Reconfigurable Intelligent Surfaces with Reflection Pattern Modulation: Beamforming Design and Performance Analysis

Shaoe Lin, Beixiong Zheng, George C Alexandropoulos, Miao Wen, Marco Di Renzo, Fangjiong Chen

Reconfigurable Intelligent Surfaces with Reflection Pattern Modulation: Beamforming Design and Performance Analysis

Shaoe Lin, Beixiong Zheng, George C. Alexandropoulos, *Senior Member, IEEE*, Miaowen Wen, *Senior Member, IEEE*, Marco Di Renzo, *Fellow, IEEE*, and Fangjiong Chen

***Abstract*—Recent research on reconfigurable intelligent surfaces (RISs) suggests that RISs can perform passive beamforming and information transfer (PBIT) simultaneously via smart reflections. In this paper, we propose an RIS-enhanced multiple-input single-output system with reflection pattern modulation (RPM) to achieve PBIT, where the joint active and passive beamforming is carefully designed by taking into account the communication outage probability. We formulate an optimization problem to maximize the *average* received signal power by jointly optimizing the active beamforming at the access point (AP) and passive beamforming at the RIS under the assumption that the RIS's state information is statistically known by the AP, and propose a high-quality suboptimal solution based on the alternating optimization technique. Moreover, a closed-form expression for the asymptotic outage probability of the proposed scheme in Rician fading is derived. The achievable rate of the proposed scheme is also investigated under the assumption that the transmitted symbols are drawn from a finite constellation. Simulation results validate the effectiveness of the proposed scheme and reveal the effect of various system parameters on the achievable rate. It is shown that the proposed scheme outperforms, in terms of achievable rate, the conventional RIS-assisted system without information transfer.**

***Index Terms*—Reconfigurable intelligent surface (RIS), information transfer, metasurface, outage probability, passive beamforming, achievable rate, reflection modulation.**

I. Introduction

TO meet the demanding requirements for the fifth generation (5G) wireless communication in, e.g., enhanced data rate, massive connectivity, low latency, ultra reliability, etc., multiple key technologies including millimeter wave (mmWave) communications, massive multiple-input multiple-output (MIMO) systems, and ultra-dense networks (UDNs) have been extensively investigated in the last decade [1]. However, most of those technologies generally require increased implementation complexity and result in considerably increased energy consumption. Leveraging the recent advances in reconfigurable metasurfaces [2], [3], reconfigurable intelligent surfaces (RISs) (a.k.a. intelligent reflecting surfaces) have emerged as a revolutionary technology for improving the coverage and energy/spectrum efficiency of future wireless communications [4]–[6]. Specifically, RISs are planar metasurfaces consisting of a large number of low-cost unit cell elements, each of which reflects the incident signals according to its reflection state [7]. By configuring the reflection amplitude and/or the phase shift at each unit cell element according to the dynamic wireless channels, the signals reflected by the RIS can be constructively or destructively combined, at the receiver, with the signals from other paths for signal power enhancement or co-channel interference suppression [8]–[10]. Compared to traditional relays, RISs are envisioned to work in a full-duplex mode without being impaired by self-interference and thermal noise, and yet possessing substantially reduced hardware cost and energy consumption owing to their nearly passive components [11]–[13].

To achieve the theoretical passive beamforming gain offered by RIS-assisted systems, the perfect channel state information (CSI) of all the cascaded RIS-aided channels was assumed to be available for optimizing the reflection coefficients of the RIS [14]–[16], which is practically difficult to implement. The acquisition/estimation of the cascaded RIS-aided channels is in practice a challenging task due to the absence of signal processing capability in nearly-passive RISs and due to the large number of unit cell elements [17]–[25]. Existing channel acquisition methods for RIS-aided systems can be classified into two categories. The first category [17] encompasses RISs that are equipped with some receive radio frequency (RF) chains, which enable them to explicitly estimate the channels from the transmitters and the receivers. The second category encompasses RISs that do not explicitly estimate the two links from the transmitter to the RIS and from the RIS to the receiver, but estimate only the cascaded transmitter-RIS-receiver channels, which are explicitly observable at the receiver. In [19] and [20], for example, an element-by-element ON/OFF-based reflection pattern scheme at the RIS is considered, which

The work of M. Wen, F. Chen, and B. Zheng was supported in part by the Natural Science Foundation of Guangdong Province under Grant 2018B030306005, in part by the Pearl River Nova Program of Guangzhou under Grant 201806010171, in part by the open research fund of the National Mobile Communications Research Laboratory, Southeast University under Grant 2020D03, in part by the Fundamental Research Funds for the Central Universities under Grant 2019SJ02, in part by the Guangdong Provincial Key Laboratory of Short-Range Wireless Detection and Communication under Grant 2017B030314003, in part by the National Natural Science Foundation of China under Grant 61671213 and Grant 61671211, and in part by the Project funded by China Postdoctoral Science Foundation under Grant 2018M640781 and Grant 2019T120731. M. Di Renzo's work was supported in part by the European Commission through the H2020 ARIADNE project under grant agreement 675806.

S. Lin and F. Chen are with the School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510641, China (e-mail: eeshe.lin@mail.scut.edu.cn, eefjchen@scut.edu.cn).

B. Zheng is with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore 117583 (e-mail: elezbe@nus.edu.sg).

G. C. Alexandropoulos is with the Department of Informatics and Telecommunications, National and Kapodistrian University of Athens, Panepistimiopolis Ilissia, 15784 Athens, Greece (e-mail: alexandg@di.uoa.gr).

M. Wen is with the School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510640, China, and also with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China (e-mail: eemwwen@scut.edu.cn).

M. Di Renzo is with Université Paris-Saclay, CNRS, CentraleSupélec, Laboratoire des Signaux et Systèmes, 3 Rue Joliot-Curie, 91192 Gif-sur-Yvette, France (marco.direnzo@centralesupelec.fr).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2020.3028198, IEEE Transactions on Wireless Communications

does not require costly RF chains. However, this approach incurs in a prohibitive channel estimation overhead due to the large number of RIS elements that are activated one-by-one to perform channel estimation. To reduce the channel estimation overhead, a grouping strategy for the RIS elements was proposed in [20] and [21], where adjacent RIS elements are grouped together and share a common reflection coefficient. Moreover, a channel estimation method that capitalizes on a carefully designed discrete Fourier transform (DFT)-based reflection pattern was introduced in [21] for application to frequency-selective fading channels. This approach significantly improves the channel estimation accuracy by leveraging the large aperture gain of RISs. In [22], the DFT-based reflection pattern was adopted for channel estimation in RIS-assisted narrowband systems. More recently, some preliminary works that tackle the issue of channel estimation in RIS-assisted multi-user systems have appeared [26], [27].

Prior works on RISs are mainly focused on maximizing the received signal strength of the links between the access point (AP)/base station and the users, e.g., [28]. In order to improve the spectral efficiency of RIS-assisted systems, the authors of [29] introduced RIS-aided space shift keying and RIS-aided spatial modulation schemes to implement the index modulation concept at the receiver side. In addition, the authors of [30] proposed to jointly encode the information in the transmitted signal and in the RIS reflection state for improving the spectral efficiency. In [31], an RIS-assisted system that delivers its own information to the receiver was introduced. In this case, the RIS is equipped with sensors, such as temperature and/or humidity sensors, that need to report the sensed data to an intended receiver. It is worth mentioning that besides the data acquired by sensors, there exist other potential information data that an RIS may need to report to a receiver, including the RIS control signaling and maintenance data [32]. One option for transmitting such kind of information data is to install multiple transmit RF chains on the RIS, which, however, results in increasing the hardware cost and energy consumption and is a solution that is not compatible with the nearly-passive RIS paradigm [10]. Another more cost-effective solution consists of implicitly embedding the data of the RIS into its reflection states. For example, a passive beamforming and information transfer (PBIT) scheme was proposed in [33] to implicitly convey the data sensed by the RIS by implementing spatial modulation on the RIS elements while simultaneously performing fully-passive beamforming to boost the received signal power. This PBIT scheme was further extended in [34] and [35] to the general scenario of multi-user RIS-assisted systems. In the PBIT scheme, however, the number of activated (ON-state) elements varies over time, which can result in significant fluctuations of the reflected signal power, and, consequently, in a relatively high outage probability.

To address this issue, we propose a practical modulation scheme called RIS-based reflection pattern modulation (RIS-RPM) to achieve PBIT in an efficient manner. The core idea of the RIS-RPM scheme is to activate a subset of the RIS elements for reflecting a sharp beam towards the intended destination, while exploiting the index combination of the ON-state RIS elements to implicitly convey the additional information of the RIS. In this paper, we consider an RIS-enhanced multiple-input single-output (MISO) wireless communication system as shown in Fig. 1, where a multi-antenna AP communicates with a single-antenna user with the aid of an RIS. The main contributions of this paper are summarized as follows:

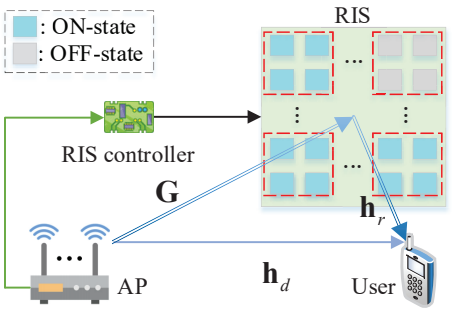


Fig. 1. The considered RIS-enhanced downlink MISO wireless communication system. The RIS adopts reflection pattern modulation (RPM) via the combination of the ON/OFF states of its elements to simultaneously enhance the communication between the AP and receiving user, and to convey its own information data.

- We consider the practical scenario where the ON/OFF state information of the RIS is statistically known by the AP, for which an optimization problem is formulated to maximize the *average* received signal power by jointly optimizing the active beamforming at the AP and passive beamforming at the RIS. The formulated problem, however, is non-convex and thus difficult to solve optimally. As such, we propose an efficient algorithm based on the alternating optimization technique to find a high-quality suboptimal solution. Moreover, we formulate and solve an optimization problem to maximize the *instantaneous* received signal power by designing the active and passive beamforming based on the instantaneous ON/OFF state information of the RIS, which serves as a performance upper bound for the practical beamforming design based on the statistical ON/OFF state information of the RIS.
- The asymptotic outage probability of the proposed RIS-RPM scheme over Rician fading channels is derived in closed-form. It is shown that the RIS is able to significantly increase the diversity gain by properly designing the phase shifts of its reconfigurable elements. Moreover, we analyze the achievable rate of the RIS-RPM scheme under the assumption that the transmitted symbols are drawn from a finite constellation input. Simulation results validate the effectiveness of the proposed RIS-RPM scheme as well as the proposed optimization algorithm, and unveil the impact of various system parameters on the achievable rate. It is shown that the RIS-RPM scheme outperforms the conventional RIS-assisted system with full-ON RIS reflection [15] in terms of achievable rate, despite the loss in the received signal strength. Moreover, the RIS-RPM scheme is shown to have the potential to strike an attractive trade-off between the received signal power and the achievable rate by varying the number of

RIS elements that are activated (ON-state).

The rest of this paper is organized as follows. In Section II, we introduce the system model and the associated channel acquisition. In Sections III and IV, we formulate two optimization problems for designing efficient active and passive beamforming solutions. The performance analysis in terms of outage probability and achievable rate is presented in Section V. Simulation results are presented in Section VI and conclusions are drawn in Section VII.

Notation: Upper and lower case boldface letters denote matrices and column vectors, respectively. $(\cdot)^{\dagger}$, $(\cdot)^{T}$, $(\cdot)^{H}$, and $(\cdot)^{-1}$ represent conjugation, transpose, Hermitian transpose, and inversion operations, respectively. $[\mathbf{X}]_{\iota,\jmath}$ denotes the $(\iota,\jmath)$-th entry of the matrix $\mathbf{X}$, and $[\mathbf{x}]_{\iota}$ denotes the ι-th entry of the vector $\mathbf{x}$. $\mathbf{I}_N$ denotes an $N \times N$ identity matrix, and $\mathbf{1}_N$ and $\mathbf{0}_N$ denote N-dimensional all-one and all-zero column vectors, respectively. $\|\mathbf{x}\|_0$ and $\|\mathbf{x}\|$ denote the zero norm and Euclidean norm of the vector $\mathbf{x}$, respectively. $\mathbb{C}^{n\times m}$ denotes the set of $n \times m$ complex-valued matrices. $\text{diag}(\mathbf{x})$ denotes a diagonal matrix whose diagonal is the vector $\mathbf{x}$. $\text{tr}(\mathbf{X})$ and $\text{rank}(\mathbf{X})$ denote the trace and rank of the matrix $\mathbf{X}$, respectively. $|\cdot|$ and $\angle(\cdot)$ denote the modulus and phase of a complex number, respectively. $p(\cdot)$ denotes the probability density function (PDF). $\lceil\cdot\rceil$ represents the ceiling function that returns the least integer greater than or equal to the argument. $\binom{n}{k}$ denotes the binomial coefficient. $\mathbb{A}\setminus\mathbb{B}$ is the complement set of the set $\mathbb{B}$ with respect to the set $\mathbb{A}$. $\mathcal{CN}(\boldsymbol{\mu},\boldsymbol{\Sigma})$ denotes the distribution of a circularly symmetric complex Gaussian random vector with covariance matrix $\boldsymbol{\Sigma}$ and mean $\boldsymbol{\mu}$. $\mathbb{E}_X\{\cdot\}$ denotes the expectation over the random variable X. $\mathbf{X} \succeq 0$ denotes that a Hermitian matrix $\mathbf{X}$ is positive semi-definite.

