ . The HBF solution is obtained by iteratively performing Algorithm 1 and the gradient descent method.
- *Relaxation-based optimization [16]*: To solve the RISbased analog beamforming subproblem, the discrete phase shifts are first relaxed to continuous ones and the alternative optimization is then performed, followed by rounding off the outcome to discrete ones.

In our simulation, we set the distance between the BS and the RIS,  $D_{0,0}^{(0)}$ , as 20m, and users are randomly deployed within a half circle of radius 60m centering at the RIS. The antenna array at the BS and the RIS are placed at angles of 15° and 30° to the x axis, respectively. The transmit power of the BS  $P_T$  is 20 W, the carrier frequency is 5.9 GHz, the antenna separation at the BS  $d_B$  is 1 m, the RIS element separation  $d_R$  is 0.03 m, and the Rician fading parameter  $\kappa$  is 4 [34]. For convenience,



Fig. 5. Sum rate v.s. SNR  $(K = N_t = 5, b = 2, N_R = 3)$ .



Fig. 6. Sum rate v.s. size of RIS  $N_R$  (SNR = 2dB,  $K = N_t = 5, b = 2$ ).

the path loss factors of both LoS and NLoS links are the same, i.e.,  $\alpha = 2$ . We set the size of the RIS  $N_R^2$  ranging between  $5^2 \sim 65^2$ , the number of antennas at the BS  $N_t$  and the number of users K between 5 ~ 15, the discreteness level of RIS b between 1 ~ 5, and the SNR (defined as  $P_T/\sigma^2$ ) between -2 dB ~ 10 dB. Specifically, for the LoS case, given the above parameters the designing rule in (33a) is only satisfied when the size of RIS is set as 40.

Fig. 5 shows the sum rate of all users versus SNR, obtained by different algorithms with an RIS of size  $3 \times 3$  (i.e.,  $N_R = 3$ ), b = 2 quantization bits for phase shifts, equal number of transmit antennas at the BS and the downlink users. The sum rate increases with SNR since more power resources are allocated by the BS. We observe that the performance of our proposed B&B based algorithm (i.e., Algorithm 2) is close to that of the simulated annealing method and much better than both the random algorithm and the proposed discrete alternative optimization algorithm. Moreover, both our proposed algorithms outperform the benchmark, i.e., the relaxationbased algorithm, since the rounding-off in the benchmark may lead to some performance loss. This indicates the efficiency of our proposed algorithms to solve the RIS-based HBF problem.

Fig. 6 shows the sum rate of all users versus the size of the RIS<sup>4</sup> with b = 2,  $N_t = K = 5$ . The suboptimal





Fig. 7. Sum rate v.s. number of quantization bits b (SNR = 2dB,  $K = N_t = 5$ ,  $N_R = 40$ ).

discrete iterative optimization is adopted for RIS-based analog beamforming. We observe that the sum rate grows rapidly with a small size of RIS and gradually flattens as the size of RIS continues to increase. <sup>5</sup> The inflection point of each curve shows up around  $N_R = 40$ , which verifies the threshold (34) given by Proposition 3 very well. When the RIS size  $N_R$  exceeds 40, though Proposition 3 does not hold, the sum rate does not drop since the RIS can always turn off those extra RIS elements to maintain the sum-rate performance. This figure also implies that though Proposition 3 is obtained in the pure LoS case, it also sheds insight into the RIS placement and array design in a more general case with small-scale fading.

Moreover, Fig. 6 also shows that the performance of RISbased beamforming with small-scale fading is much better than that of the pure LoS case when the size of the RIS is small. Such a gain comes from the reduced correlation between different channel links owing to multi-path effects. As the size of the RIS grows, Proposition 3 is satisfied such that the channel links in the pure LoS case are orthogonalized, making the gap between these two cases smaller.

Fig. 7 depicts the sum rate of all users versus the number of quantization bits b for discrete phase shifts in RIS configuration with SNR = 2 dB,  $N_t = K = 5$ , and  $N_R = 40$ . Considering the complexity of algorithms, we adopt the proposed discrete alternative optimization algorithm for RISbased analog beamforming to obtain a sub-optimal solution. As the number of quantization bits increases, the sum rate obtained by our proposed algorithm with discrete phase shifts approaches that in the continuous case. When the size of RIS grows, the gap between the discrete and continuous cases shrinks since a larger RIS usually provides more freedom of generating directional beams. Note that the implementation difficulty increases dramatically in practice with the number of quantization bits. We remark that a small number of discrete phase shifts (e.g.,  $2 \sim 3$  quantification bits) is already enough to obtain a satisfying performance.

