Asymptotic Optimality of Reconfigurable Intelligent Surfaces: Passive Beamforming and Achievable Rate

Minchae Jung,[1] Walid Saad,[1] Mérouane Debbah,[2] and Choong Seon Hong[3]
[1]Wireless@VT, Bradley Department of Electrical and Computer Engineering, Virginia Tech, VA, USA,
Emails: {hosaly, walids}@vt.edu
[2]Mathematical & Algorithmic Sciences Lab, Huawei Technologies France SASU, 92100 Boulogne-Billancourt, France,
Email: merouane.debbah@huawei.com
[3]Department of Computer Science and Engineering, Kyung Hee University, Gyeonggi-do, Rep. of Korea,
Email: cshong@khu.ac.kr

***Abstract*—Reconfigurable intelligent surfaces (RISs) have recently emerged as a promising technology that can manipulate the properties of an incident wave, such as the frequency, amplitude, and phase, without the need for complex signal processing. In this paper, the asymptotic optimality of achievable rate in a downlink RIS system is analyzed under a practical RIS environment with its associated limitations. In particular, a passive beamformer that can achieve the asymptotic optimal performance by controlling the incident wave properties is designed, under practical reflection coefficients. In order to increase the achievable system sum-rate, a modulation scheme that can be used in an RIS without interfering with existing users is proposed and its average symbol error rate is asymptotically derived. Simulation results show that the proposed schemes are in close agreement with their upper bounds in presence of a large number of RIS reflecting elements thereby verifying that the achievable rate in practical RISs satisfies the asymptotic optimality.**

I. Introduction

THE concept of a metasurface is emerging as a promising technology that can support the demand for massive connectivity, mainly driven by upcoming Internet of Things (IoT) and 6G applications [1]–[8]. A metasurface relies on a massive integration of artificial meta-atoms that are commonly made of metal structures of low-cost and passive elements. Each meta-atom can manipulate the incident electromagnetic (EM) wave impinging on it, in terms of frequency, amplitude, and phase, and reflect it to a desired destination, without additional signal processing. A metasurface can potentially provide reliable and pervasive wireless connectivity given that man-made structures, such as buildings, walls, and roads, can be equipped with metasurfaces in the near future and used for wireless transmission [2]–[4]. Moreover, a tunable metasurface can significantly enhance the signal quality at a receiver by allowing a dynamic manipulation of the incident EM wave. Tunable metasurfaces are mainly controlled by electrical, optical, mechanical, and fluid operations [5] that can be programmed in software using a field programmable gate array (FPGA) [6]. The concept of a reconfigurable intelligent surface (RIS) is essentially an electronically operated metasurface controlled by programmable software, as introduced in [5] and [6]. In wireless communication systems, a base station (BS) can send control signals to an RIS controller (i.e., FPGA) via a dedicated control link and controls the properties of the incident wave to enhance the signal quality at the receiver. In principle, the electrical size of the unit reflecting elements (i.e., meta-atoms) deployed on RIS is between $\lambda/8$ and $\lambda/4$, where λ is a wavelength of radio frequency (RF) signal [5]. Note that conventional large antenna-array systems, such as a massive multiple-input and multiple-output (MIMO) and MIMO relay system, typically require antenna spacing of greater than $\lambda/2$ [3]. Therefore, an RIS can provide more reliable and space-intensive communications compared to conventional antenna-array systems as clearly explained in [2]–[4]. Moreover, a large number of reflecting elements can be arranged on each RIS thus offering precise control of the reflection wave and allowing it to coherently align with the desired channel.

Owing to these advantages, the use of an RIS in wireless communication systems has recently received significant attention as in [6]–[8]. An RIS is typically used for two main wireless communication purposes: a) RIS as an RF chain-free transmitter and b) RIS as a passive beamformer that amplifies the incident waveform (received from a BS) and reflects it to the desired user. In [6], the authors analyzed the error rate performance of a phase-shift keying (PSK) signaling and proved that an RIS transmitter equipped with a large number of reflecting elements can convey information with high reliability. The works in [7] and [8] proposed RF chain-free transmitter architectures enabled by an RIS that can support PSK and quadrature amplitude modulation (QAM). Meanwhile, the work in [6] theoretically analyzed the average symbol error rate (SER) resulting from an ideal passive beamformer and proved that the SER decays exponentially as the number of reflecting elements on RIS increases. However, these previous studies in [6]–[8] have not considered practical RIS environments and their limitation, such as practical reflection coefficients. In fact, an RIS can manipulate the properties of an incident wave based on the resonant frequency of the tunable reflecting circuit. Then, the incident EM power is partially consumed at the resistance of the reflecting circuit according to the difference between the incident wave frequency and the resonant frequency. This results in the amplitude of the reflection coefficients less than

This research was supported by the U.S. National Science Foundation under Grants CNS-1836802 and OAC-1638283.

