

# Reconfigurable Intelligent Surface Assisted Multi-User Communications: How Many Reflective Elements Do We Need?

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

**Abstract**—Reconfigurable intelligent surfaces (RISs) consisting of multiple reflective elements are a promising technique to enhance communication quality as they can create favorable propagation conditions. In this letter, we characterize the fundamental relations between the number of reflective elements and the system sum-rate in RIS-assisted multi-user communications. It is known from previous works that the received signal-to-noise ratio (SNR) can linearly increase with the squared number of RIS reflective elements, but how many elements are sufficient to provide an acceptable system sum-rate still remains an open problem. To this end, we derive the asymptotic capacity with zero-forcing (ZF) precoding, and then analyze how many reflective elements are required so that the ratio of the system sum-rate to the capacity can exceed a predefined threshold. Numerical results verify our analysis.

**Index Terms**—Reconfigurable intelligent surface, multi-user communications, reflective elements.

## I. INTRODUCTION

RECONFIGURABLE intelligent surfaces (RISs) have emerged as a cost-effective solution to address the need for increased capacity that is expected for future wireless systems [1]. By controlling the electromagnetic response (e.g., phase shifts) for a large number of low-cost elements, an RIS can provide a favorable propagation environment between transmitters and receivers, thus improving the link quality [2].

In the literature, most existing works focus on the phase shift optimization/analysis in RIS assisted wireless communications. In [3], the authors proposed a hybrid beamforming

Manuscript received December 20, 2020; accepted February 7, 2021. Date of publication February 11, 2021; date of current version May 10, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 61625101, Grant 61829101, and Grant 61941101; and in part by the U.S. National Science Foundation under Grant EARS-1839818, Grant CNS-1717454, Grant CNS-1731424, Grant CNS-1702850, and Grant CCF-1908308. The associate editor coordinating the review of this article and approving it for publication was A. Shojaeifard. (*Corresponding author: Lingyang Song.*)

Hongliang Zhang is with the Department of Electronics, Peking University, Beijing 100871, China, and also with the Department of Electrical Engineering, Princeton University, NJ 08544 USA (e-mail: hongliang.zhang92@gmail.com).

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

Zhu Han is with the Electrical and Computer Engineering Department, University of Houston, Houston, TX 77004 USA, and also with the Department of Computer Science and Engineering, Kyung Hee University, Seoul 446-701, South Korea (e-mail: zhuhan22@gmail.com).

H. Vincent Poor are with the Department of Electrical Engineering, Princeton University, NJ 08544 USA (e-mail: poor@princeton.edu).

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

Digital Object Identifier 10.1109/LWC.2021.3058637

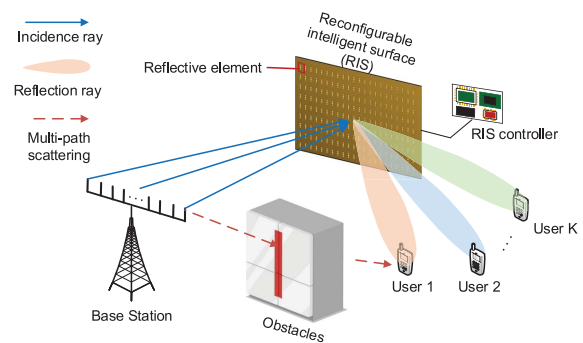


Fig. 1. System model of an RIS assisted downlink multi-user MISO network.

scheme for a multi-user RIS assisted multi-input single-output (MISO) system together with a phase shift optimization algorithm to maximize the system sum-rate. The authors in [4] optimized the power allocation at the base station (BS) as well as phase shifts at the RIS to maximize the data rate. The authors in [5] analyzed the effect of the limited phase shifts on the system data rate. As suggested in [5], the size of the RIS will also influence the system sum-rate. However, how many RIS reflective elements are sufficient to provide an acceptable system sum-rate remains an open problem.

In this letter, we consider an RIS assisted downlink multi-user cellular network. To quantify the impact of the number of RIS reflective elements on the system sum-rate, we first provide an asymptotic analysis of the system capacity for the RIS-assisted downlink multi-user MISO communications with zero-forcing (ZF) precoding which is easy to implement. Based on this capacity analysis, we further analyze how many RIS reflective elements are sufficient to provide an acceptable system sum-rate.

**Notation:** Boldface lower and upper case symbols denote vectors and matrices, respectively.  $\|\cdot\|_F$  represents the Frobenius norm.  $\mathbb{C}^{M \times N}$  denotes a complex matrix with dimensions  $M \times N$ . Conjugate transpose and matrix inverse operators are denoted by  $(\cdot)^H$  and  $(\cdot)^{-1}$ , respectively. We use  $I_K$  to denote the  $K \times K$  identity matrix.  $\mathbb{E}[\cdot]$ ,  $\text{Tr}\{\cdot\}$ , and  $\det(\cdot)$  are expectation, trace, and determinant operators, respectively. We also use  $[A]_{k,m}$  to represent the element in the  $k$ -th row and the  $m$ -th column of a matrix  $A$ .

## II. SYSTEM MODEL

### A. Scenario Description

Consider a narrow-band downlink multi-user MISO network as shown in Fig. 1, where one BS with  $M$  antennas serves  $K$

single-antenna users, where  $M \geq K$ . Due to the dynamic wireless environment involving unexpected fading and potential obstacles, the Line-of-Sight (LoS) link between the cellular users and the BS may not be stable or even falls into a complete outage.<sup>1</sup> To enhance the QoS of the communication link, an RIS is adopted to reflect the signal from the BS and directly project the signals to the users by actively shaping the propagation environment into a desirable form.