Denote the number of iterations (i.e., the number of linear integer programs) required for Algorithm 2 to converge as Y.

<sup>&</sup>lt;sup>5</sup>We do not show the simulated annealing algorithm in this figure due to its high complexity with a large RIS.



Fig. 8. C.D.F. of the number of iterations to converge for Algorithm 2 ( $K = N_t = 5$  and b = 2).

Fig. 8 shows the cumulative distribution function (C.D.F.) of Y v.s. Y for different RIS sizes with SNR = 2 dB,  $N_t = K = 5$ , and b = 2. We observe that the convergence speed grows faster when the RIS size becomes smaller. The algorithm can converge within 10 iterations for most cases, e.g., only 6 iterations are enough for 90% cases to converge when the RIS size is 4 ×4.

## VII. CONCLUSION AND DISCUSSION

In this paper, we have studied an RIS-based downlink multi-user multi-antenna system in the absence of direct links between the BS and users. The BS transmits signals to users via the reflection-based RIS with limited discrete phase shifts. Considering the close coupling between channel propagation and the RIS configuration pattern selection, we have carried out an HBF scheme for sum rate maximization. Continuous digital beamforming has been performed at the BS and discrete analog beamforming has been achieved inherently at the RIS via configuration pattern selection. The sum rate maximization problem has been decomposed into two subproblems and solved iteratively by our proposed SRM algorithm.

Three remarks can be drawn from the theoretical analysis and numerical results, providing insights for RIS-based system design.

- The sum rate of the RIS-based system with discrete phase shifts increases rapidly when the number of quantization bits b is small, and gradually approaches the sum rate achieved in the continuous case if b is large enough.
- The sum rate increases with the size of the RIS and converges to a stable value as the size of the RIS grows to reach the threshold determined by Proposition 3.
- The minimum number of transmit antennas at the BS required to achieve any fully digital beamforming scheme is only half of that in traditional HBF schemes, implying that the RIS-based HBF scheme can greatly reduce the cost of dedicated hardware.

The above observations indicate that when designing RISbased systems, a moderate sized RIS and a very small number of quantization bits are enough to achieve a satisfactory sum rate at low cost.

We let  $\operatorname{Tr}((\tilde{F}\tilde{F}^{H})^{-1}) = \operatorname{Tr}(\frac{w}{K}I_{K})$  and then prove that minimizing  $\operatorname{Tr}((\tilde{F}\tilde{F}^{H})^{-1})$  is equivalent to minimizing w. Inspired by Schur complement, we construct a new matrix

$$\boldsymbol{Z} = \begin{bmatrix} \frac{w}{K} \boldsymbol{I}_{K} & \boldsymbol{I}_{K} \\ \boldsymbol{I}_{K} & \tilde{\boldsymbol{F}} \tilde{\boldsymbol{F}}^{H} \end{bmatrix}.$$
 (36)

Since Z is symmetric and  $\tilde{F}\tilde{F}^{H}$  is invertible, we have

$$Z \succeq 0 \iff \tilde{F}\tilde{F}^{H} \succeq 0 \text{ and } \left(\tilde{F}\tilde{F}^{H}\right)^{-1} \preceq \frac{w}{K}I_{K}.$$
 (37)

Therefore, we have

$$\operatorname{Tr}((\tilde{\boldsymbol{F}}\tilde{\boldsymbol{F}}^{H})^{-1}) \leq \operatorname{Tr}(\frac{w}{K}\boldsymbol{I}_{K}) = w.$$
(38)

This completes the proof.