978-1-7281-5089-5/20/$31.00 ©2020 IEEE

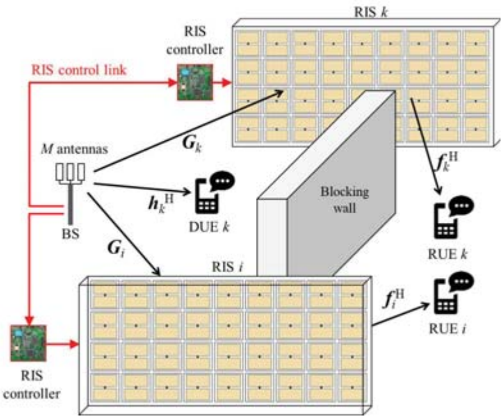

Fig. 1. Illustrative system model of the considered RIS-based MISO system.

or equal to one depending on the phase shifts of the incident wave. However, the work in [6] assumed an ideal RIS whose amplitude of the reflection coefficients are always equal to one which is impractical for an RIS. Moreover, the signals from the RIS transmitters proposed in [7] and [8] can be undesired interference for existing cellular network, given that those RISs operate as an underlay coexistence with cellular networks. Therefore, there is a need for new analysis of practical RISs when dealing with practical reflection coefficients that can verify the asymptotic optimality of realistic RISs.

The main contribution of this paper is a rigorous optimality analysis of the data rates that can be achieved by an RIS under consideration of practical reflection coefficients. In this regard, we first design a passive beamformer that achieves asymptotic signal-to-noise ratio (SNR) optimality, regardless of the reflection power loss. We then propose a new modulation scheme that can be used in an RIS to achieve sum-rate higher than the one achieved by a conventional network without RIS. In the proposed modulation scheme, each RIS utilizes an ambient RF signal, convert it into desired signal by controlling the properties of incident wave, and transmit it to the desired user, without interfering with existing users. We also prove that the achievable SNR from the proposed modulation converges to the asymptotic SNR resulting from a conventional massive MIMO or MIMO relay system, as the number of reflecting elements on an RIS increases. Our simulations show that the proposed schemes can asymptotically achieve the performance resulting from an ideal RIS and its upper bound.

The rest of this paper is organized as follows. Section II presents the RIS-based system model. Section III describes the optimality of achievable rate in downlink RIS system. Simulation results are provided in Section IV to support and verify the analyses, and Section V concludes the paper.

II. System Model

Consider a single BS multiple-input single-output (MISO) system that consists of a set $\mathcal{K}$ of K single-antenna user equipments (UEs) and multiple RISs each of which having N reflecting elements, as shown in Fig. 1. The BS is equipped with M antennas and serves one UE at each time slot based on a time-division multiple access (TDMA). Also, the BS

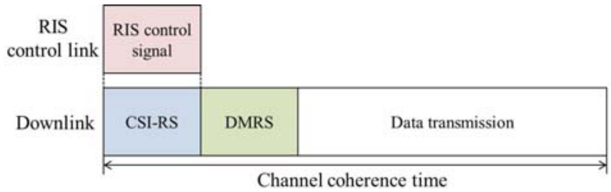

Fig. 2. Downlink frame structure at BS in considered RIS system.

transmits the downlink signal to scheduled UE through the transmit beamforming. In our system model, we consider two types of UEs: a) UEs directly connected to the BS (called DUEs) and b) UEs connected to the BS via an RIS (called RUEs). Each UE can measure the downlink channel quality information (CQI) and transmit this information to the BS as done in existing cellular systems [9]. For UEs whose CQI exceeds a pre-determined threshold, the BS will directly transmit downlink signals to these UEs (which are now DUEs) without using the RIS. When the CQI is below a pre-determined threshold (i.e., the direct BS-UE channel is poor), the BS will have to allocate, respectively, suitable RISs to those UEs (that become RUEs) experiencing this poor CQI and, then, send a control signal to each RIS controller via a dedicated control link. Given the received BS control signal, the RIS controller determines N bias direct-current (DC) voltages for all reflecting elements and then, the varactor capacitance can be controlled, resulting in phase shifts of the reflection wave. Note that an RIS cannot coherently align, simultaneously, with the desired channels of all RUEs which, in turn, limits system performance [2]–[4]. For densely located RISs, we assume that each RUE is connected to different RISs (i.e., one RUE per one RIS) depending on the location of each RUE. Also, given a practical range of mobility and carrier frequency, we consider that all channels are generated from quasi-static block fading whose coherence time covers the downlink transmission period, as shown in Fig. 2. In accordance with 3GPP LTE specification [10], we consider two types of reference signals (RSs): Channel state information-RS (CSI-RS) and demodulation RS (DMRS). A CSI-RS is used to estimate the CSI and report CQI back to the BS, and a DMRS is a beamformed RS used to estimate an effective CSI for demodulation [10]. In order to estimate accurate CSI, an RIS will not operate during the CSI-RS period and, simultaneously, the BS can send a control signal to the RIS controller via a dedicated control link during this period. Then, the RIS operates based on this control signal, and reflects the DMRS and data signal with controlled phase shifts. The RUE receives the phase-shifted DMRS and estimates the effective CSI, and eventually, the downlink signal can be decoded. Hereinafter, we assume that the channel state of each wireless link follows a stationary stochastic process under a perfect CSI at the BS and, hence, our analysis will result in a performance bound of practical channel estimation scenarios.