II. System Description and Channel Estimation

As illustrated in Fig. 1, we consider a MISO wireless communication system, where an RIS composed of L unit cell elements is deployed to enhance the communication link from an AP equipped with N transmit antennas to a single-antenna user. An RIS controller is attached to the RIS, which is responsible for reconfiguring the phase shifts and/or reflection amplitudes of the unit cell elements [36] and exchanging information with the AP for the realization of the simultaneous passive beamforming and reflection modulation. We assume that the RIS controller is equipped with a number of sensors in order to monitor/collect environmental data, which is typically low-rate bursty data for the user. In this paper, we consider quasi-static block fading channels for the AP-user, AP-RIS, and RIS-user links, where all the channels remain unchanged over the coherence block equal to the duration of T_c symbol sampling periods, and are independent and identically distributed (i.i.d.) across different coherence blocks. This assumption is valid since RISs are practically used to support low-mobility users in their vicinity.

A. System Model

For each symbol duration, only K out of L $(K \leq L)$ unit cell elements are turned ON to reflect the incident signals so that the indices of the ON-state elements can implicitly convey the RIS information. With the proposed RPM scheme, there are in total $\binom{L}{K}$ combinations for determining the indices of K ON-state elements at the RIS. Therefore, $\log_2\binom{L}{K}$ information bits can be implicitly conveyed. An example of the proposed RPM scheme with an RIS of $L = 4$ unit cell elements ($K = 3$ ON-state elements) is illustrated in Table I. However, the number of unit cell elements at the RIS can be practically large, leading to an enormous calculation and storage overhead for the mapping between information bits and ON-state elements as the number of K-combinations increases exponentially with L. To address this issue, we adopt the RIS-elements grouping method, where a total number of L unit cell elements are divided into G groups, each of which consists of $\bar{L} = L/G$ adjacent elements sharing a common reflection coefficient. For notational convenience, we assume that $\bar{L}$ is an integer that we refer to as the grouping size. As such, the information of the RIS is conveyed through the indices of the ON-state groups. Specifically, for each symbol duration, $\bar{K}$ out of G $(\bar{K} \leq G)$ groups are randomly turned ON for reflecting the incident signals, and the remaining $(G - \bar{K})$ groups are deliberately turned OFF for realizing the proposed RPM scheme. Therefore, a total number of $\log_2\binom{G}{\bar{K}}$ information bits can be conveyed. Let $\mathbb{I} \triangleq \{i_1, i_2, \ldots, i_{\bar{K}}\}$ denote the indices of the $\bar{K}$ ON-state groups, where $i_k \in \mathbb{G} \triangleq \{1, 2, \ldots, G\}$ for $k = 1, 2, \ldots, \bar{K}$ with $i_1 < i_2 < \ldots < i_{\bar{K}}$. Given the indices of the ON-state groups, the common reflection coefficient for the g-th group can be characterized by

$$\theta_g = \begin{cases} \beta_g e^{-j\phi_g}, & g \in \mathbb{I} \\ 0, & g \in \mathbb{G}\setminus\mathbb{I} \end{cases} \qquad g = 1, 2, \ldots, G \quad (1)$$

where $\beta_g \in [0, 1]$ and $\phi_g \in (0, 2\pi]$ represent the common reflection amplitude and phase shift for the g-th group, respectively. To enhance the reflected signal power and ease the hardware design, the reflection amplitudes of the ON-state groups are set to the maximum value, i.e., $\beta_g = 1$, $\forall g \in \mathbb{I}$. Let $\boldsymbol{\theta} \triangleq [\theta_1, \theta_2, \ldots, \theta_G]^T$ denote the RIS reflection vector after grouping, which characterizes the equivalent interaction of the RIS with the incident signals. In particular, $\|\boldsymbol{\theta}\|_0 = \bar{K}$.

Let $\mathbf{G} \triangleq [\boldsymbol{\Delta}_1, \boldsymbol{\Delta}_2, \ldots, \boldsymbol{\Delta}_G]^H \in \mathbb{C}^{L\times N}$, $\mathbf{h}_r^H \triangleq [\mathbf{r}_1^H, \mathbf{r}_2^H, \ldots, \mathbf{r}_G^H] \in \mathbb{C}^{1\times L}$, and $\mathbf{h}_d^H \in \mathbb{C}^{1\times N}$ denote the baseband channels from the AP to RIS, from the RIS to user, and from the AP to user, respectively, where $\boldsymbol{\Delta}_g^H \in \mathbb{C}^{\bar{L}\times N}$ and $\mathbf{r}_g^H \in \mathbb{C}^{1\times\bar{L}}$ denote the baseband element-wise channels from the AP to the g-th group and from the g-th group to the user, respectively. Moreover, due to the severe path loss and high attenuation, the signals reflected by the RIS more

TABLE I
An Example of the Proposed RPM Scheme with an RIS of $L = 4$ Unit Cell Elements ($K = 3$ ON-state Elements). The Index of the OFF-state Element is Used to Implicitly Convey the Information Bits of the RIS.

Information bits of the RIS	00	01	10	11
Index of the OFF-state element	{4}	{3}	{2}	{1}
Reflection state of the RIS				

than once have negligible power and hence can be ignored. Accordingly, the received signal at the user is given by

$$y = \sqrt{P_t}\left(\sum_{g=1}^{G} \mathbf{r}_g^H \theta_g \mathbf{\Delta}_g^H + \mathbf{h}_d^H\right)\mathbf{w}x + n \tag{2}$$

where P_t is the maximum transmit power of the AP, $\mathbf{w} \in \mathbb{C}^{N\times 1}$ stands for the active beamforming at the AP, x is the transmitted symbol, which is drawn from an M-ary constellation $\mathbb{A}$ with normalized power, θ_g is the common reflection coefficient for the g-th group, and $n \sim \mathcal{CN}(0, \sigma^2)$ is the additive white Gaussian noise (AWGN) with σ^2 being the noise power. Since the RIS elements in the same group share a common reflection coefficient, by denoting with $\bar{\mathbf{h}}_g^H = \mathbf{r}_g^H \mathbf{\Delta}_g^H$ the equivalent cascaded channel of the AP-RIS-user link via the g-th group with zero phase shift, (2) can be rewritten as

$$y = \sqrt{P_t}\left(\boldsymbol{\theta}^T \mathbf{H} + \mathbf{h}_d^H\right)\mathbf{w}x + n \tag{3}$$

where $\mathbf{H} = \left[\bar{\mathbf{h}}_1, \bar{\mathbf{h}}_2, \ldots, \bar{\mathbf{h}}_G\right]^H \in \mathbb{C}^{G\times N}$ denotes the cascaded AP-RIS-user channel matrix without the effect of the RIS reflection. From (3), it can be observed that the information delivered to the user consists of two parts: the first part is from the AP *explicitly* expressed as x and the second part is from the RIS *implicitly* embedded in $\boldsymbol{\theta}$.

B. Channel Estimation

To boost the received signal power by jointly optimizing the active beamforming at the AP and passive beamforming at the RIS, the knowledge of $\mathbf{H}$ and $\mathbf{h}_d$ is required. By assuming channel reciprocity and using a time-division duplex (TDD) protocol, the CSI of $\mathbf{H}$ and $\mathbf{h}_d$ can be obtained at the AP. Specifically, during the channel training, $(G+1)$ consecutive pilot symbols are sent by the user. For the duration of the i-th $(i = 0, 1, \ldots, G)$ pilot symbol, the received pilot signal vector at the AP can be expressed as

$$\mathbf{y}_p^{(i)} = \sqrt{P_p}\left(\sum_{g=1}^{G} \mathbf{\Delta}_g \psi_g^{(i)} \mathbf{r}_g + \mathbf{h}_d\right) x_p^{(i)} + \mathbf{n}^{(i)} \tag{4}$$

where P_p is the transmit power of the user, $x_p^{(i)}$ is the i-th pilot symbol with normalized power, $\psi_g^{(i)}$ is the common phase shift for the g-th group during the transmission of the i-th pilot symbol, and $\mathbf{n}^{(i)} \sim \mathcal{CN}(\mathbf{0}_N, \sigma^2 \mathbf{I}_N)$ is the AWGN vector. We note that $|\psi_g^{(i)}| = 1$. By letting $\boldsymbol{\psi}^{(i)} \triangleq \left[\psi_1^{(i)}, \psi_2^{(i)}, \ldots, \psi_G^{(i)}\right]^T$ denote the RIS phase-shift state during the transmission of the i-th pilot symbol, (4) can be rewritten as

$$\begin{aligned}\mathbf{y}_p^{(i)} =& \sqrt{P_p}\left(\mathbf{H}^H \boldsymbol{\psi}^{(i)} + \mathbf{h}_d\right) x_p^{(i)} + \mathbf{n}^{(i)} \\ =& \sqrt{P_p}\tilde{\mathbf{H}}\tilde{\boldsymbol{\psi}}^{(i)} x_p^{(i)} + \mathbf{n}^{(i)}\end{aligned} \tag{5}$$

where $\tilde{\mathbf{H}} = \begin{bmatrix}\mathbf{h}_d & \mathbf{H}^H\end{bmatrix} \in \mathbb{C}^{N\times(G+1)}$ and $\tilde{\boldsymbol{\psi}}^{(i)} = \begin{bmatrix}1 \\ \boldsymbol{\psi}^{(i)}\end{bmatrix}$. By stacking $(G+1)$ consecutive received pilot signals, we have

$$\begin{aligned}\mathbf{Y}_p =& \left[\mathbf{y}_p^{(0)}, \mathbf{y}_p^{(1)}, \ldots, \mathbf{y}_p^{(G)}\right] \\ =& \sqrt{P_p}\tilde{\mathbf{H}}\mathbf{\Psi}\mathrm{diag}\left(\mathbf{x}_p\right) + \mathbf{\Omega}\end{aligned} \tag{6}$$

where $\mathbf{\Psi} = \left[\tilde{\boldsymbol{\psi}}^{(0)}, \tilde{\boldsymbol{\psi}}^{(1)}, \ldots, \tilde{\boldsymbol{\psi}}^{(G)}\right]$ is the RIS phase-shift pattern, $\mathbf{x}_p = \left[x_p^{(0)}, x_p^{(1)}, \ldots, x_p^{(G)}\right]^T$ denotes the pilot sequence, and $\mathbf{\Omega} = \left[\mathbf{n}^{(0)}, \mathbf{n}^{(1)}, \ldots, \mathbf{n}^{(G)}\right]$ is the AWGN matrix. By using the DFT-based phase-shift pattern [22], the CSI of $\mathbf{H}$ and $\mathbf{h}_d$ can be estimated as

$$\hat{\tilde{\mathbf{H}}} = \begin{bmatrix}\hat{\mathbf{h}}_d & \hat{\mathbf{H}}^H\end{bmatrix} = \frac{1}{\sqrt{P_p}(G+1)}\mathbf{Y}_p \mathrm{diag}\left(\mathbf{x}_p\right)^{-1}\mathbf{F}_{G+1}^H \tag{7}$$

where $\left[\mathbf{F}_{G+1}\right]_{\iota,\jmath} = e^{-j\frac{2\pi\iota\jmath}{G+1}}$, $\iota, \jmath = 0, 1, \ldots, G$.

After acquiring the CSI of $\mathbf{H}$ and $\mathbf{h}_d$, the AP jointly optimizes the active beamforming vector $\mathbf{w}$ and the phase shifts of all the groups to boost the received signal power, and then reports the optimized phase shifts to be implemented to the RIS controller. In the following section, we formulate an optimization problem to maximize the *average* received signal power by jointly designing the active beamforming at the AP and the passive beamforming at the RIS based on the statistical ON/OFF state information of the RIS.