# APPENDIX B PROOF OF PROPOSITION 1

We first rewrite F as

$$F = \sum_{l_1, l_2} q_{l_1, l_2} G_{l_1, l_2}, \qquad (39)$$

where  $G_{l_{1},l_{2}} = \left[\theta^{(1)}h_{l_{1},l_{2}}^{(1)}, \cdots, \theta^{(K)}h_{l_{1},l_{2}}^{(K)}\right]^{T}$  and  $h_{l_{1},l_{2}}^{(k)} = \left[h_{l_{1},l_{2}}^{(k,1)}, \cdots, h_{l_{1},l_{2}}^{(k,N_{T})}\right]$ . We then have  $\tilde{F}\tilde{F}^{H}$   $= P^{-\frac{1}{2}}\sum_{l_{1},l_{2}}q_{l_{1},l_{2}}G_{l_{1},l_{2}}\sum_{l_{1}',l_{2}'}q_{l_{1}',l_{2}'}^{H}G_{l_{1}',l_{2}'}P^{-\frac{1}{2}}$   $= \sum_{l_{1},l_{2}}q_{l_{1},l_{2}}\sum_{l_{1}',l_{2}'}q_{l_{1}',l_{2}'}^{H}P^{-\frac{1}{2}}G_{l_{1},l_{2}}G_{l_{1}',l_{2}'}^{H}P^{-\frac{1}{2}}$   $= \sum_{l_{1},l_{2}}\sum_{l_{1}',l_{2}'}\frac{(j+e^{j\theta_{l_{1},l_{2}}})(-j+e^{-j\theta_{l_{1}',l_{2}'}})}{4}A_{l_{1},l_{2},l_{1}',l_{2}'}$   $= \sum_{l_{1},l_{2}'}\sum_{l_{1}',l_{2}'}\frac{A_{l_{1},l_{2},l_{1}',l_{2}'}}{4}\left(\cos(\theta_{l_{1},l_{2}}-\theta_{l_{1}',l_{2}'})+j\sin(\theta_{l_{1},l_{2}}-\theta_{l_{1}',l_{2}'})\right)$   $+ \sum_{l_{1},l_{2}}\sum_{l_{1}',l_{2}'}\frac{A_{l_{1},l_{2},l_{1}',l_{2}'}}{4}\left(\cos(\Delta\theta_{l_{1},l_{2},l_{1}',l_{2}'})+j\sin(\Delta\theta_{l_{1},l_{2},l_{1}',l_{2}'})\right)$   $+ \sum_{l_{1},l_{2}}\sum_{l_{1}',l_{2}'}\frac{A_{l_{1},l_{2},l_{1}',l_{2}'}}{4}\left(\cos(\theta_{l_{1}',l_{2}})+\cos(\theta_{l_{1},l_{2}})+1\right),$ (40)

where  $A_{l_1,l_2,l'_1,l'_2} = P^{-\frac{1}{2}} G_{l_1,l_2} G^H_{l'_1,l'_2} P^{-\frac{1}{2}}$ . Note that  $\tilde{F} \tilde{F}^H$  is not linear with respect to  $\theta_{l_1,l_2}$ . Taking the advantage of the discrete property of  $\theta_{l_1,l_2}$ , we can further transform the non-linear functions into linear ones.

With the definitions of  $x^{l_1,l_2}$ , we have

$$\cos(\theta_{l_1,l_2}) = \boldsymbol{x}^{l_1,l_2} \boldsymbol{c}^T, \quad \sin(\theta_{l_1,l_2}) = \boldsymbol{x}^{l_1,l_2} \boldsymbol{s}^T, \quad (41)$$

with  $\|\boldsymbol{x}^{l_1,l_2}\|_1 = 1$ . It is also worthwhile to point out that the value of  $\theta_{l_1,l_2}$  only falls in the range  $[0, 2\pi)$ , and thus, we have

$$e^T x^{l_1, l_2} = 0. (42)$$

Similarly, according to the definitions of  $y^{l_1,l_2,l_1',l_2'}$ , we have

$$\cos(\Delta \theta_{l_1, l_2, l_1', l_2'}) = \boldsymbol{y}^{l_1, l_2, l_1', l_2'} \boldsymbol{c}^T, \sin(\Delta \theta_{l_1, l_2, l_1', l_2'}) = \boldsymbol{y}^{l_1, l_2, l_1', l_2'} \boldsymbol{s}^T,$$
(43)

with

$$\boldsymbol{a}^{T}(\boldsymbol{x}^{l_{1},l_{2}}-\boldsymbol{x}^{l_{1}',l_{2}'}) = \boldsymbol{a}^{T}\boldsymbol{y}^{l_{1},l_{2},l_{1}',l_{2}'}.$$
(44)

With these transformations,  $\tilde{F}\tilde{F}^{H}$  is linear with respect to  $x^{l_1,l_2}$  and  $y^{l_1,l_2,l_1',l_2'}$ .