We divide the UE set into two sets, such that $\mathcal{K} = \mathcal{D} \cup \mathcal{R}$ where $\mathcal{D}$ is the set of DUEs and $\mathcal{R}$ is the set of RUEs. Then, the received signal at UE k is obtained as

$$y_k = \begin{cases} \sqrt{P}\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{w}_k x_k^{\mathrm{d}} + n_k^{\mathrm{d}}, \text{ if } k \in \mathcal{D}, \\ \sqrt{P}\boldsymbol{f}_k^{\mathrm{H}}\boldsymbol{\Phi}_k\boldsymbol{G}_k\boldsymbol{w}_k x_k^{\mathrm{r}} + n_k^{\mathrm{r}}, \text{ if } k \in \mathcal{R}, \end{cases} \quad (1)$$

where P is the BS transmit power and $\boldsymbol{h}_k \in \mathbb{C}^{M\times 1}$, $\boldsymbol{G}_k \in \mathbb{C}^{N\times M}$, and $\boldsymbol{f}_k \in \mathbb{C}^{N\times 1}$ are, respectively, the fading channels between the BS and DUE k, between the BS and RIS k, and between RIS k and RUE k. Also, $\boldsymbol{w}_k \in \mathbb{C}^{N\times 1}$ is the transmit beamforming vector and x_k^{d} and x_k^{r} are downlink transmit symbols for DUE k and RUE k, respectively, with noise terms $n_k^{\mathrm{d}} \sim \mathcal{CN}(0, N_0)$ and $n_k^{\mathrm{r}} \sim \mathcal{CN}(0, N_0)$. In (1), $\boldsymbol{\Phi}_k \in \mathbb{C}^{N\times N}$ is a reflection matrix (i.e., passive beamformer) that includes reflection amplitudes and phases resulting from N reflecting elements. This reflection matrix is controlled by the RIS control signal from the BS and then, $\boldsymbol{\Phi}_k$ can be obtained as follows:

$$\boldsymbol{\Phi}_k = \mathrm{diag}\left(A(\angle\Gamma_1)e^{j\angle\Gamma_1}, \cdots, A(\angle\Gamma_N)e^{j\angle\Gamma_N}\right). \quad (2)$$

In [11], $A(\angle\Gamma_n)$ is approximated when SMV1231-079 varactor is used at each RIS element with $f_{\mathrm{c}} = 2.4$ GHz, as follows:

$$A(\angle\Gamma_n) \approx 0.8\left(\frac{\sin(\angle\Gamma_n - 0.43\pi) + 1}{2}\right)^{1.6} + |\Gamma|_{\min},$$

where Γ_n, $|\Gamma_n|$, and $\angle\Gamma_n$ are the reflection coefficient, amplitude, and phase, respectively, and $|\Gamma|_{\min} = 0.2$ is the minimum reflection amplitude. Hence, the instantaneous SNR at UE k can be obtained as follows:

$$\gamma_k = \begin{cases} PE_k^{\mathrm{d}}\left|\boldsymbol{h}_k^{\mathrm{H}}\boldsymbol{w}_k\right|^2/N_0, & \text{if } k \in \mathcal{D}, \\ PE_k^{\mathrm{r}}\left|\boldsymbol{f}_k^{\mathrm{H}}\boldsymbol{\Phi}_k\boldsymbol{G}_k\boldsymbol{w}_k\right|^2/N_0, & \text{if } k \in \mathcal{R}, \end{cases} \quad (3)$$

where E_k^{d} and E_k^{r} are the average energy per symbol for DUE k and RUE k, respectively. Given this practical RIS model, our goal is to maximize (3) and eventually achieve (asymptotically) the SNR of an ideal RIS as $N \to \infty$. In most prior studies such as [6], the properties of the reflection wave, such as the frequency, amplitude, and phase, are assumed to be independently controlled, however, these properties are closely related to each other as discussed in [12]. Hence, their relationship should be considered in the system model to accurately verify the potential of practical RISs. Given this practical RIS model, we will propose a novel passive beamformer and a new modulation scheme that can achieve an ideal performance, asymptotically.

III. Optimality of the Achievable Downlink Rate in an RIS

We analyze the optimality of the achievable rate using practical RISs under consideration of the practical reflection coefficients, as N increases to infinity. As proved in [6], given an ideal RIS that reflects the incident wave without power loss, the downlink SNR of the RIS achieves, asymptotically, the order of $\mathcal{O}\left(N^2\right)$, as N increases to infinity. However, the downlink SNRs of a conventional massive MIMO or MIMO relay system, each of which equipped with N antennas, equally increase with $\mathcal{O}(N)$ as proved in [13]. This squared SNR gain of RIS will analytically result in twice as much performance as conventional array systems in terms of achievable rate, without additional radio resources. In order to prove the optimality of the achievable rate using practical RISs under the aforementioned limitation, we first design a passive beamformer that achieves the SNR order of $\mathcal{O}\left(N^2\right)$ asymptotically. We then design a modulation scheme which can be used in an RIS that uses ambient RF signals to transmit data without additional radio resource and achieves the asymptotic SNR in order of $\mathcal{O}(N)$ like a conventional massive MIMO (or MIMO relay) system.