III. Beamforming Design Based on Statistical ON/OFF State Information

A. Problem Formulation

We consider the practical assumption that the ON/OFF state information of the RIS is statistically known by the AP. Our objective is to minimize the outage probability of the combined AP-user channel by jointly optimizing the active beamforming at the AP and the passive beamforming at the RIS under some constraints for the maximum transmit power of the AP and for the unit-modulus reflection of the RIS. Specifically, given the indices of the ON-state groups $\mathbb{I}$, let $\mathbf{s} \triangleq [s_1, s_2, \ldots, s_G]^T$ denote the ON/OFF states of all groups of the RIS in a vector form, with each entry given by

$$s_g = \begin{cases}1, & g \in \mathbb{I} \\ 0, & g \in \mathbb{G}\setminus\mathbb{I}\end{cases} \quad g = 1, 2, \ldots, G \tag{8}$$

where $s_g = 1$ and $s_g = 0$ represent the ON and OFF states, respectively. Based on (8), the RIS reflection vector $\boldsymbol{\theta}$ can be rewritten as $\boldsymbol{\theta} = \mathbf{\Phi}\mathbf{s}$, in which $\mathbf{\Phi} = \mathrm{diag}\left(\boldsymbol{\varphi}\right)$ is the diagonal phase-shift matrix of the RIS with the g-th diagonal entry given by $\varphi_g = e^{-j\phi_g}$, $\phi_g \in (0, 2\pi]$, $g = 1, 2, \ldots, G$. The combined channel from the AP to user is given as $\boldsymbol{\theta}^T\mathbf{H} + \mathbf{h}_d^H = \mathbf{s}^T\mathbf{\Phi}\mathbf{H} + \mathbf{h}_d^H$, and thus the achievable rate of the combined AP-user channel conditioned on $\mathbf{H}$ and $\mathbf{h}_d$ can be expressed as $\log_2\left(1 + \frac{P_t}{\sigma^2}\left|\left(\mathbf{s}^T\mathbf{\Phi}\mathbf{H} + \mathbf{h}_d^H\right)\mathbf{w}\right|^2\right)$ bits per second per Hertz (bps/Hz). The corresponding outage probability for a target rate R is given by

$$\begin{aligned}p_{\text{out}}(R) &= \mathbb{P}\left\{\log_2\left(1 + \frac{P_t}{\sigma^2}\left|\left(\mathbf{s}^T\mathbf{\Phi}\mathbf{H} + \mathbf{h}_d^H\right)\mathbf{w}\right|^2\right) < R\right\} \\ &= \mathbb{P}\left\{P_t\left|\left(\mathbf{s}^T\mathbf{\Phi}\mathbf{H} + \mathbf{h}_d^H\right)\mathbf{w}\right|^2 < (2^R - 1)\sigma^2\right\}.\end{aligned} \tag{9}$$

The optimization problem for minimizing the outage probability in (9) under the constraints of maximum transmit power at the AP and the unit-modulus reflection at the RIS can be formulated as follows.

$$\text{(P0)}: \min_{\mathbf{w}, \mathbf{\Phi}} \; t \tag{10}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2020.3028198, IEEE Transactions on Wireless Communications

$$\text{s.t. } \mathbb{P}\left\{P_t\left|\left(\mathbf{s}^T\mathbf{\Phi}\mathbf{H}+\mathbf{h}_d^H\right)\mathbf{w}\right|^2<(2^R-1)\sigma^2\right\}\le t \tag{11}$$

$$\|\mathbf{w}\|^2\le 1 \tag{12}$$

$$|\varphi_g|=1,\quad g=1,2,\ldots,G. \tag{13}$$

However, this problem is non-convex and difficult to solve due to the outage probability and unit-modulus constraints. On the other hand, it can be observed from (9) that, given any ON/OFF-state vector $\mathbf{s}$, we need to optimize the active beamforming at the AP and passive beamforming at the RIS to maximize the received signal power for minimizing the outage probability. Since the instantaneous ON/OFF state information of the RIS is unavailable to the AP, however, we instead optimize the active beamforming and passive beamforming to maximize the *average* received signal power. We note that this approach is effective in improving the performance in terms of outage probability, as it will be shown in the simulation results. Let $\mathbf{A}=\mathbb{E}_{\mathbf{s}}\left\{\mathbf{s}\mathbf{s}^T\right\}$ and $\mathbf{a}=\mathbb{E}_{\mathbf{s}}\left\{\mathbf{s}\right\}$ denote the covariance matrix and mean vector of $\mathbf{s}$, respectively. Based on (8) and assuming that all possible index combinations are equiprobable, the elements of $\mathbf{A}$ and $\mathbf{a}$ can be derived as follows.

$$[\mathbf{A}]_{i,j}=\begin{cases}\dfrac{\bar{K}}{G}, & i=j\\ \dfrac{\bar{K}(\bar{K}-1)}{G(G-1)}, & i\neq j\end{cases}\qquad i,j=1,2,\ldots,G \tag{14}$$

and

$$[\mathbf{a}]_i=\frac{\bar{K}}{G},\quad i=1,2,\ldots,G. \tag{15}$$

Thus, the average received signal power normalized by the transmit power is given by

$$\begin{aligned}&\mathbb{E}_{\mathbf{s}}\left\{\left|\left(\mathbf{s}^T\mathbf{\Phi}\mathbf{H}+\mathbf{h}_d^H\right)\mathbf{w}\right|^2\right\}\\ =&\mathbf{w}^H\left(\mathbf{H}^H\mathbf{\Phi}^H\mathbb{E}_{\mathbf{s}}\{\mathbf{s}\mathbf{s}^T\}\mathbf{\Phi}\mathbf{H}+\mathbf{H}^H\mathbf{\Phi}^H\mathbb{E}_{\mathbf{s}}\{\mathbf{s}\}\mathbf{h}_d^H\right.\\ &\left.+\mathbf{h}_d\mathbb{E}_{\mathbf{s}}\{\mathbf{s}^T\}\mathbf{\Phi}\mathbf{H}+\mathbf{h}_d\mathbf{h}_d^H\right)\mathbf{w}\\ =&\mathbf{w}^H\begin{bmatrix}\mathbf{\Phi}\mathbf{H}\\ \mathbf{h}_d^H\end{bmatrix}^H\underbrace{\begin{bmatrix}\mathbf{A} & \mathbf{a}\\ \mathbf{a}^T & 1\end{bmatrix}}_{\tilde{\mathbf{A}}}\begin{bmatrix}\mathbf{\Phi}\mathbf{H}\\ \mathbf{h}_d^H\end{bmatrix}\mathbf{w}.\end{aligned} \tag{16}$$

Accordingly, the corresponding optimization problem based on the estimated CSI of $\hat{\mathbf{H}}$ and $\hat{\mathbf{h}}_d$ is formulated as follows.

$$\text{(P1):}\quad\max_{\mathbf{w},\mathbf{\Phi}}\quad\mathbf{w}^H\begin{bmatrix}\mathbf{\Phi}\hat{\mathbf{H}}\\ \hat{\mathbf{h}}_d^H\end{bmatrix}^H\tilde{\mathbf{A}}\begin{bmatrix}\mathbf{\Phi}\hat{\mathbf{H}}\\ \hat{\mathbf{h}}_d^H\end{bmatrix}\mathbf{w} \tag{17}$$

$$\text{s.t.}\quad\|\mathbf{w}\|^2\le 1 \tag{18}$$

$$|\varphi_g|=1,\quad g=1,2,\ldots,G. \tag{19}$$

It can be readily verified that problem (P1) is a non-convex problem as well, since the objective function of (17) is non-concave with respect to both $\mathbf{w}$ and $\mathbf{\Phi}$, and the constraint in (19) is non-convex. Moreover, due to the mutual coupling between $\mathbf{w}$ and $\mathbf{\Phi}$ in the objective function of (17), problem (P1) becomes even more difficult to solve. To circumvent the above difficulties, we develop an alternating optimization algorithm to find a high-quality suboptimal solution to problem (P1), which iteratively optimizes $\mathbf{w}$ and $\mathbf{\Phi}$.

B. Joint Beamforming Design

For any given RIS phase-shift matrix $\mathbf{\Phi}$, problem (P1) can be rewritten as

$$\text{(P1.1):}\quad\max_{\mathbf{w}}\quad\mathbf{w}^H\tilde{\mathbf{R}}\mathbf{w} \tag{20}$$

$$\text{s.t.}\quad\|\mathbf{w}\|^2\le 1 \tag{21}$$

where

$$\tilde{\mathbf{R}}=\begin{bmatrix}\mathbf{\Phi}\hat{\mathbf{H}}\\ \hat{\mathbf{h}}_d^H\end{bmatrix}^H\tilde{\mathbf{A}}\begin{bmatrix}\mathbf{\Phi}\hat{\mathbf{H}}\\ \hat{\mathbf{h}}_d^H\end{bmatrix}. \tag{22}$$

Since $\tilde{\mathbf{R}}$ is Hermitian, for any non-zero $\mathbf{w}$, we have the following inequality

$$\mathbf{w}^H\tilde{\mathbf{R}}\mathbf{w}\le\lambda_{\max}(\tilde{\mathbf{R}})\|\mathbf{w}\|^2 \tag{23}$$

where $\lambda_{\max}(\tilde{\mathbf{R}})$ denotes the maximum eigenvalue of the matrix $\tilde{\mathbf{R}}$. Let $\mathbf{v}_{\max}$ denote the eigenvector corresponding to the maximum eigenvalue of the matrix $\tilde{\mathbf{R}}$. Then, it can be readily verified that the optimal solution to problem (P1.1) is given by $\mathbf{w}=\mathbf{v}_{\max}/\|\mathbf{v}_{\max}\|$.

Next, we optimize $\boldsymbol{\varphi}$ based on the optimal active beamforming vector to problem (P1.1). Specifically, given $\mathbf{w}$, by defining $\mathbf{\Lambda}=\text{diag}\left(\hat{\mathbf{H}}\mathbf{w}\right)$ and $g_d=\hat{\mathbf{h}}_d^H\mathbf{w}$, problem (P1) can be rewritten as (irrelevant terms are omitted for simplicity)

$$\text{(P1.2):}\quad\max_{\boldsymbol{\varphi}}\quad\boldsymbol{\varphi}^H\mathbf{\Lambda}^H\mathbf{A}\mathbf{\Lambda}\boldsymbol{\varphi}+g_d\boldsymbol{\varphi}^H\mathbf{\Lambda}^H\mathbf{a}+g_d^H\mathbf{a}^T\mathbf{\Lambda}\boldsymbol{\varphi} \tag{24}$$

$$\text{s.t.}\quad|\varphi_g|=1,\quad g=1,2,\ldots,G. \tag{25}$$

From (24) and (25), we observe that problem (P1.2) is a non-convex quadratically constrained quadratic program (QCQP), which can be reformulated as a homogeneous QCQP by introducing an auxiliary variable t [37], i.e.,

$$\text{(P1.3):}\quad\max_{\tilde{\boldsymbol{\varphi}}}\quad\tilde{\boldsymbol{\varphi}}^H\mathbf{\Xi}\tilde{\boldsymbol{\varphi}} \tag{26}$$

$$\text{s.t.}\quad|\varphi_g|=1,\quad g=1,2,\ldots,G \tag{27}$$

where

$$\mathbf{\Xi}=\begin{bmatrix}\mathbf{\Lambda}^H\mathbf{A}\mathbf{\Lambda} & g_d\mathbf{\Lambda}^H\mathbf{a}\\ g_d^H\mathbf{a}^T\mathbf{\Lambda} & 0\end{bmatrix},\quad\tilde{\boldsymbol{\varphi}}=\begin{bmatrix}\boldsymbol{\varphi}\\ t\end{bmatrix}. \tag{28}$$

The objective function of (26) can be rewritten as $\tilde{\boldsymbol{\varphi}}^H\mathbf{\Xi}\tilde{\boldsymbol{\varphi}}=\text{tr}\left(\mathbf{\Xi}\mathbf{Q}\right)$ with $\mathbf{Q}=\tilde{\boldsymbol{\varphi}}\tilde{\boldsymbol{\varphi}}^H$. In particular, $\mathbf{Q}$ is a positive semidefinite matrix with rank $(\mathbf{Q})=1$. However, as the rank-one constraint is non-convex, we apply the semi-definite relaxation (SDR) method to relax this constraint and reformulate (P1.3) as

$$\text{(P1.4):}\quad\max_{\mathbf{Q}}\quad\text{tr}\left(\mathbf{\Xi}\mathbf{Q}\right) \tag{29}$$

$$\text{s.t.}\quad[\mathbf{Q}]_{g,g}=1,\ g=1,2,\ldots,G+1 \tag{30}$$

$$\mathbf{Q}\succeq 0 \tag{31}$$

which is a standard convex semi-definite programming (SDP) problem and can be solved via existing convex optimization solvers such as CVX [38]. It is worth pointing out that after the relaxation, the optimal solution $\mathbf{Q}^*$ to problem (P1.4) may

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2020.3028198, IEEE Transactions on Wireless Communications

Algorithm 1 Alternating Optimization Algorithm for Solving Problem (P1)

Input: $\hat{\mathbf{H}}$, $\hat{\mathbf{h}}_d$, threshold ϵ, and the maximum iteration number I