# APPENDIX C Proof of Proposition 2

We first prove that when  $N_t = K$ , any fully digital beamforming scheme can be achieved by the RIS-based HBF scheme if  $N_R^2 \ge KN_t$  holds. We then state that it is not possible to achieve digital beamforming when  $N_t < K$ . Therefore,  $N_t \ge 2K$  is a sufficient condition.

i) Denote the channel matrix between the RIS and the users as  $H_{FD} \in {}^{K \times N_R^2}$ . Since  $N_t = K$ , the *k*th column of the digital beamformer can be expressed by  $\mathbf{V}_{D,k} = \begin{bmatrix} 0^T, v_{k,k}, 0^T \end{bmatrix}^T$ , where  $v_{k,k}$  is the *k*th element of the the *k*th column in  $\mathbf{V}_D \in \mathbb{R}^{N_t \times K}$ . To satisfy  $\mathbf{H}_{FD}\mathbf{V}_{FD} = \mathbf{F}\mathbf{V}_D$ , we have

$$[\dots, f_{m,k}, \dots] \begin{bmatrix} 0\\ \vdots\\ v_{k,k}\\ \vdots\\ 0 \end{bmatrix} = (\mathbf{H}_{FD} \mathbf{V}_{FD})_{m,k}, \qquad (45)$$

i.e.,

$$\sum_{l_1,l_2} \frac{1}{2} \phi^{(k)} \left[ \left( j + e^{j\theta_{l_1,l_2}} \right) g_{l_1,l_2}^{m,k} \right] v_{k,k} = \left( H_{FD} V_{FD} \right)_{m,k}, \quad (46)$$

for all  $0 \le m, k \le K$ , where  $f_{m,k}$  is the element of matrix **F**. Note that the term multiplied by  $v_{k,k}$  in (46) can achieve different magnitudes owing to the linear combination of channel coefficients, which is different from the traditional HBF [23]. At least one solution can be found for this set of equations if the number of equations is no smaller than the number of variables, i.e.,  $N_R^2 \ge KN_t$ . This also holds for  $N_t > K$  since we can always use the solution for  $N_t = K$  and set those extra variables to be zero.

ii) We observe that  $rank(\mathbf{H}_{FD}\mathbf{V}_{FD}) = K$  and  $rank(\mathbf{F}\mathbf{V}_D) = \min\{K, N_t\}$ . If  $N_t < K$ , then  $rank(\mathbf{H}_{FD}\mathbf{V}_{FD}) > rank(\mathbf{F}\mathbf{V}_D)$ , implying that the RIS-based HBF cannot implement the fully digital beamforming scheme. This completes the proof.

#### APPENDIX D

## **PROOF OF PROPOSITION 3**

Note that each user k goes through different path losses due to various positions, which naturally varies the corresponding channel coefficients. Therefore, we only focus on orthogonalizing different links from the BS to the same user. Since the path loss is the same for different links with respect to user k, we consider the channel response between one BS antenna n and all RIS elements as

$$\mathbf{f}^{(n,k)} = \left[ q_{0,0} e^{-j\frac{2\pi}{\lambda} \left( D_{0,0}^{(n)} + d_{0,0}^{(k)} \right)}, q_{0,1} e^{-j\frac{2\pi}{\lambda} \left( D_{0,1}^{(n)} + d_{0,1}^{(k)} \right)}, \cdots, q_{N_R - 1, N_R - 1} e^{-j\frac{2\pi}{\lambda} \left( D_{N_R - 1, N_R - 1}^{(n)} + d_{N_R - 1, N_R - 1}^{(k)} \right)} \right].$$
(47)

To keep any two different links orthogonal to each other, the following condition should be satisfied,

$$\mathbf{f}^{(n_a,k)}\Big]^H \cdot \mathbf{f}^{(n_b,k)} = 0, \forall n_a \neq n_b,$$
(48)

where  $n_a$  and  $n_b$  denote two different transmit antennas. By substituting (47), (48) can be rewritten by

$$\sum_{l_2=0}^{N_R-1} \sum_{l_1=0}^{N_R-1} q_{l_1,l_2} \cdot (q_{l_1,l_2})^* e^{j\frac{2\pi}{\lambda} \left( D_{l_1,l_2}^{(n_a)} - D_{l_1,l_2}^{(n_b)} \right)} = 0.$$
(49)

We substitute (1) and (8) into (49) and obtain the following

$$\sum_{l_1=0}^{N_R-1} \left[ e^{j\frac{2\pi}{\lambda D_{0,0}^{(0)}} l_1 d_R^{(1)} d_B (n_b - n_a) \cos \theta_R \cos \theta_B} \right] = 0, \quad (50)$$

where  $n_b - n_a \in \mathbb{Z}$ . Since this condition should hold for any two transmit antennas  $n_a$  and  $n_b$ , we then have (33) based on the principle of geometric sums.