Algorithm 1 Reflection Phase Selection Algorithm

1: **Initialization:** Select $\hat{m} = \underset{1\le m\le M}{\arg\max}\,\|\boldsymbol{g}_m\|$ and set $s_0 = 0$ and $i = 1$.
2: **Reflection phase selection:** $\hat{\theta} = \underset{\theta\in[-\pi,\pi]}{\arg\max}\,|s_{i-1} + a_{i\hat{m}}\phi(\theta)|$.
3: **Update reference vector:** $s_i = s_{i-1} + a_{i\hat{m}}\phi(\hat{\theta})$.
4: Select $\phi_i = \phi(\hat{\theta})$.
5: Set $i \leftarrow i+1$ and go to Step 2 until $i = N+1$.
6: Return $\hat{\boldsymbol{\Phi}}_k = \mathrm{diag}(\phi_1, \phi_2, \cdots, \phi_N)$.

A. Passive Beamformer Design

The maximum instantaneous SNR at DUE k can be achieved by using a maximum ratio transmission (MRT) where $\boldsymbol{w}_k = \boldsymbol{h}_k/\|\boldsymbol{h}_k\|$, which yields an SNR $\gamma_k = PE_k^{\mathrm{d}}\|\boldsymbol{h}_k\|^2/N_0$. We then formulate an optimization problem whose goal is to maximize instantaneous SNR at RUE k with respect to $\boldsymbol{\Phi}_k$ and $\boldsymbol{w}_k$, as follows:

$$\max_{\boldsymbol{\Phi}_k,\boldsymbol{w}_k} \frac{PE_k^{\mathrm{r}}\left|\boldsymbol{f}_k^{\mathrm{H}}\boldsymbol{\Phi}_k\boldsymbol{G}_k\boldsymbol{w}_k\right|^2}{N_0}, \quad (4)$$

$$\text{s.t. } |\Gamma|_{\min} \le A(\angle\Gamma_n) \le 1, \forall n, \quad (4a)$$

$$-\pi \le \angle\Gamma_n \le \pi, \forall n. \quad (4b)$$

For any given $\boldsymbol{\Phi}_k$, it is well known that the MRT precoder is the optimal solution to problem (4) such that $\boldsymbol{w}_k = \frac{\boldsymbol{G}_k^{\mathrm{H}}\boldsymbol{\Phi}_k^{\mathrm{H}}\boldsymbol{f}_k}{\left\|\boldsymbol{G}_k^{\mathrm{H}}\boldsymbol{\Phi}_k^{\mathrm{H}}\boldsymbol{f}_k\right\|}$ [14]. Then, we formulate an optimization problem with respect to $\boldsymbol{\Phi}_k$ as follows:

$$\max_{\boldsymbol{\Phi}_k} \frac{PE_k^{\mathrm{r}}\left\|\boldsymbol{f}_k^{\mathrm{H}}\boldsymbol{\Phi}_k\boldsymbol{G}_k\right\|^2}{N_0}, \quad \text{s.t. (4a), (4b).} \quad (5)$$

This problem is non-convex since $\left\|\boldsymbol{f}_k^{\mathrm{H}}\boldsymbol{\Phi}_k\boldsymbol{G}_k\right\|^2$ is not concave with respect to $\boldsymbol{\Phi}_k$. Let $\boldsymbol{f}_k = [f_1, \cdots, f_N]^{\mathrm{T}}$ and $\boldsymbol{G}_k = [\boldsymbol{g}_1, \cdots \boldsymbol{g}_M]$, where $\boldsymbol{g}_m \in \mathbb{C}^{N\times 1} = [g_{1m}, \cdots, g_{Nm}]^{\mathrm{T}}$ is the channel between BS antenna m and RIS k. Then, we have

$$\left\|\boldsymbol{f}_k^{\mathrm{H}}\boldsymbol{\Phi}_k\boldsymbol{G}_k\right\|^2 = \sum\nolimits_{m=1}^{M}\left|\sum\nolimits_{n=1}^{N} a_{nm}\cdot\phi(\angle\Gamma_n)\right|^2, \quad (6)$$

where $a_{nm} = |f_n||g_{nm}|e^{j(\angle f_n^* + \angle g_{nm})}$ and $\phi(\angle\Gamma_n) = A(\angle\Gamma_n)e^{j\angle\Gamma_n}$. Here, we propose Algorithm 1 that can achieve $\mathcal{O}\left(N^2\right)$ as $N \to \infty$, asymptotically. In this algorithm, we first select antenna $\hat{m}$ which has the largest channel gain among M channels between the BS and the RIS. Since a reflection amplitude is always 1 when its phase equals to $-\pi$, $\phi(-\pi)$ is selected as a reflection phase ϕ_1 and we determine $a_{1\hat{m}}\phi(-\pi)$ as a reference vector, s_1, in the first round. Note that $\left\|\boldsymbol{f}_k^{\mathrm{H}}\boldsymbol{\Phi}_k\boldsymbol{G}_k\right\|^2$ is calculated based on the sum of N vectors such as $\sum_{n=1}^{N} a_{nm}\phi\left(\angle\Gamma_n\right)$, as shown in (6). Therefore, when $i > 1$, we compare the Euclidean norm of vector additions between s_{i-1} and θ shifted candidates, i.e., $\left|s_{i-1} + a_{i\hat{m}}\phi\left(\theta\right)\right|, \forall\theta \in [-\pi, \pi]$, and select $\hat{\theta}$ with the maximum Euclidean norm. Therefore, we can derive the suboptimal solution such that $\phi_i = \phi\big(\hat{\theta}\big)$ for each reflecting element i. Algorithm 1 results in a suboptimal solution and will not achieve the optimal performance that can be obtained by the exhaustive search method. However, we can prove the following result related to the asymptotic optimality of Algorithm 1.