1: Initialize the diagonal phase-shift matrix $\mathbf{\Phi}^{(1)} := \mathbf{I}_G$ and set the iteration number $n := 1$
2: **repeat**
3: Substitute $\mathbf{\Phi}^{(n)}$ into (22) to get $\tilde{\mathbf{R}}^{(n)}$, then find the eigenvector corresponding to the maximum eigenvalue of $\tilde{\mathbf{R}}^{(n)}$ to obtain the active beamforming $\mathbf{w}^{(n)}$
4: Given $\mathbf{w}^{(n)}$, solve problem (P1.4) via a convex optimization solver and the Gaussian randomization of (32) to obtain the RIS phase-shift matrix $\mathbf{\Phi}^{(n+1)}$
5: Update $n := n + 1$
6: **until** The fractional increase of (17) is less than ϵ or $n > I$

Output: $\mathbf{w}^*$ and $\mathbf{\Phi}^*$

not be a rank-one solution. Therefore, we retrieve $\tilde{\boldsymbol{\varphi}}^*$ from $\mathbf{Q}^*$ as follows.

$$\tilde{\boldsymbol{\varphi}}^* = \begin{cases} \mathbf{U}\mathbf{D}^{1/2}\mathbf{1}_{G+1}, & \text{rank}\,(\mathbf{Q}^*) = 1 \\ \mathbf{U}\mathbf{D}^{1/2}\boldsymbol{\varkappa}, & \text{rank}\,(\mathbf{Q}^*) \neq 1 \end{cases} \tag{32}$$

where $\mathbf{Q}^* = \mathbf{U}\mathbf{D}\mathbf{U}^H$ is the eigenvalue decomposition of $\mathbf{Q}^*$ and $\boldsymbol{\varkappa} \sim \mathcal{CN}(\mathbf{0}_{G+1}, \mathbf{I}_{G+1})$ is a random vector. Finally, the suboptimal solution to problem (P1.2) is given by

$$\varphi_g^* = \frac{[\tilde{\boldsymbol{\varphi}}^*]_g \Big/ [\tilde{\boldsymbol{\varphi}}^*]_{G+1}}{\left|[\tilde{\boldsymbol{\varphi}}^*]_g \Big/ [\tilde{\boldsymbol{\varphi}}^*]_{G+1}\right|}, \quad g = 1, 2, \ldots, G. \tag{33}$$

The algorithm proceeds by iteratively solving subproblems (P1.1) and (P1.4) in an alternating manner until the convergence criterion is met, i.e., the fractional increase of (17) is less than a small positive number ϵ, or the maximum number of iterations is reached. Algorithm 1 summarizes this iterative procedure. The convergence of the proposed algorithm can be guaranteed by the fact that the objective value of problem (P1) is non-decreasing over the iterations and it is upper-bounded by a finite value due to the limited transmit power. As far as the computational complexity is concerned, problem (P1.1) involving the eigenvalue decomposition of an $N \times N$ matrix can be solved with a complexity of $\mathcal{O}(N^3)$, and the SDP problem (P1.4) can be solved with a worst-case complexity of $\mathcal{O}((G+1)^{4.5})$ [39]. Given the number of iterations, denoted by I, the total complexity for solving problem (P1) is thus given by $\mathcal{O}\left((N^3 + (G+1)^{4.5})I\right)$. After obtaining the optimized phase-shift matrix $\mathbf{\Phi}^*$, the RIS performs passive beamforming with $\mathbf{\Phi}_{\mathbb{I}}^*$ only (while the RIS elements belonging to the remaining $(G - \bar{K})$ groups are turned OFF) according to the selection of $\mathbb{I}$ by the RIS.

IV. Beamforming Design Based on Instantaneous ON/OFF State Information

In this section, we characterize the upper bound of the received signal power, which serves as a benchmark for the beamforming design based on the statistical ON/OFF state information of the RIS. We assume that the AP knows the ON/OFF state information of the RIS exactly. It is worth mentioning that this assumption is similar to that in [30]. Under this assumption, we formulate an optimization problem to maximize the *instantaneous* received signal power by jointly designing the active beamforming at the AP and the passive beamforming at the RIS based on the instantaneous ON/OFF state information of the RIS.

Given the indices of the ON-state groups $\mathbb{I}$, the corresponding optimization problem based on the estimated CSI can be formulated as follows (P_t is omitted for brevity).

$$\text{(P2)}: \max_{\mathbf{w},\boldsymbol{\theta}} \quad \left|\left(\boldsymbol{\theta}^T\hat{\mathbf{H}} + \hat{\mathbf{h}}_d^H\right)\mathbf{w}\right|^2 \tag{34}$$

$$\text{s.t.} \quad \|\mathbf{w}\|^2 \leq 1, \tag{35}$$

$$|\theta_g| = 1, \quad \forall\, g \in \mathbb{I}, \tag{36}$$

$$\theta_g = 0, \quad \forall\, g \in \mathbb{G} \setminus \mathbb{I}. \tag{37}$$

By eliminating the constraint in (37), (P2) is equivalent to

$$\text{(P2.1)}: \max_{\mathbf{w},\boldsymbol{\theta}_{\mathbb{I}}} \quad \left|\left(\boldsymbol{\theta}_{\mathbb{I}}^T\hat{\mathbf{H}}_{\mathbb{I}} + \hat{\mathbf{h}}_d^H\right)\mathbf{w}\right|^2 \tag{38}$$

$$\text{s.t.} \quad \|\mathbf{w}\|^2 \leq 1, \tag{39}$$

$$|\theta_{i_k}| = 1,\ i_k \in \mathbb{I},\ k = 1, 2, \ldots, \bar{K} \tag{40}$$

where $\boldsymbol{\theta}_{\mathbb{I}}$ is the subvector consisting of the $\bar{K}$ entries of $\boldsymbol{\theta}$ indexed by $\mathbb{I}$, and $\hat{\mathbf{H}}_{\mathbb{I}}$ is the submatrix consisting of the $\bar{K}$ rows of $\hat{\mathbf{H}}$ indexed by $\mathbb{I}$. It can be readily verified that problem (P2.1) is non-convex as well, since the objective function of (38) is non-concave with respect to both $\mathbf{w}$ and $\boldsymbol{\theta}_{\mathbb{I}}$, as well as the constraint in (40) is not convex. However, problem (P2.1) can be solved suboptimally by leveraging the alternating optimization technique similarly to Algorithm 1. Specifically, based on the alternating optimization technique, at each iteration, one of $\mathbf{w}$ and $\boldsymbol{\theta}_{\mathbb{I}}$ is optimized while the other is kept fixed. In detail, the algorithmic procedure can be summarized as follows. For any given active beamforming vector $\mathbf{w}$, problem (P2.1) can be reformulated as follows.

$$\text{(P2.2)}: \max_{\boldsymbol{\theta}_{\mathbb{I}}} \quad \left|\sum_{k=1}^{\bar{K}} \theta_{i_k}\hat{\tilde{\mathbf{h}}}_{i_k}^H\mathbf{w} + \hat{\mathbf{h}}_d^H\mathbf{w}\right|^2 \tag{41}$$

$$\text{s.t.} \quad |\theta_{i_k}| = 1,\ i_k \in \mathbb{I},\ k = 1, 2, \ldots, \bar{K}. \tag{42}$$

By exploiting the triangle inequality, we have

$$\left|\sum_{k=1}^{\bar{K}} \theta_{i_k}\hat{\tilde{\mathbf{h}}}_{i_k}^H\mathbf{w} + \hat{\mathbf{h}}_d^H\mathbf{w}\right| \leq \sum_{k=1}^{\bar{K}} \left|\theta_{i_k}\hat{\tilde{\mathbf{h}}}_{i_k}^H\mathbf{w}\right| + \left|\hat{\mathbf{h}}_d^H\mathbf{w}\right| \tag{43}$$

with equality if and only if $\angle\left(\theta_{i_k}\hat{\tilde{\mathbf{h}}}_{i_k}^H\mathbf{w}\right) = \angle\left(\hat{\mathbf{h}}_d^H\mathbf{w}\right)$ for all $i_k \in \mathbb{I}$. Therefore, the optimal phase shift for the k-th ($k = 1, 2, \ldots, \bar{K}$) ON-state group is given by

$$\phi_{i_k} = \angle\left(\hat{\tilde{\mathbf{h}}}_{i_k}^H\mathbf{w}\right) - \angle\left(\hat{\mathbf{h}}_d^H\mathbf{w}\right), \quad i_k \in \mathbb{I}. \tag{44}$$

For a given $\boldsymbol{\theta}_{\mathbb{I}}$ in (44), it can be readily obtained that the optimal active beamforming vector is given by

$$\mathbf{w}_{\text{MRT}} \triangleq \frac{1}{\left\|\boldsymbol{\theta}_{\mathbb{I}}^T\hat{\mathbf{H}}_{\mathbb{I}} + \hat{\mathbf{h}}_d^H\right\|}\left(\boldsymbol{\theta}_{\mathbb{I}}^T\hat{\mathbf{H}}_{\mathbb{I}} + \hat{\mathbf{h}}_d^H\right)^H \tag{45}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2020.3028198, IEEE Transactions on Wireless Communications

which is the well-known maximum-ratio transmission (MRT). Therefore, $\boldsymbol{\theta}_{\mathbb{I}}$ and $\mathbf{w}$ are optimized according to (44) and (45) in an alternating manner until the convergence criterion is met. We note that the proposed iterative algorithm is guaranteed to converge since the objective value of problem (P2) is non-decreasing over the iterations and the optimal objective value of problem (P2) is finite.

V. Performance Analysis

In this section, we investigate the proposed RIS-RPM scheme in terms of outage probability and achievable rate.

A. Outage Probability

For ease of exposition, we assume $N = 1$ with $\mathbf{G} \equiv \mathbf{g}$ and $\mathbf{h}_d^H \equiv h_d^\dagger$, such that the active beamforming vector $\mathbf{w}$ can be dropped. Moreover, we consider the Rician fading channel model for all the channels involved, where each channel coefficient is the superposition of a determined line-of-sight (LoS) component and a non-LoS component characterized by a complex Gaussian random variable. Let κ_{AR}, κ_{Ru}, and κ_{Au} denote the Rician factors of the AP-RIS, RIS-user, and AP-user links, respectively. In particular, the RIS is generally installed on the walls/ceilings to establish a LoS link with the AP to boost the signal strength in its vicinity, while the user is usually in a relatively rich scattering environment. Therefore, we assume $\kappa_{\text{AR}} = \infty$, $\kappa_{\text{Ru}} = 0$, and $\kappa_{\text{Au}} = 0$, so that the AP-RIS channel has only a fixed LoS component while the AP-user and RIS-user channels can be well characterized by Rayleigh fading. Let σ_g^2 denote the power of the LoS component of the AP-RIS channel and assume $\mathbf{h}_r^H \sim \mathcal{CN}(\mathbf{0}_L, \sigma_h^2 \mathbf{I}_L)$ and $h_d^\dagger \sim \mathcal{CN}(0, \sigma_d^2)$. Thus, we have $\mathbf{H} \equiv \mathbf{h} \sim \mathcal{CN}(\mathbf{0}_G, \sigma_r^2 \mathbf{I}_G)$ with $\sigma_r^2 = \bar{L}\sigma_h^2\sigma_g^2$. For ease of notation, we assume $\sigma_d^2 = \sigma_r^2 = 1$. Moreover, we resort to a unified definition of signal-to-noise ratio (SNR) as $\gamma = P_t/\sigma^2$ to draw essential insights. Accordingly, the outage probability in (9) can be simplified as

$$p_{\text{out}}(R) = \mathbb{P}\left\{\left|\mathbf{s}^T\boldsymbol{\Phi}\mathbf{h} + h_d^\dagger\right|^2 < \frac{2^R - 1}{\gamma}\right\}. \tag{46}$$

For ease of writing, we use the notation $\phi_0 = \angle(h_d^\dagger)$ and $\boldsymbol{\chi} \triangleq [\chi_0, \chi_1, \ldots, \chi_G]^T = \begin{bmatrix} h_d^\dagger \\ \boldsymbol{\Phi}\mathbf{h} \end{bmatrix}$. First, we assume perfect CSI available at the AP for setting the phase shifts of the RIS elements. The optimal phase shifts for the case of $N = 1$ correspond to perfectly co-phasing the signals reflected by the RIS elements with the signal of the direct link to the intended user, regardless of the ON/OFF-state information of the RIS. Thus, we have $\boldsymbol{\chi} = e^{j\phi_0}\left[|\chi_0|, |\chi_1|, \ldots, |\chi_G|\right]^T$, where $\{|\chi_g|\}_{g=0}^G$ are i.i.d. Rayleigh random variables with parameter $\sqrt{1/2}$. We define $X = \left|\mathbf{s}^T\boldsymbol{\Phi}\mathbf{h} + h_d^\dagger\right|^2$, which is the square of the sum of $(\bar{K}+1)$ i.i.d. Rayleigh random variables. In particular, X has a Gamma distribution with parameters

$$k_X = \frac{\mathbb{E}\{X\}^2}{\mathbb{E}\{X^2\} - \mathbb{E}\{X\}^2}, \quad \theta_X = \frac{\mathbb{E}\{X^2\} - \mathbb{E}\{X\}^2}{\mathbb{E}\{X\}} \tag{47}$$