#### REFERENCES

- E. W. Paper, "More than 50 billion connected devices," Utah Educ. Network, Salt Lake City, UT, USA, Tech. Rep. 284 23-3149, Feb. 2011.
- [2] L. Song, "Relay selection for two-way relaying with amplify- andforward protocols," *IEEE Trans. Veh. Technol.*, vol. 60, no. 4, pp. 1954–1959, May 2011.
- [3] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.
- [4] D. R. Smith, O. Yurduseven, L. P. Mancera, P. Bowen, and N. B. Kundtz, "Analysis of a waveguide-fed metasurface antenna," *Phys. Rev. A, Gen. Phys. Appl.*, vol. 8, no. 5, pp. 1–16, Nov. 2017.
- [5] M. D. Renzo *et al.*, "Smart radio environments empowered by reconfigurable AI meta-surfaces: An idea whose time has come," *EURASIP J. Wireless Commun. Netw.*, vol. 2019, no. 1, pp. 1–20, May 2019.
- [6] T. Cui, D. Smith, and R. Liu, *Metamaterials: Theory, Design, and Applications.* New York, NY, USA: Springer, 2010.
- [7] L. Li *et al.*, "Machine-learning reprogrammable metasurface imager," *Nature Commun.*, vol. 10, no. 1, pp. 1–8, Dec. 2019.
  [8] H.-T. Chen, A. J. Taylor, and N. Yu, "A review of metasurfaces:
- [8] H.-T. Chen, A. J. Taylor, and N. Yu, "A review of metasurfaces: Physics and applications," *Rep. Prog. Phys.*, vol. 79, no. 7, Jun. 2016, Art. no. 076401.
- [9] N. Shlezinger, O. Dicker, Y. C. Eldar, I. Yoo, M. F. Imani, and D. R. Smith, "Dynamic metasurface antennas for uplink massive MIMO systems," *IEEE Trans. Commun.*, vol. 67, no. 10, pp. 6829–6843, Oct. 2019.
- [10] 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, Aug. 2019.
- [11] X. Yu, D. Xu, and R. Schober, "MISO wireless communication systems via intelligent reflecting surfaces," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC)*, Changchun, China, Aug. 2019, pp. 735–740.