Proposition 1. Algorithm 1 can achieve an instantaneous SNR in order of $\mathcal{O}\left(N^2\right)$ as $N \to \infty$.

Proof: Due to space limitations, the detailed proof can be found in our technical report in [12]. ■

Proposition 1 shows that the instantaneous SNR resulting from Algorithm 1 can achieve the SNR order of an ideal RIS. Moreover, Algorithm 1 requires a complexity of $\mathcal{O}\left(N\right)$ resulting in simpler RIS control compared to the existing work on RIS in [11], as will be discussed in Section IV.

Next, we analyze the average SNR of a downlink RIS system, under consideration of the RIS reflection matrix derived from Algorithm 1. We assume that $\boldsymbol{f}_k \sim \mathcal{CN}\left(0, \boldsymbol{I}_N\right)$ and $\boldsymbol{g}_m \sim \mathcal{CN}\left(0, \boldsymbol{I}_M\right)$ considering Rayleigh fading channels. From (6), the instantaneous SNR at RUE k is obtained by

$$\gamma_k = \frac{PE_k^{\mathrm{r}}}{N_0}\sum_{m=1}^{M}\left|\sum_{n=1}^{N}\left|f_n\right|\left|g_{nm}\right| A\left(\angle\Gamma_n\right) e^{j\left(\angle\Gamma_n + \angle f_n^* + \angle g_{nm}\right)}\right|^2.$$

By selecting $\angle\Gamma_n = \theta_n = -\angle f_n^* - \angle g_{nm_0}$ for $0 \leq m_0 \leq M$ and $\forall n$, we have the following:

$$\frac{PE_k^{\mathrm{r}}}{N_0}\left\|\boldsymbol{f}_k^{\mathrm{H}}\hat{\boldsymbol{\Phi}}_k\boldsymbol{G}_k\right\|^2 \underset{(b)}{\geq} \alpha\left(\beta + \sum_{m\neq m_0}^{M}\left|\sum_{n=1}^{N}\left|f_n\right|\left|g_{nm}\right| e^{j\Delta g_{n\mathrm{d}}}\right|^2\right), \tag{7}$$

where $\alpha = \frac{PE_k^{\mathrm{r}}|\Gamma|_{\min}^2}{N_0}$, $\beta = \left|\sum_{n=1}^{N}\left|f_n\right|\left|g_{nm_0}\right|\right|^2$, and $\Delta g_{n\mathrm{d}} = \angle g_{nm} - \angle g_{nm_0}$. In (7), $\hat{\boldsymbol{\Phi}}_k$ and (b) result from Algorithm 1 and the minimum reflection amplitude, i.e., $A\left(\theta_n\right) \geq |\Gamma|_{\min}$, $\forall n$, respectively. We refer to the right hand side of (7) as an instantaneous SNR lower bound $\underline{\gamma}_k$. For notational convenience, we define $\underline{\gamma}_{\mathrm{l}} = \sum_{n=1}^{N}\left|f_n\right|\left|g_{nm_0}\right|$ and $\underline{\gamma}_{\mathrm{r},m} = \sum_{n=1}^{N}\left|f_n\right|\left|g_{nm}\right| e^{j\Delta g_{n\mathrm{d}}}$ in (7). Then, the random variable $\underline{\gamma}_k$ follows Lemma 1.

Lemma 1. Based on the central limit theorem, as $N \to \infty$, $\underline{\gamma}_k$ follows a non-central chi-squared distribution with one degree of freedom and its mean converges to

$$E\left[\underline{\gamma}_k\right] = \frac{N^2 PE_k^{\mathrm{r}}\pi^2\left|\Gamma\right|_{\min}^2}{16N_0}. \tag{8}$$

Proof: Due to space limitations, the detailed proof can be found in our technical report in [12]. ■

Lemma 1 shows that the lower bound of the average SNR increases with $\mathcal{O}\left(N^2\right)$ and the average SNR gains from $M-1$ antennas become negligible compared to those from m_0, as $N \to \infty$. Moreover, we can observe from (8) that this lower bound is equal to the single antenna case in [6]. Since $\mathcal{O}\left(N^2\right)$ can also be achieved by the instantaneous SNR resulting from Algorithm 1, as proved in Proposition 1, the average SNR gains from $M-1$ antennas are also negligible compared to those from m_0, as $N \to \infty$. Therefore, the average SNR resulting from Algorithm 1 will also converge to that of the single antenna case. Given this convergence of the average SNR, we can use $\boldsymbol{w}_k$ as an antenna selection that can achieve full multi-antenna diversity with a low-cost and low-complexity instead of a MRT. By selecting the BS antenna whose channel gain has the maximum value such as in Step 1 of Algorithm 1, we can determine the transmit precoding vector as follows:

$$\begin{cases} w_m = 1, \text{ if } m = \hat{m}, \\ w_m = 0, \text{ if } m \neq \hat{m}, \end{cases} \tag{9}$$

where $\hat{m} = \underset{1\leq m\leq M}{\arg\max}\left\|\boldsymbol{g}_m\right\|$. Although an MRT precoder achieves the optimal performance for a single-user MISO system, it requires multiple RF chains associated with multiple antennas resulting in higher cost and hardware complexity compared to the transmit antenna selection scheme. Since the average SNR will converge to the single antenna system as N increases, a large N results in a performance convergence between MRT and transmit antenna selection. Moreover, we can observe from Lemma 1 that the SER resulting from Algorithm 1 decays exponentially as a function of N, as proved in [6].