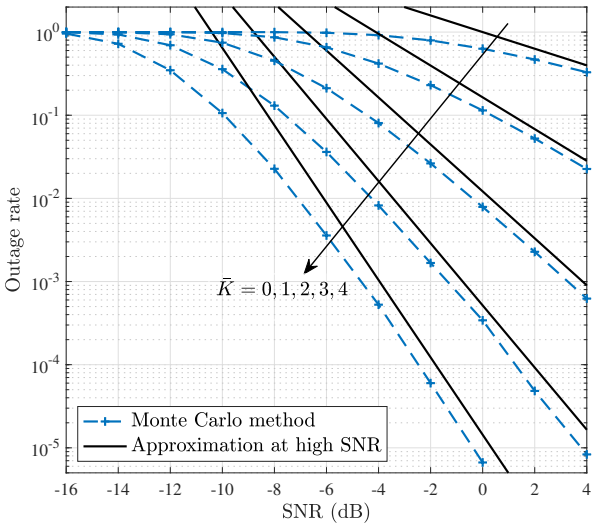


Fig. 2. Outage rate of the proposed RIS-RPM scheme in the case of $N = 1$ assuming $G = 4$, $R = 1$, $\bar{K} = \{0, 1, \ldots, G\}$, $\sigma_d^2 = \sigma_r^2 = 1$ and perfect CSI available at the AP.

where

$$\mathbb{E}\{X\} = (\bar{K}+1)(1 + \frac{\pi}{4}\bar{K}) \tag{48}$$

$$\begin{aligned}\mathbb{E}\{X^2\} =& 2(\bar{K}+1) + (\frac{3\pi}{2} + 3)(\bar{K}+1)\bar{K} \\ &+ \frac{3\pi}{2}(\bar{K}+1)\bar{K}(\bar{K}-1) \\ &+ \frac{\pi^2}{16}(\bar{K}+1)\bar{K}(\bar{K}-1)(\bar{K}-2).\end{aligned} \tag{49}$$

Its PDF is given by

$$p(x) = x^{k_X - 1}\frac{e^{-x/\theta_X}}{(\theta_X)^{k_X}\Gamma(k_X)}, \quad x \geq 0. \tag{50}$$

Approximating e^{-x/θ_X} by 1 for $x \to 0$, we have

$$\mathbb{P}\{X < \delta\} \approx \frac{\delta^{k_X}}{(\theta_X)^{k_X}\Gamma(k_X + 1)} \tag{51}$$

for a small positive value $\delta \to 0$. Hence at high SNR the outage probability can be approximated by

$$p_{\text{out}}(R) \approx p_{\text{out}}^{\text{U}}(R) \triangleq \frac{(2^R - 1)^{k_X}}{(\theta_X)^{k_X}\Gamma(k_X + 1)}\gamma^{-k_X}. \tag{52}$$

From (52), we see a diversity gain of k_X. It can be readily verified that k_X linearly increases with $\bar{K}$ and satisfies $1 \leq k_X \leq \bar{K} + 1$ with equality if and only if $\bar{K} = 0$. As shown in Fig. 2, $p_{\text{out}}^{\text{U}}(R)$ well approximates $p_{\text{out}}(R)$ in the high SNR region, and an increase of $\bar{K}$ leads to an increase in the diversity gain.

As a benchmark, we consider that the phases of the RIS are not configurable but are fixed, e.g., $\boldsymbol{\Phi} = \mathbf{I}_G$, which does not require any CSI for optimization and thus can dispense with the channel acquisition. In this case, X is the square of the sum of $\bar{K} + 1$ independent $\mathcal{CN}(0, 1)$ random variables and follows a Gamma distribution with parameters $k_x = 1$ and $\theta_x = \bar{K} + 1$. Hence at high SNR the outage probability with

$$\bar{R}_{\text{RIS-RPM}}=\log_2 J+\log_2 M-\log_2 e - \frac{1}{J}\frac{1}{M}\sum_{j=1}^{J}\sum_{m=1}^{M}\mathbb{E}_{\tilde{\mathbf{g}}_r,\tilde{g}_d,v}\left\{\log_2\sum_{j'=1}^{J}\sum_{m'=1}^{M}e^{-\frac{\left|\frac{v}{\sqrt{P_t}}+\left(\sum\limits_{k\in\mathbb{S}_j}[\tilde{\mathbf{g}}_r]_k+\tilde{g}_d\right)a_m-\left(\sum\limits_{k\in\mathbb{S}_{j'}}[\tilde{\mathbf{g}}_r]_k+\tilde{g}_d\right)a_{m'}\right|^2}{\sigma^2/P_t}}\right\} \tag{56}$$

a unit phase shift design can be approximated by

$$p_{\text{out}}^{\text{unit}}(R)\approx\frac{2^R-1}{(\bar{K}+1)\gamma}. \tag{53}$$

Comparing (52) with (53), we can see an increase in the diversity gain by appropriately optimizing the phase shifts of the RIS.

B. Achievable Rate

In practice, the constellation $\mathbb{A}$ is a finite and discrete complex signal set of cardinality M with normalized power, i.e., $\mathbb{A}\triangleq\{a_m\}_{m=1}^{M}$ with $\mathbb{E}\{|a_m|^2\}=1$, where the constellation points are independent and equiprobable. Let $\mathbb{S}\triangleq\{\mathbb{S}_j\}_{j=1}^{J}$ denote the index set of all the possible combinations of the $\bar{K}$ ON-state groups with cardinality $J=\binom{G}{\bar{K}}$, where $\mathbb{S}_j$ is the j-th element of set $\mathbb{S}$. We assume that all the combinations in $\mathbb{S}$ are independent and equiprobable.

1) Beamforming Design Based on Statistical ON/OFF State Information: First, we focus on the RIS-RPM scheme with the practical beamforming design based on statistical ON/OFF state information of the RIS. We rewrite (3) as

$$y=\sqrt{P_t}\left(\mathbf{s}^T\mathbf{\Phi}\mathbf{H}\mathbf{w}+\mathbf{h}_d^H\mathbf{w}\right)x+n. \tag{54}$$

We define $\tilde{\mathbf{g}}_r=\mathbf{\Phi}\mathbf{H}\mathbf{w}$, whose entries are the effective channels of the cascaded AP-RIS-user links observed at the user associated with the corresponding RIS-elements groups. Let $\tilde{g}_d=\mathbf{h}_d^H\mathbf{w}$ denote the effective channel of the direct AP-user link observed at the user. Before detecting the information data transmitted by the AP and RIS, the user has to acquire $\tilde{\mathbf{g}}_r$ and $\tilde{g}_d$. In particular, $\tilde{\mathbf{g}}_r$ can be obtained at the user based on downlink pilot symbols by using $\mathbf{\Phi}$ as the RIS reflection pattern, while $\tilde{g}_d$ can be obtained by applying conventional channel estimation methods with all RIS-elements groups turned OFF. Therefore, we assume that $\tilde{\mathbf{g}}_r$ and $\tilde{g}_d$ are available at the user side, and the achievable rate of the RIS-RPM scheme with practical beamforming design is given by

$$\bar{R}_{\text{RIS-RPM}}=\mathbb{E}_{\tilde{\mathbf{g}}_r,\tilde{g}_d}\left\{\mathrm{I}\left(\mathbb{I},x;y|\tilde{\mathbf{g}}_r,\tilde{g}_d\right)\right\} \tag{55}$$

where $\mathrm{I}(X,Y;Z)$ denotes the mutual information between the random vector (X,Y) and the random variable Z.

Proposition 1. *The achievable rate of the RIS-RPM scheme with practical beamforming design is given by (56), which is shown at the top of the page, where v is a complex Gaussian random variable distributed as $\mathcal{CN}(0,\sigma^2)$.*

Proof. According to the definition of mutual information, (55) can be derived as

$$\bar{R}_{\text{RIS-RPM}}=\mathrm{H}\left(\mathbb{I},x\right)-\mathbb{E}_{\tilde{\mathbf{g}}_r,\tilde{g}_d}\left\{\mathrm{H}\left(\mathbb{I},x|y,\tilde{\mathbf{g}}_r,\tilde{g}_d\right)\right\} \tag{57}$$

where $\mathrm{H}(\cdot)$ and $\mathrm{H}(\cdot|\cdot)$ denote the marginal entropy and conditional entropy, respectively. Due to the independence between the selection of $\mathbb{I}$ and the symbol modulation at the AP, we have $\mathrm{H}(\mathbb{I},x)=\log_2 J+\log_2 M$. The last term in the right hand side of (57) can be expressed according to the definition of conditional entropy as

$$\mathrm{H}(\mathbb{I},x|y,\tilde{\mathbf{g}}_r,\tilde{g}_d)=\frac{1}{J}\frac{1}{M}\sum_{j=1}^{J}\sum_{m=1}^{M}\int_y p\left(y|\mathbb{I}=\mathbb{S}_j,x=a_m,\tilde{\mathbf{g}}_r,\tilde{g}_d\right)\times\log_2\frac{p(y)}{\frac{1}{J}\frac{1}{M}p\left(y|\mathbb{I}=\mathbb{S}_j,x=a_m,\tilde{\mathbf{g}}_r,\tilde{g}_d\right)}dy \tag{58}$$

where

$$p\left(y|\mathbb{I}=\mathbb{S}_j,x=a_m,\tilde{\mathbf{g}}_r,\tilde{g}_d\right)=\frac{1}{\pi\sigma^2}e^{-\frac{\left|y-\sqrt{P_t}\left(\sum\limits_{k\in\mathbb{S}_j}[\tilde{\mathbf{g}}_r]_k+\tilde{g}_d\right)a_m\right|^2}{\sigma^2}} \tag{59}$$

and

$$p(y)=\frac{1}{J}\frac{1}{M}\frac{1}{\pi\sigma^2}\sum_{j'=1}^{J}\sum_{m'=1}^{M}e^{-\frac{\left|y-\sqrt{P_t}\left(\sum\limits_{k\in\mathbb{S}_{j'}}[\tilde{\mathbf{g}}_r]_k+\tilde{g}_d\right)a_{m'}\right|^2}{\sigma^2}}. \tag{60}$$

Replacing y with $v\triangleq y-\sqrt{P_t}\left(\sum\limits_{k\in\mathbb{S}_j}[\tilde{\mathbf{g}}_r]_k+\tilde{g}_d\right)a_m$ yields

$$\mathrm{H}\left(\mathbb{I},x|y,\tilde{\mathbf{g}}_r,\tilde{g}_d\right)=-\log_2\pi\sigma^2+\frac{1}{J}\frac{1}{M}\sum_{j=1}^{J}\sum_{m=1}^{M}\int_v p(v)\times\log_2\sum_{j'=1}^{J}\sum_{m'=1}^{M}e^{-\frac{\left|\frac{v}{\sqrt{P_t}}+\left(\sum\limits_{k\in\mathbb{S}_j}[\tilde{\mathbf{g}}_r]_k+\tilde{g}_d\right)a_m-\left(\sum\limits_{k\in\mathbb{S}_{j'}}[\tilde{\mathbf{g}}_r]_k+\tilde{g}_d\right)a_{m'}\right|^2}{\sigma^2/P_t}}dv + \frac{1}{J}\frac{1}{M}\sum_{i=1}^{J}\sum_{j=1}^{M}\int_v p(v)\log_2\frac{1}{p(v)}dv \tag{61}$$

where $p(v)=\frac{1}{\pi\sigma^2}\exp(-\frac{|v|^2}{\sigma^2})$ is the PDF of a complex Gaussian random variable with zero mean and variance σ^2. Since the last term in the right hand side of (61) is the differential entropy of a $\mathcal{CN}(0,\sigma^2)$ random variable, which is equal to $\log_2\pi\sigma^2 e$, we finally obtain the expression of $\bar{R}_{\text{RIS-RPM}}$ in (56). □

2) Beamforming Design Based on Instantaneous ON/OFF State Information: We characterize the upper bound of the achievable rate of the RIS-RPM scheme under the assumption

$$\bar{R}^{\text{UB}}_{\text{RIS-RPM}} = \log_2 J + \log_2 M - \log_2 e - \frac{1}{J}\frac{1}{M}\sum_{j=1}^{J}\sum_{m=1}^{M}\mathbb{E}_{\mathbf{H},\mathbf{h}_d,v}\left\{\log_2 \sum_{j'=1}^{J}\sum_{m'=1}^{M} e^{-\frac{\left|v+\sqrt{P_t}f(\mathbb{S}_j,\mathbf{H},\mathbf{h}_d)a_m-\sqrt{P_t}f(\mathbb{S}_{j'},\mathbf{H},\mathbf{h}_d)a_{m'}\right|^2}{\sigma^2}}\right\} \tag{64}$$

that the active and passive beamforming vectors are optimized based on the instantaneous ON/OFF state information of the RIS. In this case, the beamforming $\mathbf{w}$ and $\boldsymbol{\theta}_{\mathbb{I}}$ highly depend on the selection of $\mathbb{I}$. We define

$$f(\mathbb{I},\mathbf{H},\mathbf{h}_d) \triangleq \left(\boldsymbol{\theta}_{\mathbb{I}}^T\mathbf{H}_{\mathbb{I}} + \mathbf{h}_d^H\right)\mathbf{w} \tag{62}$$

which is the effective channel observed at the user that can be obtained by downlink channel training. By assuming that $f(\mathbb{I},\mathbf{H},\mathbf{h}_d)$ is available at the user, the achievable rate of the RIS-RPM scheme with the beamforming design presented in Section IV is given by

$$\bar{R}^{\text{UB}}_{\text{RIS-RPM}} = \mathbb{E}_{\mathbf{H},\mathbf{h}_d}\left\{\text{I}\left(\mathbb{I},x;y|\mathbf{H},\mathbf{h}_d\right)\right\}. \tag{63}$$