- [12] Y. Han, W. Tang, S. Jin, C.-K. Wen, and X. Ma, "Large intelligent surface-assisted wireless communication exploiting statistical CSI," *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 8238–8242, Aug. 2019.
- [13] C. Huang, G. C. Alexandropoulos, A. Zappone, M. Debbah, and C. Yuen, "Energy efficient multi-user MISO communication using low resolution large intelligent surfaces," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Abu Dhabi, UAE, Dec. 2018, pp. 1–6.
- [14] S. Nie, J. M. Jornet, and I. F. Akyildiz, "Intelligent environments based on ultra-massive MIMO platforms for wireless communication in millimeter wave and terahertz bands," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Brighton, U.K., May 2019, pp. 7849–7853.
- [15] Q.-U.-U. Nadeem, A. Kammoun, A. Chaaban, M. Debbah, and M.-S. Alouini, "Intelligent reflecting surface assisted wireless communication: Modeling and channel estimation," 2019, arXiv:1906.02360. [Online]. Available: http://arxiv.org/abs/1906.02360
- [16] Q. Wu and R. Zhang, "Beamforming optimization for wireless network aided by intelligent reflecting surface with discrete phase shifts," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1838–1851, Mar. 2020.
- [17] H. Guo, Y.-C. Liang, J. Chen, and E. G. Larsson, "Weighted sumrate optimization for intelligent reflecting surface enhanced wireless networks," 2019, arXiv:1905.07920. [Online]. Available: http://arxiv. org/abs/1905.07920
- [18] A. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Press, 2005.
- [19] F. Sohrabi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 501–513, Apr. 2016.
- [20] O. Ozdogan, E. Bjornson, and E. G. Larsson, "Intelligent reflecting surfaces: Physics, propagation, and pathloss modeling," *IEEE Wireless Commun. Lett.*, vol. 9, no. 5, pp. 581–585, May 2020.
- [21] F. Bohagen, P. Orten, and G. Oien, "Optimal design of uniform rectangular antenna arrays for strong line-of-sight MIMO channels," *EURASIP J. Wireless Commun. Network.*, vol. 2007, no. 2, Jan. 2007, Art. no. 045084.
- [22] H. Zhang, B. Di, L. Song, and Z. Han, "Reconfigurable intelligent surfaces assisted communications with limited phase shifts: How many phase shifts are enough?" *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 4498–4502, Apr. 2020.
- [23] X. Zhang, A. F. Molisch, and S.-Y. Kung, "Variable-phase-shift-based RF-baseband codesign for MIMO antenna selection," *IEEE Trans. Signal Process.*, vol. 53, no. 11, pp. 4091–4103, Nov. 2005.
- [24] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.
- [25] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication—Part I: Channel inversion and regularization," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 195–202, Jan. 2005.
- [26] D. Tse and P. Viswanath, Fundamentals of Wireless Communication. Cambridge, U.K.: Cambridge Univ. Press, 2005.
- [27] N. J. Higham, Functions of Matrices: Theory and Computation. Philadelphia, PA, USA: SIAM, 2008.
- [28] J. Gallier, The Schur Complement and Symmetric Positive Semidefinite (and Definite) Matrices. Danboro, PA, USA: Penn Engineering, 2010.
- [29] T. Gally, M. E. Pfetsch, and S. Ulbrich, "A framework for solving mixedinteger semidefinite programs," *Optim. Methods Softw.*, vol. 33, no. 3, pp. 594–632, May 2017.
- [30] S. Zhang, H. Zhang, B. Di, and L. Song, "Cellular UAV-to-X communications: Design and optimization for multi-UAV networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 2, pp. 1346–1359, Feb. 2019.
- [31] A. Alkhateeb, J. Mo, N. Gonzalez-Prelcic, and R. W. Heath, Jr., "MIMO precoding and combining solutions for millimeter-wave systems," *IEEE Commun. Mag.*, vol. 52, no. 12, pp. 122–131, Dec. 2014.
- [32] M. Shehata, A. Mokh, M. Crussière, M. Helard, and P. Pajusco, "On the equivalence of hybrid beamforming to full digital zero forcing in mmWave MIMO," in *Proc. Int. Conf. Telecommun. (ICT)*, Hanoi, Vietnam, Apr. 2019, pp. 1–7.
- [33] B. Di, L. Song, and Y. Li, "Sub-channel assignment, power allocation, and user scheduling for non-orthogonal multiple access networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 11, pp. 7686–7698, Nov. 2016.
- [34] H. El-Sallabi, M. Abdallah, and K. Qaraqe, "Modelling of parameters of Rician fading distribution as a function of polarization parameter in reconfigurable antenna," in *Proc. IEEE/CIC Int. Conf. Commun. China* (*ICCC*), Shanghai, China, Oct. 2014, pp. 534–538.



**Boya Di** (Member, IEEE) received the B.S. degree in electronic engineering from Peking University in 2014 and the Ph.D. degree from the Department of Electronics, Peking University, China, in 2019. So far, she has contributed as the first author for nine journal articles and one of her journal papers is currently listed as ESI highly cited articles. Her current research interests include edge computing, vehicular networks, non-orthogonal multiple access, and 5G wireless networks. She has also served as a TPC Member in GLOBECOM 2016, ICCC 2017, ICC 2016, ICC 2018, and VTC 2019.



Hongliang Zhang (Member, IEEE) received the B.S. and Ph.D. degrees from the School of Electrical Engineering and Computer Science, Peking University, in 2014 and 2019, respectively. He is currently a Post-Doctoral Fellow with the Electrical and Computer Engineering Department and the Computer Science Department, University of Houston, TX, USA. His current research interest includes cooperative communications, the Internet of Things networks, hypergraph theory, and optimization theory. He has served as a TPC member for many IEEE

conferences, such as GLOBECOM, ICC, and WCNC. He is currently an Editor of *IET Communications*. He also serves as a Guest Editor for the IEEE INTERNET OF THINGS JOURNAL special issue on Internet of UAVs over Cellular Networks.