B. Modulation for Unscheduled RUE

Next, we devolop a modulation scheme that can be used to increase the achievable RIS sum-rate. In our downlink TDMA scheme, the BS can transmit the downlink signal to only one scheduled UE at each time slot. However, by reflecting the ambient RF signals generated from the BS, each RIS also can send the downlink signal to its unscheduled RUE at each time slot, as done in ambient backscatter communications [15]. Different from ambient backscatter communications, an RIS can convey information by reflecting the ambient RF signal without interfering with existing scheduled UEs. In addition to the data rates obtained from the BS's downlink DUEs, we can obtain additional data rates at unscheduled RUEs without additional radio resources, resulting in higher achievable sum-rate than a conventional network without RIS. For convenience, we refer to each unscheduled RUE and its connected RIS as uRUE and uRIS, respectively. As shown in [6] and [7], an RIS can transmit PSK signals to serving user by controlling its reflection phase. As such, we consider that each uRIS controls its reflection phase to send the downlink signal to each corresponding uRUE by reflecting the incident wave from the BS, whenever the DUE is scheduled. In order to avoid undesired interference from uRISs to the scheduled DUE, we consider the following procedure. First, the BS sends the RIS control signals, which are related to the data symbols for each uRUE, to uRISs during the CSI-RS period (see Fig. 2). Each uRIS controls N bias DC voltages for all reflecing elements depending on the control signal and reflects the incident wave from the BS with the same reflection phase during the DMRS and data transmission periods. Meanwhile, each scheduled DUE (sDUE) receives the DMRS from the BS and estimates the effective CSI which includes the transmit precoding at the BS and the phase shifts from the uRISs. Since all uRISs keep using the same reflection phase from the beginning of the DMRS to the end of the data transmission, this effective CSI will not change during the downlink data transmission and this results in zero interference. Similarly, each uRUE receives the CSI-RS from the BS and estimates the CSI between the BS and each uRUE without controlled phase shifts. Note that the BS broadcasts the information about the transmit precoder and the modulation of sDUE by using the downlink control indicator (DCI), based on the LTE specification in [16]. Given the estimated CSI and the broadcast DCI, the uRUE can calculate the effective CSI and eventually decode the downlink signal transmitted from the uRIS. By using this procedure, each uRIS needs to transmit only one symbol during the channel coherence time, however, the uRUEs can achieve the asymptotic SNR in order of $\mathcal{O}(N)$ as will be proved in Theorem 1. Moreover, since the uRUE receives the same symbols from the uRIS during the data transmission period, the uRUE will achieve a diversity gain proportional to the length of the data transmission period. Furthermore, since the BS can send the RIS control signal related to one symbol for each uRIS during the entire channel coherence time, we can reduce the link burden of the RIS control link. Given this procedure, the received downlink signal at uRUE i can be obtained by:

$$y_i = \sqrt{P}\boldsymbol{f}_i^{\mathrm{H}}\bar{\boldsymbol{\Phi}}_i\boldsymbol{G}_i\boldsymbol{w}_k x_k^{\mathrm{d}} + n_i^{\mathrm{r}}, \tag{10}$$

where $i \in \mathcal{R}$, $k \in \mathcal{D}$, and $\boldsymbol{w}_k$ and x_k^{r} are the transmit precoder and the downlink signal for sDUE k, respectively. To modulate the downlink data bits into the reflection matrix, we propose the following modulated reflection matrix for the uRUE i:

$$\bar{\boldsymbol{\Phi}}_i = A(\omega_i)\, e^{j\omega_i}\boldsymbol{I}_N,\ \omega_i \in \mathcal{M}, \tag{11}$$

where $e^{j\omega_i}$ is the proposed M_{o}-PSK modulation symbol for uRUE i and $\mathcal{M} = [\mu_1, \cdots, \mu_{2^{M_{\mathrm{o}}}}]$ is a set of the corresponding M_{o}-PSK modulation. For two examples of QPSK signaling, we have $\mathcal{M} = \left\{0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}\right\}$ when x_k^{d} is BPSK symbol and $\mathcal{M} = \left\{0, \frac{\pi}{8}, \frac{\pi}{4}, \frac{3\pi}{8}\right\}$ when x_k^{d} is QPSK symbol. When x_k^{d} is a QAM symbol, we can also obtain $\mathcal{M}$ according to the minimum angle between adjacent QAM symbols. Hence, using (10) and (11), we obtain:

$$y_i = \sqrt{P}\boldsymbol{f}_i^{\mathrm{H}}\boldsymbol{G}_i\boldsymbol{w}_k A(\omega_i)\, e^{j\omega_i} x_k^{\mathrm{d}} + n_i^{\mathrm{r}}. \tag{12}$$

Note that $\boldsymbol{f}_i^{\mathrm{H}}\boldsymbol{G}_i$ can be estimated from the CSI-RS, and the modulation scheme of x_k^{d} and $\boldsymbol{w}_k$ are known at the uRUE from the broadcasted DCI. Assuming that all RISs are equipped with the same passive elements resulting in identical reflection coefficient models, the uRUE can calculate $A(\omega_i)$ for $|\mathcal{M}|$ symbols and eventually estimate ω_i resulting in additional data rate in the considered RIS system. From (12), we can prove the following result related to the average SER at the uRUE.