Proposition 2. *The achievable rate of the RIS-RPM scheme with the beamforming design presented in Section IV is given by (64), which is shown at the top of the page, where v is a complex Gaussian random variable distributed as $\mathcal{CN}(0,\sigma^2)$.*

Proof. The proof is similar to (56) and is omitted for brevity. □

VI. Simulation Results and Discussions

In this section, simulation results are presented to evaluate the performance of the proposed schemes. We consider a three dimensional coordinate system, where the centers of the AP and RIS are located at $(0,0,0)$ and $(0,d_0,0)$, respectively, and the user is located at $(0,d_y,d_z)$. For the AP, we consider a uniform linear array of $N=4$ antennas with an antenna spacing of half-wavelength, located in the x-axis. For the RIS, we consider a uniform square array of $L=12\times 12=144$ elements with an element spacing of half-wavelength, deployed in the x-z plane. The fading channel model for all the channels involved is given by

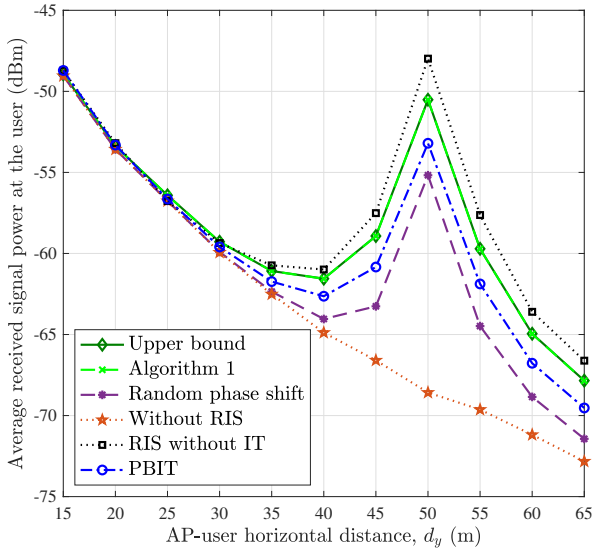
$$h = \sqrt{\frac{\kappa\varpi}{\kappa+1}}e^{-j2\pi d/\vartheta} + \sqrt{\frac{1}{\kappa+1}}\mathcal{CN}(0,\varpi)$$

where κ is the Rician factor, ϖ is the path loss, d is the signal propagation distance, and ϑ is the signal wavelength. We use the path loss model $\varpi = C_0 d^{-\alpha}$, where C_0 is the path loss at a reference distance of 1 meter (m), and α is the path loss exponent. The Rician factors of the AP-user, AP-RIS, and RIS-user links are set to $\kappa_{\text{Au}}=0$, $\kappa_{\text{AR}}=\infty$, and $\kappa_{\text{Ru}}=0$, respectively, as in Section V-A. The path loss exponents of the AP-user, AP-RIS, and RIS-user links are set to 3.8, 2.2, and 2.4, respectively. The noise power is set to $\sigma^2=-80$ dBm. Other parameters are set as follows: $d_0=50$ m, $d_z=2$ m, $\vartheta=0.1$ m, $T_c=150$ symbol sampling periods, $C_0=30$ dB, and a quadrature phase-shift keying (QPSK) modulation at the AP is assumed. We consider the Zadoff-Chu sequence as the pilot sequence and set $P_p=10$ dBm for channel estimation. The maximum iteration number in Algorithm 1 is $I=5$ and the convergence threshold is $\epsilon=10^{-4}$. The parameters G and $\bar{K}$ as well as the AP transmit power level are specified next to study their effects on the system performance. The simulation results are obtained by averaging over more than 1000 independent channel realizations.

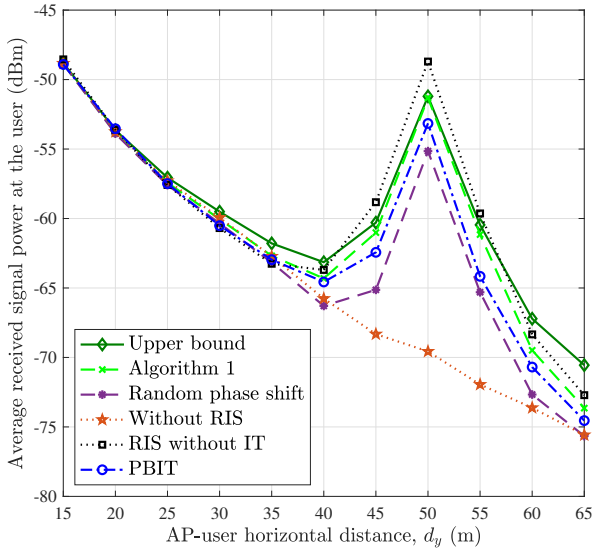
A. Performance of Algorithm 1

To evaluate the effectiveness of Algorithm 1, we consider the following schemes with fixed $G=4$ and $P_t=20$ dBm: 1) Benchmark scheme without information transfer (IT) [15] where all RIS elements are turned ON and problem (P2) with $\mathbb{I}=\mathbb{G}$ is solved to obtain the active and passive beamforming vectors; 2) Conventional MISO scheme without RIS in which $\mathbf{w}=\hat{\mathbf{h}}_d/\left\|\hat{\mathbf{h}}_d\right\|$; 3) Random phase shift scheme where the phase shifts of the diagonal entries in $\boldsymbol{\Phi}$ are randomly drawn from $(0,2\pi]$ and then problem (P1.1) is solved to obtain the active beamforming vector; 4) Upper bound that solves problem (P2); 5) PBIT scheme where the elements in the ON/OFF-state vector $\mathbf{s}$ are independently drawn from the set $\{0,1\}$ with equal probability. For the PBIT scheme, the active and passive beamforming vectors are obtained by solving problem (P1) with $\mathbf{A}=\frac{1}{4}(\mathbf{1}_G\cdot\mathbf{1}_G^T+\mathbf{I}_G)$ and $\mathbf{a}=\frac{1}{2}\cdot\mathbf{1}_G$. The number of ON-state groups in Algorithm 1 is $\bar{K}=3$. We note that the schemes 1) and 2) are two special cases of Algorithm 1 with $\bar{K}=G$ and $\bar{K}=0$, respectively. From Fig. 3, we observe that all the RIS-assisted schemes significantly enhance the average received power at the user as compared to the scheme without RIS, when the user is located in the vicinity of the RIS. Also, Algorithm 1 significantly outperforms the random phase shift scheme. Moreover, as expected from Section IV, the average received power achieved by Algorithm 1 is upper-bounded by the beamforming design based on the instantaneous ON/OFF state information of the RIS. On the other hand, it is worth pointing out that Algorithm 1 incurs a small received signal power loss as compared to the scheme without IT. This is expected since one RIS-elements group is turned OFF deliberately to convey the additional information of the RIS in Algorithm 1, while all RIS-elements groups are used for enhancing the reflected signal power in the conventional scheme. The impact of channel estimation errors on the performance of active and passive beamforming is shown in Fig. 3 as well. We observe that, compared to the upper bound, Algorithm 1 with perfect CSI at the AP achieves nearly the same performance, while with imperfect CSI it performs worse when the user is far away from both the AP and RIS. This can be explained by the fact that, given the same transmit power at the user, the received signal power for training is lower when the user moves away from both the AP and RIS, and thus the channel estimation accuracy is reduced.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2020.3028198, IEEE Transactions on Wireless Communications



(a) Perfect CSI at the AP



(b) Estimated CSI at the AP, $P_p = 10$ dBm

Fig. 3. Average received signal power at the user versus AP-user horizontal distance d_y, where $G = 4$, $P_t = 20$ dBm, and $\bar{K} = 3$ in Algorithm 1.

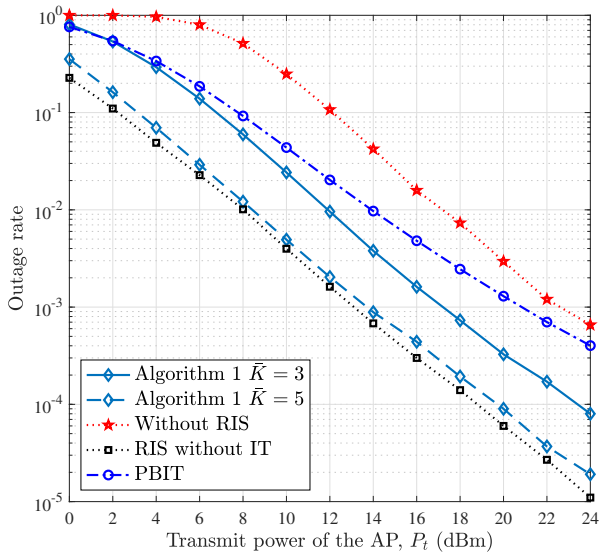


Fig. 4. Performance comparison between the proposed RIS-RPM scheme and the PBIT counterpart, where $G = 6$, $R = 1$, and $d_y = 45$ m.

B. Outage Rate Performance

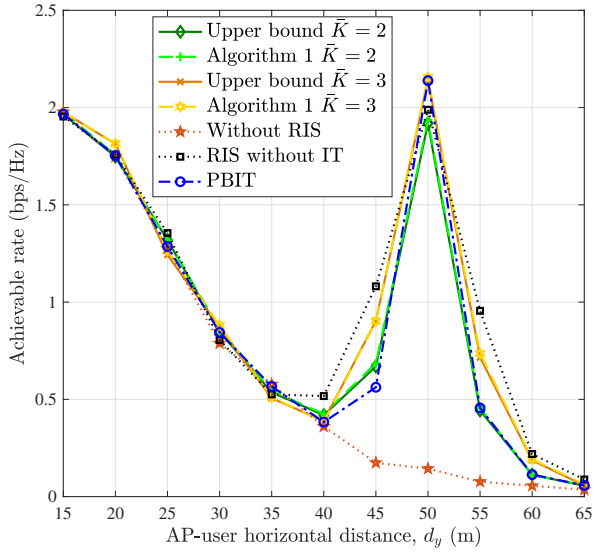
In Fig. 4, we consider the PBIT scheme by considering that the elements in $\mathbf{s}$ are independently drawn from the set $\{0, 1\}$ with equal probability. Under the PBIT scheme, the average number of ON-state groups at the RIS is $\mathbb{E}\{\mathbf{s}^T\mathbf{s}\} = G/2$. The outage rate of our proposed RIS-RPM scheme with $\bar{K} = G-1$ and $\bar{K} = G/2$ is illustrated. Two benchmark schemes are considered: 1) Benchmark scheme without IT; 2) Conventional MISO scheme without RIS. Fig. 4 shows the outage rate of different schemes versus the AP transmit power, where $G = 6$, $R = 1$ and $d_y = 45$ m. We observe that the proposed RIS-RPM scheme with $\bar{K} = G/2 = 3$ outperforms the PBIT scheme. This originates from the fact that the received power of the PBIT scheme fluctuates because the number of ON-state elements of the RIS varies. On the other hand, the proposed RIS-RPM scheme keeps the number of ON-state elements fixed, thus reducing the signal fluctuations and, thus, providing a better outage probability. Moreover, by increasing the number of ON-state groups $\bar{K}$, the outage rate of the RIS-RPM scheme can be further improved.

C. Achievable Rate Performance

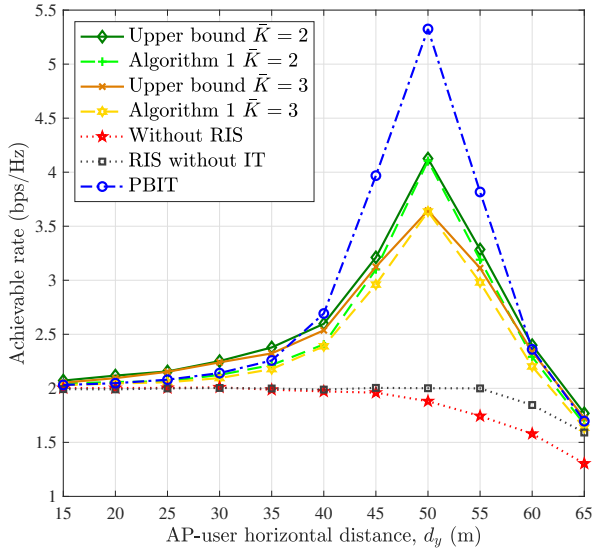
To evaluate the achievable rate of the RIS-RPM scheme, the following schemes are considered: 1) Benchmark scheme without IT; 2) Conventional MISO scheme without RIS; 3) Upper bound whose achievable rate is obtained from (64); 4) PBIT scheme whose elements in $\mathbf{s}$ are independently drawn from the set $\{0, 1\}$ with equal probability. The active and passive beamforming vectors are obtained based on the estimated CSI.