Lingyang Song (Fellow, IEEE) received the Ph.D. degree from the University of York, U.K., in 2007, where he received the K. M. Stott Prize for excellent research. He was a Research Fellow with the University of Oslo, Norway, until rejoining Philips Research, U.K., in March 2008. In May 2009, he joined the Department of Electronics, School of Electronics Engineering and Computer Science, Peking University, where he is currently a Boya Distinguished Professor. His main research interests include wireless communication and networks, sig-

nal processing, and machine learning. He was a recipient of the IEEE Leonard G. Abraham Prize in 2016 and the IEEE Asia Pacific (AP) Young Researcher Award in 2012. He has been an IEEE Distinguished Lecturer since 2015.



Yonghui Li (Fellow, IEEE) received the Ph.D. degree from the Beijing University of Aeronautics and Astronautics in November 2002. From 1999 to 2003, he was affiliated with Linkair Communication Inc., where he held a position of Project Manager with responsibility for the design of physical layer solutions for the LAS-CDMA system. Since 2003, he has been with the Centre of Excellence in Telecommunications, The University of Sydney, Australia. He is currently a Professor with the School of Electrical and Information Engineering,

The University of Sydney. His current research interests are in the area of wireless communications, with a particular focus on MIMO, millimeter wave communications, machine to machine communications, coding techniques, and cooperative communications. He holds a number of patents granted and pending in these fields. He was a recipient of the Australian Queen Elizabeth II Fellowship in 2008 and the Australian Future Fellowship in 2012. He received the best paper awards from the IEEE International Conference on Communications (ICC) 2014, the IEEE PIMRC 2017, and the IEEE Wireless Days Conferences (WD) 2014. He is currently an Editor of the IEEE TRANSACTIONS ON COMMUNICATIONS and the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. He was also the Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS Special issue on Millimeter Wave Communications for Future Mobile Networks.



**Zhu Han** (Fellow, IEEE) received the B.S. degree in electronic engineering from Tsinghua University in 1997, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Maryland at College Park, College Park, in 1999 and 2003, respectively.

From 2000 to 2002, he was an Research and Development Engineer of JDSU, Germantown, MD, USA. From 2003 to 2006, he was a Research Associate with the University of Maryland. From 2006 to 2008, he was an Assistant Professor with

Boise State University, ID, USA. He is currently a John and Rebecca Moores Professor with the Electrical and Computer Engineering Department and in the Computer Science Department, University of Houston, TX, USA. His research interests include wireless resource allocation and management, wireless communications and networking, game theory, big data analysis, security, and smart grid.

Dr. Han received an NSF Career Award in 2010, the Fred W. Ellersick Prize of the IEEE Communication Society in 2011, the EURASIP Best Paper Award for the *Journal on Advances in Signal Processing* in 2015, the IEEE Leonard G. Abraham Prize in the field of Communications Systems (Best Paper Award in IEEE JSAC) in 2016, and several best paper awards in IEEE conferences. He was an IEEE Communications Society Distinguished Lecturer from 2015 to 2018. He has been an AAAS fellow since 2019 and has also been an ACM Distinguished Member since 2019. He has been a 1% Highly Cited Researcher since 2017 according to Web of Science.



**H. Vincent Poor** (Life Fellow, IEEE) received the Ph.D. degree in electronics engineering and computer science (EECS) from Princeton University in 1977. From 1977 to 1990, he was on the faculty of the University of Illinois at Urbana–Champaign. Since 1990, he has been on the faculty at Princeton, where he is currently the Michael Henry Strater University Professor of Electrical Engineering. From 2006 to 2016, he served as the Dean for Princeton's School of Engineering and Applied Science. He has also held visiting appointments at several

other universities, including most recently at Berkeley and Cambridge. His research interests are in the areas of information theory, signal processing and matching learning, and their applications in wireless networks, energy systems, and related fields. Among his publications in these areas is the recent book *Multiple Access Techniques for 5G Wireless Networks and Beyond* (Springer, 2019).

Dr. Poor is a member of the National Academy of Engineering and the National Academy of Sciences, and is a foreign member of the Chinese Academy of Sciences, the Royal Society, and other national and international academies. Recent recognition of his work includes the 2017 IEEE Alexander Graham Bell Medal and a D.Eng. *honoris causa* from the University of Waterloo, awarded in 2019.