Theorem 1. The uRUE achieves an average SNR in order of $\mathcal{O}(N)$ as $N \to \infty$ and the average SER with the proposed M_{o}-PSK signaling can be approximated by

$$P_{\mathrm{e}} = \frac{1}{2^{M_o}}\sum_{p=1}^{2^{M_o}}\int_0^{\pi - \frac{\Delta_\mu}{2}} \frac{1}{1 - \frac{N N_{\mathrm{s}} E_{\mathrm{s}} t(\theta) A(\mu_p)^2}{N_0}} d\theta, \tag{13}$$

where $t(\theta) = \frac{-\sin^2(\Delta_\mu/2)}{\sin^2\theta}$, Δ_μ is the angular spacing of the proposed M_{o}-PSK symbols, and N_{s} is the number of transmitted symbols during the downlink data transmission period.

Proof: Due to space limitations, the detailed proof can be found in our technical report in [12]. ∎

Theorem 1 shows that the uRUE can achieve an asymptotic SNR in order of $\mathcal{O}(N)$ by using the proposed modulation scheme resulting in several implications. First, an RIS can provide the same asymptotic SNR as a conventional massive MIMO or MIMO relay system for all unscheduled RUEs, simultaneously. For the scheduled RUEs, an RIS also provides an asymptotic SNR in order of $\mathcal{O}\left(N^2\right)$, which is much higher than that of conventional MIMO systems, as proved in Proposition 1. Hence, an RIS can support the demand for massive connectivity and high data traffic, without additional radio resources. Moreover, Theorem 1 shows that the average SER of uRUE can be obtained based on deterministic values and we can evaluate the reliability of the considered RIS system without extensive simulations. In particular, we can observe from (13) that the average SER decreases as N increases and it eventually reaches zero as $N \to \infty$, resulting in reliable communication regardless of E_{s}/N_0 even at the uRUEs.

IV. Simulation Results and Analysis

We run extensive simulations to assess the downlink performance, in terms of the ergodic rate and SER, under a practical-sized RIS environment with finite N. We assume that all channels are generated by Rayleigh fading, resulting in $\boldsymbol{f}_k \sim \mathcal{CN}(0, \boldsymbol{I}_N)$ and $\boldsymbol{g}_m \sim \mathcal{CN}(0, \boldsymbol{I}_M), \forall k, m$, where $M = 2$. Note that all numerical results are obtained from Monte Carlo simulations that are statistically averaged over a large number of independent runs.

Fig 3 shows the ratio between the ergodic rate at the RUE, R_k^{r}, resulting from Algorithm 1 and theoretical upper bound. The theoretical upper bound is derived as follows:

$$\hat{R}_k = \mathrm{E}\left[\log\left(1 + \frac{P}{N_0}\sum_{m=1}^{M}\left|\sum_{n=1}^{N}|f_n|\,|g_{nm}|\right|^2\right)\right]. \tag{14}$$

We compare the results with the AO algorithm proposed in [11] which is shown up to $N = 500$ due to its computational complexity. As shown in Fig. 3, the ergodic rate ratios resulting from the proposed scheme increase toward 1 as N increases, verifying the optimality of Algorithm 1 as proved in Proposition 1. Although the AO algorithm can also achieve the upper bound performance, asymptotically, it requires very high complexity. In the AO algorithm, a

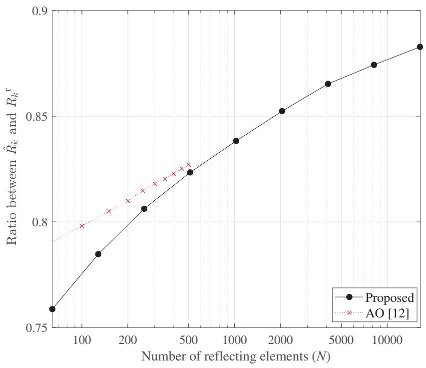

Fig. 3. Performance comparison of the ergodic rates resulting from Algorithm 1 with different b values when $PE_k^{\mathrm{r}}/N_0 = 1$.

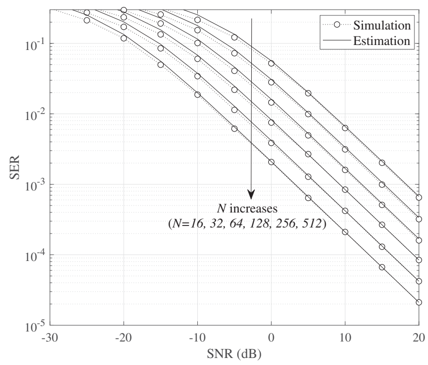

Fig. 4. Performance comparison of the average SERs resulting from the proposed modulation with different N values.

complexity in order of $\mathcal{O}\left(N^2\right)$ is required for each iteration and it continues until convergence. However, the proposed Algorithm 1 requires $\mathcal{O}\left(N\right)$ resulting in a much simpler operation at the BS especially for a large N. Moreover, the AO algorithm uses MRT which requires multiple RF chains resulting in higher cost and hardware complexity compared to the proposed algorithm.