1) Effect of AP-User Horizontal Distance d_y: In Fig. 5, we evaluate the effect of d_y on the achievable rate, where $G = 4$, $\bar{K} = \{2, 3\}$ and $P_t = \{0 \text{ dBm}, 20 \text{ dBm}\}$. One can observe from Fig. 5(a) that, when the transmit power of the AP is very low, the achievable rate of the scheme without RIS decreases as the user moves away from the AP. In contrast, the achievable rate of RIS-aided schemes increases significantly as the user moves toward the RIS, and decrease as the user moves away from both the AP and RIS. This is because a user close to either the AP or RIS is able to receive a stronger transmitted/reflected signals from the AP/RIS. This finding implies that cell-edge users can benefit from an RIS deployed in their neighborhood, i.e., the rate of cell-edge users can be enhanced by deploying an RIS, instead of increasing the AP transmit power or deploying an expensive AP/relay. Also, we observe from Fig. 5(a) that the RIS-RPM scheme with $\bar{K} = 3$ has the potential to

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2020.3028198, IEEE Transactions on Wireless Communications



(a) $P_t = 0$ dBm



(b) $P_t = 20$ dBm

Fig. 5. Effect of AP-user horizontal distance d_y on the achievable rate, where $G = 4$, and (a) $P_t = 0$ dBm; (b) $P_t = 20$ dBm.

outperform the scheme without IT even though the received signal power of the RIS-RPM scheme is lower as shown in Fig. 3(b). As for the RIS-RPM scheme with $\bar{K} = 2$, since a large number of RIS elements are turned OFF deliberately for information transfer, the additional information from the RIS cannot compensate for the information reduction caused by the loss of the received signal power, thus leading to a smaller achievable rate than the scheme without IT. Moreover, it can be observed from Fig. 5(b) that, when the AP transmit power is high, the RIS-RPM scheme exhibits significantly superior rate performance over the scheme without IT, which is attributed to the additional information delivered by the RIS. This implies that the RIS-RPM scheme provides a mechanism to enable a flexible tradeoff between the received signal power and the

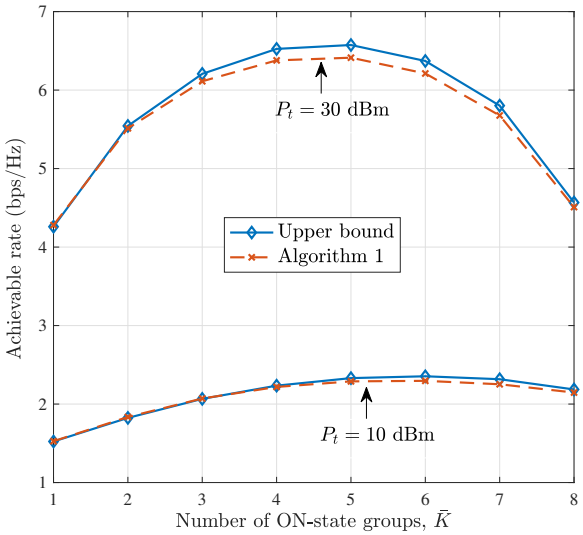


Fig. 6. Effect of $\bar{K}$ on the achievable rate, where $G = 9$, $d_y = 45$ m, and P_t equal to 10 dBm and 30 dBm are considered.

TABLE II
GROUPING RATIO AND CHANNEL ESTIMATION OVERHEAD RATIO

G	2	4	6
ρ	1/72	1/36	1/24
ξ	3/150	5/150	7/150
$L_x \times L_z$	12×6	6×6	6×4

Each group consists of $\bar{L} = L_x \times L_z$ elements with L_x elements along x-axis and L_z elements along z-axis.

achievable rate by varying the number of OFF-state groups at the RIS. Moreover, when the transmit power of the AP is high, the PBIT scheme achieves the maximum achievable rate, as expected.

2) Effect of the Number of ON-State RIS-elements Groups $\bar{K}$: We compare the achievable rate of the RIS-RPM scheme versus $\bar{K}$ in Fig. 6, by assuming $G = 9$, $d_y = 45$ m, and $P_t = \{10 \text{ dBm}, 30 \text{ dBm}\}$. We observe that there exists an optimal $\bar{K}$, which varies with the AP transmit power level. For $P_t = 30$ dBm, the optimal $\bar{K}$ is 5 whereas the optimal $\bar{K}$ for $P_t = 10$ dBm is 6. Generally, it is expected that when the AP transmit power is high enough, the optimal $\bar{K}$ is likely to be $\lceil G/2 \rceil$, since the entropy of the RIS is maximized. Furthermore, when the AP transmit power is very high, due to the assumption of a finite-alphabet input, the achievable rate for different values of $\bar{K}$ is the sum of the corresponding uncoded transmitted information rates of the AP and RIS. In contrast, when the AP transmit power becomes very low, the optimal $\bar{K}$ is likely to be $(G-1)$, since the reflected signal power is maximized while the RIS still can convey its information through the RPM scheme.

3) Effect of RIS-elements Grouping Ratio: The RIS-elements grouping ratio is defined as $\rho \triangleq 1/\bar{L}$. Let $\xi \triangleq (G+1)/T_c$ denote the ratio of the time overhead for channel estimation to the coherence time normalized to the symbol

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2020.3028198, IEEE Transactions on Wireless Communications

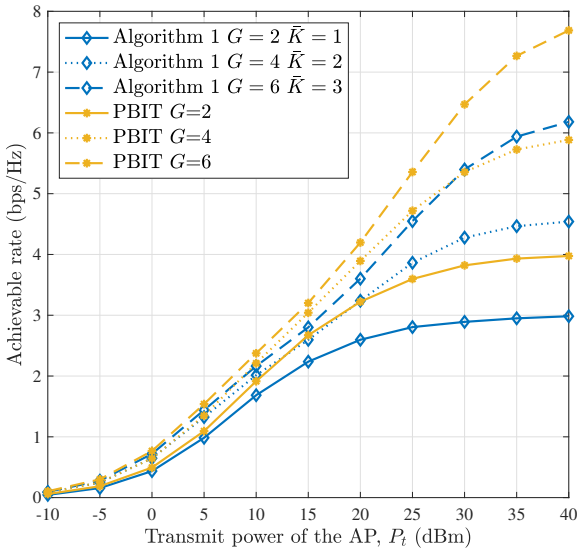


Fig. 7. Effect of RIS-elements grouping ratio on the achievable rate, where $d_y = 45$ m and the average number of ON-state RIS elements at each time is $\bar{L} \times \bar{K} = 72$.

sampling period. In Fig. 7, we examine the effect of ρ on the achievable rate, where $d_y = 45$ m, $G = \{2, 4, 6\}$, and $\bar{K} = G/2 = \{1, 2, 3\}$ for the proposed RIS-RPM scheme. In particular, the average number of ON-state RIS elements is kept constant for different schemes. We observe that the RIS-RPM scheme with a large G achieves a better achievable rate. This can be justified by observing that the degrees of freedom for designing the RIS reflection coefficients and the amount of additional information that can be transmitted by the proposed scheme increases as the grouping ratio increases. However, as shown in Table II, the pilot overhead ratio ξ increases with the grouping ratio, which results in more time for channel estimation and less time for data transmission, thus reducing the average achievable rate. On the other hand, we can observe that a large AP transmit power level is needed for the PBIT scheme to be competitive.

VII. Conclusions

In this paper, we considered an RIS-enhanced MISO wireless communication system and proposed the RIS-RPM scheme to achieve PBIT. A practical beamforming design based on the statistical ON/OFF state information of the RIS was proposed to maximize the *average* received signal power at the user, for which an efficient algorithm based on the alternating optimization technique was proposed to obtain a high-quality solution. Also, we formulated an optimization problem to maximize the *instantaneous* received signal power by designing active and passive beamforming vectors based on the instantaneous ON/OFF state information of the RIS, which yields as upper bound for the received signal power of the RIS-RPM scheme. Moreover, the asymptotic outage probability of the RIS-RPM scheme in Rician fading was derived in closed-form. In particular, the RIS was shown to be able to increase the diversity gain by properly designing the phase shifts of its elements. The achievable rate of the RIS-RPM scheme was analyzed for the case where the transmitted symbols are drawn from a finite constellation. Finally, simulation results corroborated the effectiveness of the proposed optimization algorithms for optimizing the RIS-RPM scheme and revealed the effect of different system parameters on the achievable rate. It was shown that the RIS-RPM scheme is able to improve the achievable rate even though it undergoes a loss in the received signal power compared to the conventional RIS-assisted system with full-ON reflection.

References

[1] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[2] F. Liu, A. Pitilakis, M. S. Mirmoosa, O. Tsilipakos, X. Wang, A. C. Tasolamprou, S. Abadal, A. Cabellos-Aparicio, E. Alarcón, C. Liaskos, N. V. Kantartzis, M. Kafesaki, E. N. Economou, C. M. Soukoulis, and S. Tretyakov, "Programmable metasurfaces: State of the art and prospects," in *Proc. IEEE ISCAS*, Florence, Italy, May 2018, pp. 1–5.

[3] T. J. Cui, M. Q. Qi, X. Wan, J. Zhao, and Q. Cheng, "Coding metamaterials, digital metamaterials and programmable metamaterials," *L. Sci. & Appl.*, vol. 3, no. 10, pp. e218, Oct. 2014.

[4] X. Tan, Z. Sun, D. Koutsonikolas, and J. M. Jornet, "Enabling indoor mobile millimeter-wave networks based on smart reflect-arrays," in *Proc. IEEE INFOCOM*, Honolulu, HI, USA, Apr. 2018, pp. 270–278.

[5] M. Di Renzo and J. Song, "Reflection probability in wireless networks with metasurface-coated environmental objects: An approach based on random spatial processes," *EURASIP J. Wireless Commun. Netw.*, no. 99, Apr. 2019.

[6] N. S. Perović, M. Di Renzo, and M. F. Flanagan, "Channel capacity optimization using reconfigurable intelligent surfaces in indoor mmWave environments," in *Proc. IEEE ICC*, Dublin, Ireland, Jun. 2020, pp. 1–7.

[7] C. Liaskos, S. Nie, A. Tsioliaridou, A. Pitsillides, S. Ioannidis, and I. Akyildiz, "A new wireless communication paradigm through software-controlled metasurfaces," *IEEE Commun. Mag.*, vol. 56, no. 9, pp. 162–169, Sep. 2018.

[8] M. Di Renzo, A. Zappone, M. Debbah, M.-S. Alouini, C. Yuen, J. de Rosny, and S. Tretyakov, "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and road ahead," *IEEE J. Select. Areas Commun.*, early access, 2020.

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

[10] C. Huang, S. Hu, G. C. Alexandropoulos, A. Zappone, C. Yuen, R. Zhang, M. Di Renzo, and M. Debbah, "Holographic MIMO surfaces for 6G wireless networks: Opportunities, challenges, and trends," *IEEE Wireless Commun.*, early access, 2020.

[11] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 4157–4170, Aug. 2019.

[12] E. Björnson, Ö. Özdogan, and E. G. Larsson, "Intelligent reflecting surface versus decode-and-forward: How large surfaces are needed to beat relaying?" *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 244–248, Feb. 2020.

[13] M. Di Renzo, K. Ntontin, J. Song, F. H. Danufane, X. Qian, F. Lazarakis, J. De Rosny, D. Phan-Huy, O. Simeone, R. Zhang, M. Debbah, G. Lerosey, M. Fink, S. Tretyakov, and S. Shamai, "Reconfigurable intelligent surfaces vs. relaying: Differences, similarities, and performance comparison," *IEEE Open J. Commun. Societ.*, vol. 1, pp. 798–807, Jun. 2020.

[14] B. Zheng, Q. Wu, and R. Zhang, "Intelligent reflecting surface-assisted multiple access with user pairing: NOMA or OMA?" *IEEE Commun. Lett.*, vol. 24, no. 4, pp. 753–757, Apr. 2020.

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

[16] B. Zheng, C. You, and R. Zhang, "Double-IRS assisted multi-user MIMO: Cooperative passive beamforming design," *arXiv preprint arXiv:2008.13701*, 2020.

[17] A. Taha, M. Alrabeiah, and A. Alkhateeb, "Deep learning for large intelligent surfaces in millimeter wave and massive MIMO systems," in *Proc. IEEE GLOBECOM*, Waikoloa, HI, USA, Dec. 2019, pp. 1–6.

[18] Z. 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.

[19] D. Mishra and H. Johansson, "Channel estimation and low-complexity beamforming design for passive intelligent surface assisted MISO wireless energy transfer," in *Proc. IEEE ICASSP*, Brighton, United Kingdom, May 2019, pp. 4659–4663.

[20] Y. Yang, B. Zheng, S. Zhang, and R. Zhang, "Intelligent reflecting surface meets OFDM: Protocol design and rate maximization," *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 4522–4535, Jul. 2020.

[21] B. Zheng and R. Zhang, "Intelligent reflecting surface-enhanced OFDM: Channel estimation and reflection optimization," *IEEE Wireless Commun. Lett.*, vol. 9, no. 4, pp. 518–522, Apr. 2020.

[22] T. L. Jensen and E. De Carvalho, "An optimal channel estimation scheme for intelligent reflecting surfaces based on a minimum variance unbiased estimator," in *Proc. IEEE ICASSP*, Barcelona, Spain, May 2020, pp. 5000–5004.

[23] C. You, B. Zheng, and R. Zhang, "Intelligent reflecting surface with discrete phase shifts: Channel estimation and passive beamforming," in *Proc. IEEE ICC*, Dublin, Ireland, Jun. 2020, pp. 1–6.

[24] S. Lin, B. Zheng, G. C. Alexandropoulos, M. Wen, F. Chen, and S. Mumtaz, "Adaptive transmission for reconfigurable intelligent surface-assisted OFDM wireless communications," *IEEE J. Select. Areas Commun.*, early access, 2020.

[25] B. Zheng, C. You, and R. Zhang, "Fast channel estimation for IRS-assisted OFDM," *arXiv preprint arXiv:2008.04476*, 2020.

[26] L. Wei, C. Huang, G. C. Alexandropoulos, and C. Yuen, "Parallel factor decomposition channel estimation in RIS-assisted multi-user MISO communication," in *Proc. IEEE SAM*, Hangzhou, China, Jun. 2020.

[27] B. Zheng, C. You, and R. Zhang, "Intelligent reflecting surface assisted multi-user OFDMA: Channel estimation and training design," *IEEE Trans. Wireless Commun.*, early access, 2020.

[28] C. Huang, R. Mo, and C. Yuen, "Reconfigurable intelligent surface assisted multiuser MISO systems exploiting deep reinforcement learning," *IEEE J. Select. Areas Commun.*, vol. 38, no. 8, pp. 1839–1850, Aug. 2020.

[29] E. Basar, "Reconfigurable intelligent surface-based index modulation: A new beyond MIMO paradigm for 6G," *IEEE Trans. Commun.*, vol. 68, no. 5, pp. 3187–3196, May 2020.

[30] R. Karasik, O. Simeone, M. Di Renzo, and S. Shamai Shitz, "Beyond max-SNR: Joint encoding for reconfigurable intelligent surfaces," in *Proc. IEEE ISIT*, Los Angeles, USA, Jun. 2020, pp. 2965–2970.

[31] M. Di Renzo, M. Debbah, D.-T. Phan-Huy, A. Zappone, M.-S. Alouini, C. Yuen, V. Sciancalepore, G. C. Alexandropoulos, J. Hoydis, H. Gacanin, J. de Rosny, A. Bounceur, G. Lerosey, and M. Fink, "Smart radio environments empowered by reconfigurable AI meta-surfaces: An idea whose time has come," *EURASIP J. Wireless Commun. Netw.*, vol. 2019, no. 1, May 2019.

[32] X. Yuan, Y.-J. Zhang, Y. Shi, W. Yan, and H. Liu, "Reconfigurable-intelligent-surface empowered 6G wireless communications: Challenges and opportunities," *arXiv preprint arXiv:2001.00364*, 2020.

[33] W. Yan, X. Yuan, and X. Kuai, "Passive beamforming and information transfer via large intelligent surface," *IEEE Wireless Commun. Lett.*, vol. 9, no. 4, pp. 533–537, Apr. 2020.

[34] W. Yan, X. Yuan, Z. He, and X. Kuai, "Large intelligent surface aided multiuser MIMO: Passive beamforming and information transfer," in *Proc. IEEE ICC*, Dublin, Ireland, Jun. 2020, pp. 1–7.

[35] W. Yan, X. Yuan, Z.-Q. He, and X. Kuai, "Passive beamforming and information transfer design for reconfigurable intelligent surfaces aided multiuser MIMO systems," *IEEE J. Select. Areas Commun.*, vol. 38, no. 8, pp. 1793–1808, Aug. 2020.

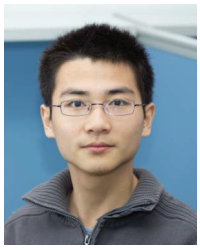
[36] H. Yang, X. Chen, F. Yang, S. Xu, X. Cao, M. Li, and J. Gao, "Design of resistor-loaded reflectarray elements for both amplitude and phase control," *IEEE Antennas Wireless Propag. Lett.*, vol. 16, pp. 1159–1162, Nov. 2016.

[37] A. Man-Cho, J. Zhang, and Y. Ye, "On approximating complex quadratic optimization problems via semidefinite programming relaxations," *Mathematical Programming*, vol. 110, no. 1, pp. 93–110, Jun. 2007.

[38] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming," 2016. [Online]. Available: http://cvxr.com/cvx

[39] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.

Shaoe Lin received the B.S. and M.S. degrees from the South China University of Technology, Guangzhou, China, in 2013 and 2016, respectively. She is currently pursuing the Ph.D. degree with the South China University of Technology, Guangzhou, China. From 2019 to 2020, she was a Visiting Student Research Collaborator with University of Waterloo, Waterloo, ON, Canada. Her recent research interests include MIMO systems, OFDM, and index modulation.



Beixiong Zheng (M'18) received the B.S. and Ph.D. degrees from the South China University of Technology, Guangzhou, China, in 2013 and 2018, respectively. He is currently a research fellow with the Department of Electrical and Computer Engineering, National University of Singapore. His recent research interests include intelligent reflecting surface (IRS), index modulation (IM), and non-orthogonal multiple access (NOMA).

From 2015 to 2016, he was a Visiting Student Research Collaborator with Columbia University, New York, NY, USA. He was a recipient of the Best Paper Award from the IEEE International Conference on Computing, Networking and Communications in 2016, the Best Ph.D. Thesis Award from China Education Society of Electronics in 2018, and the Outstanding Reviewer of Physical Communication in 2019.

George C. Alexandropoulos (S'07–M'10–SM'15) was born in Athens, Greece, in 1980. He received the Engineering Diploma degree in computer engineering and informatics, the M.A.Sc. degree (with distinction) in signal processing and communications, and the Ph.D. degree in wireless communications from the University of Patras (UoP), Rio-Patras, Greece, in 2003, 2005, and 2010, respectively. From 2001 to 2014, he has held research positions at various Greek universities and research institutes (UoP, University of Peloponnese, Technical University of Crete, National Center for Scientific Research "Demokritos," National Observatory of Athens, and the Athens Information Technology Center for Research and Education), where he technically managed several national, European, and international R&D projects, and lectured mathematics and computer engineering courses. From 2014 till January 2019, he was Senior Research Engineer at the Mathematical and Algorithmic Sciences Lab, Paris Research Center, Huawei Technologies France SASU delivering RAN solutions for 5G NR and beyond, and mainly for full duplex and massive MIMO, hybrid A/D beamforming management, and reconfigurable metasurfaces. Currently, he is Assistant Professor for communication systems and signal processing at the Department of Informatics and Telecommunications, National and Kapodistrian University of Athens, Greece. His research and development activities span the general areas of algorithmic design, optimization, and performance analysis for wireless communication networks with emphasis on multi-antenna systems, transceiver hardware architectures, high frequency communications, and distributed machine learning algorithms.

Dr. Alexandropoulos is a Senior Member of the IEEE Communications and Signal Processing Societies as well as a Professional Engineer of the Technical Chamber of Greece. He currently serves as an Editor for IEEE Transactions on Wireless Communications, IEEE Communications Letters, ELSEVIER Computer Networks, Frontiers in Communications and Networks, and ITU Journal on Future and Evolving Technologies. He has received scholarships for his postgraduate and PhD studies, a student travel grant for the 2010 IEEE Global Telecommunications Conference, the Best Ph.D. Thesis Award 2010 by a Greek University in the fields of informatics and telecommunications, and the IEEE Communications Society Best Young Professional in Industry Award 2018. More information is available at www.alexandropoulos.info.

Miaowen Wen (SM'18) received the Ph.D. degree from Peking University, Beijing, China, in 2014. From 2012 to 2013, he was a Visiting Student Research Collaborator with Princeton University, Princeton, NJ, USA. He is currently an Associate Professor with South China University of Technology, Guangzhou, China, and a Hong Kong Scholar with the University of Hong Kong, Hong Kong. He has published a Springer book entitled *Index Modulation for 5G Wireless Communications* and more than 100 IEEE journal papers. His research interests include a variety of topics in the areas of wireless and molecular communications.

Dr. Wen was the recipient of four Best Paper Awards from the IEEE ITST'12, the IEEE ITSC'14, the IEEE ICNC'16, and the IEEE ICCT'19, and was the winner in data bakeoff competition (Molecular MIMO) at IEEE Communication Theory Workshop (CTW) 2019, Selfoss, Iceland. He has served on the Editorial Boards of the IEEE ACCESS, and the *EURASIP Journal on Wireless Communications and Networking*. He served as a Guest Editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (Special Issue on Spatial Modulation for Emerging Wireless Systems) and for the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING (Special Issue on Index Modulation for Future Wireless Networks: A Signal Processing Perspective). He is currently serving as an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, the IEEE COMMUNICATIONS LETTERS, the *Physical Communication* (Elsevier), and the FRONTIERS IN COMMUNICATIONS AND NETWORKS, and a Guest Editor for the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING (Special Issue on Advanced Signal Processing for Local and Private 5G Networks).

Marco Di Renzo (S'05-AM'07-M'09-SM'14-F'20) was born in L'Aquila, Italy, in 1978. He received the Laurea (cum laude) and Ph.D. degrees in electrical engineering from the University of L'Aquila, Italy, in 2003 and 2007, respectively, and the Habilitation a Diriger des Recherches (Doctor of Science) degree from University Paris-Sud, France, in 2013.

Since 2010, he has been with the French National Center for Scientific Research (CNRS), where he is a CNRS Research Director (CNRS Professor) in the Laboratory of Signals and Systems (L2S) of Paris-Saclay University – CNRS and CentraleSupelec, Paris, France. Also, he is a Nokia Foundation Visiting Professor at Aalto University, Finland, and was a Honorary Professor at University Technology Sydney, Sydney, Australia, and a Visiting Professor at the University of L'Aquila, Italy.

He serves as the Editor-in-Chief of IEEE Communications Letters. He served as an Editor of IEEE Transactions on Communications, IEEE Transactions on Wireless Communications, IEEE Communications Letters, and as the Associate Editor-in-Chief of IEEE Communications Letters. He is a Distinguished Lecturer of the IEEE Vehicular Technology Society and IEEE Communications Society.

He is a recipient of several awards, including the 2013 IEEE-COMSOC Best Young Researcher Award for Europe, Middle East and Africa, the 2013 NoE-NEWCOM# Best Paper Award, the 2014-2015 Royal Academy of Engineering Distinguished Visiting Fellowship, the 2015 IEEE Jack Neubauer Memorial Best System Paper Award, the 2015 CNRS Award for Excellence in Research and Ph.D. Supervision, the 2016 MSCA Global Fellowship (declined), the 2017 SEE-IEEE Alain Glavieux Award, the 2018 IEEE-COMSOC Young Professional in Academia Award, the 2019 Nokia Foundation Visiting Professorship, and 8 Best Paper Awards at IEEE conferences (2012 and 2014 IEEE CAMAD, 2013 IEEE VTC-Fall, 2014 IEEE ATC, 2015 IEEE ComManTel, 2017 IEEE SigTelCom, EAI 2018 INISCOM, IEEE ICC 2019).

He is a Highly Cited Researcher (Clarivate Analytics, Web of Science) and a Fellow of the IEEE.

Fangjiong Chen (M'06) received the B.S. degree in electronics and information technology from Zhejiang University, Hangzhou, China, in 1997, and the Ph.D. degree in communication and information engineering from the South China University of Technology, Guangzhou, China, in 2002. He was with the School of Electronics and Information Engineering, South China University of Technology. From 2002 to 2005, he was a Lecturer, and from 2005 to 2011, he was an Associate Professor with the South China University of Technology.

He is currently a full-time Professor with the School of Electronics and Information Engineering, South China University of Technology. He is a vice director of Engineering Center for Short-Range Wireless Communication and Network, Ministry of Education. His research interests include signal detection and estimation, array signal processing, and wireless communication.

Prof. Chen received the National Science Fund for Outstanding Young Scientists in 2013. He was elected to the New Century Excellent Talent Program of MOE, China, in 2012.