In Fig. 4, Theorem 1 is verified in the following scenario. The BS transmits BPSK signals to the sDUE via a wireless channel, $\boldsymbol{h}_k$, and also sends the RIS control signals related to the data symbols for the uRIS via a dedicated RIS control link. Data symbols for the uRIS are modulated based on the proposed modulation technique assuming the BPSK signaling (i.e., $M_0 = 2$). As shown in Fig. 4, the asymptotic SERs derived from Theorem 1 are close to the results of our simulations. Moreover, the SERs linearly decrease as N increases given that the SNR difference is always equal to 3 dB when N is doubled. For instance, when the target SER is $2 \cdot 10^{-2}$, the corresponding SNRs are 5, 2, -1, -4, -7, -10 dB for $N = 16, 32, 64, 128, 256, 512$, respectively. This result shows that the SER can be reduced by increasing N and eventually it converges to zero as $N \to \infty$.

V. Conclusions

In this paper, we have asymptotically analyzed the optimality of the achievable rate using practical RISs in presence of limitations such as practical reflection coefficients. In particular, we have designed a passive beamformer that can achieve the asymptotic optimal SNR under discrete reflection phases with a practical reflection power loss, and shown that it achieves a SNR optimality. We have also proposed a modulation scheme that can be used in a downlink RIS system resulting in higher achievable sum-rate than a conventional network without RIS. Moreover, we have derived the approximated SER of the proposed modulation scheme, showing that it achieves an asymptotic SNR of a conventional massive array systems such as a massive MIMO or MIMO relay system. Simulation results have shown that the result of our algorithm converges to the asymptotic upper bound as $N \to \infty$ and the approximated SER is in close agreement with the result from our simulations. Therefore, we expect that our approach will provide an important solution for future wireless networks achieveing optimality and supporting massive connectivity.

References

[1] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," *IEEE Network*, to appear, 2020.

[2] M. Jung, W. Saad, Y. Jang, G. Kong, and S. Choi, "Reliability analysis of large intelligent surfaces (LISs): Rate distribution and outage probability," *IEEE Wireless Communications Letters*, vol. 8, no. 6, pp. 1662–1666, Dec. 2019.

[3] M. Jung, W. Saad, Y. Jang, G. Kong, and S. Choi, "Performance analysis of large intelligent surfaces (LISs): Asymptotic data rate and channel hardening effects," *IEEE Transactions on Wireless Communications*, to appear, 2020.

[4] M. Jung, W. Saad, and G. Kong, "Large intelligent surfaces (LISs): Uplink spectral efficiency and pilot training," *available online: arxiv.org/abs/1904.00453*, Mar. 2019.

[5] T. J. Cui, S. Liu, and L. Zhang, "Information metamaterials and metasurfaces," *Journal of Materials Chemistry C*, vol. 5, pp. 3644–3668, 2017.

[6] E. Basar, M. D. Renzo, J. de Rosny, M. Debbah, M. S. Alouini, and R. Zhang, "Wireless communications through reconfigurable intelligent surfaces," *IEEE Access*, vol. 7, pp. 116 753–117 773, Sep. 2019.

[7] W. Tang, J. Y. Dai, M. Chen, X. Li, Q. Cheng, S. Jin, K. Wong, and T. J. Cui, "Programmable metasurface-based RF chain-free 8PSK wireless transmitter," *Electronics Letters*, vol. 55, no. 7, pp. 417–420, 2019.

[8] W. Tang, X. Li, J. Y. Dai, S. Jin, Y. Zeng, Q. Cheng, and T. J. Cui, "Wireless communications with programmable metasurface: Transceiver design and experimental results," *available online: arxiv.org/abs/1811.08119*, Nov. 2018.

[9] 3rd Generation Partnership Project, "Technical Specification Group Radio Access Network; User Equipment (UE) conformance specification; Radio transmission and reception; Part 1: Conformance testing," TS 36.521-1, V14.3.0, Aug. 2017.

[10] ——, "Technical Specification Group Radio Access Network; Physical layer procedures for data," TS 38.214, V15.2.0, Jul. 2018.

[11] S. Abeywickrama, R. Zhang, and C. Yuen, "Intelligent reflecting surface: Practical phase shift model and beamforming optimization," *available online: arxiv.org/abs/1907.06002*, Aug. 2019.

[12] M. Jung, W. Saad, M. Debbah, and C. S. Hong, "On the optimality of reconfigurable intelligent surfaces (RISs): Passive beamforming, modulation, and resource allocation," *available online: arxiv.org/abs/1910.00968*, Oct. 2019.

[13] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Transactions on Communications*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[14] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge Univ. Press, 2005.

[15] V. C. Liu, A. N. Parks, V. Talla, S. Gollakota, D. Wetherall, and J. S. R. Smith, "Ambient backscatter: wireless communication out of thin air," in *ACM SIGCOMM*, 2013.

[16] 3rd Generation Partnership Project, "Technical Specification Group Radio Access Network; Physical layer procedures," TS 36.213, V14.6.0, Apr. 2018.