The RIS is composed of  $N$  electrically controlled RIS reflective elements with the size of  $a \times b$ . Each reflective element can adjust the phase shift by leveraging positive-intrinsic-negative (PIN) diodes. A PIN diode can be switched between ‘‘ON’’ and ‘‘OFF’’ states, based on which the metal plate can add a different phase shift to the reflected signal. Define  $\theta_n$  as the phase shift for reflective element  $n$ , and the reflection factor of reflective element  $n$  can be written by  $\Gamma_n = \Gamma e^{-j\theta_n}$ , where the reflection amplitude  $\Gamma \in [0, 1]$  is a constant.

### B. Channel Model

Let  $\mathbf{G} \in \mathbb{C}^{K \times M}$  be the channel matrix between the BS and users, where  $g^{km} \triangleq [\mathbf{G}]_{km}$  is the channel coefficient between user  $k$  and antenna  $m$  at the BS. Since each reflective element will reflect the signals from the BS to the users, the channel  $g_{km}$  consists of  $N$  paths, and we denote the channel gain from antenna  $m$  to user  $k$  through the  $n$ -th reflective element as  $g_n^{km}$ .

The channel matrix  $\mathbf{G}$  models independent fast fading, path loss, and RIS response. To be specific,  $g_n^{km}$  can be written by

$$g_n^{km} = (\beta_n^k)^{-1/2} h_n^{km} \Gamma_n. \quad (1)$$

Here,  $h_n^{km}$  is independent fast fading coefficient with  $\mathbb{E}[h_n^{km}] = 0$  and  $\mathbb{E}[(h_n^{km})^2] = 1$ . Based on the results in [6], the path loss is related with the incident angle  $\alpha_n^k$ , i.e.,  $\beta_n^k = \frac{64\pi^3}{Aab\lambda^2} \frac{(l_n d_n^k)^2}{\cos^3(\alpha_n^k)}$ , where  $A$  is the antenna gain,  $\lambda$  is the wavelength,  $l_n$  is the distance between the BS and the  $n$ -th reflective element,<sup>2</sup> and  $d_n^k$  is the distance between user  $k$  and the  $n$ -th reflective element. Therefore, we have

$$g^{km} = \sum_{n=1}^N g_n^{km} = (\bar{\beta}^k)^{-1/2} \sum_{n=1}^N h_n^{km} \Gamma_n, \quad (2)$$

where  $\bar{\beta}^k$  is the equivalent path loss between the BS and user  $k$  through the RIS. Therefore, channel matrix  $\mathbf{G}$  can be written by

$$\mathbf{G} = \mathbf{B}^{-1/2} \mathbf{H}, \quad (3)$$

where  $\mathbf{B}$  is a diagonal matrix whose  $k$ -th diagonal reflective element being  $\bar{\beta}^k$  and the reflective element at the  $k$ -th column and the  $m$ -th row of  $\mathbf{H}$  is equal to  $\sum_{n=1}^N h_n^{km} \Gamma_n$ .

### C. Achievable Rate With Zero-Forcing Precoding

To serve  $K$  users simultaneously, the BS first encodes the data symbols for different users with a normalized precoding matrix  $\mathbf{W} \in \mathbb{C}^{M \times K}$ . In this letter, we adopt the ZF precoder

<sup>1</sup>The RIS is more likely to be deployed in the cases where the LoS link does not exist to further extend the coverage.

<sup>2</sup>We assume that the distance between the BS and the RIS is much larger than the margin of two antennas, and thus the distances between the RIS and different antennas are assumed to be the same.

as it can obtain a near-optimal solution with a low complexity [3]. At the BS, the signals are assumed to be transmitted under a normalized power allocation matrix  $\mathbf{\Lambda}$  which satisfies  $\text{Tr}\{\mathbf{\Lambda}\mathbf{\Lambda}^H\} = 1$ . Denote the intended signal vector for  $K$  users as  $\mathbf{s}$  satisfying  $\mathbb{E}[\mathbf{s}\mathbf{s}^H] = \frac{P}{K} \mathbf{I}_K$ , where  $P$  is the transmitted power. Therefore, the transmitted signals at the BS can be given by  $\mathbf{x} = \mathbf{W}\mathbf{\Lambda}\mathbf{s}$ . Due to ZF precoding, after the reflection by the RIS, the received signal at the users can be written by

$$\mathbf{y} = \mathbf{G}\mathbf{W}\mathbf{\Lambda}\mathbf{s} + \mathbf{n}, \quad (4)$$

where the  $k$ -th reflective element of  $\mathbf{y}$  is the received signal for user  $k$  and  $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma^2)$  is the additive white Gaussian noise. Assume that the channel information is available for the transmitter. For ZF precoding, the  $k$ -th column of  $\mathbf{W}$  can be written by

$$\mathbf{w}_k = \frac{\mathbf{v}_k}{\|\mathbf{v}_k\|_F}, \quad (5)$$

where  $\mathbf{v}_k$  is the  $k$ -th column of matrix  $\mathbf{V}$  where

$$\mathbf{V} = \mathbf{G}^H (\mathbf{G}\mathbf{G}^H)^{-1}. \quad (6)$$

Consequently, with the ZF precoding, the signal-to-noise ratio (SNR) for user  $k$  can be expressed as

$$\gamma_k = \frac{P\Lambda_k}{K\sigma^2} (\mathbf{g}_k \mathbf{w}_k)(\mathbf{g}_k \mathbf{w}_k)^H, \quad (7)$$

where  $\mathbf{g}_k$  is the  $k$ -th row of  $\mathbf{G}$  and  $\Lambda_k = [\mathbf{\Lambda}\mathbf{\Lambda}^H]_{k,k}$ . Here, we assume that the channel state information is perfectly obtained by the BS, and thus  $\mathbf{\Lambda}$  is assumed to be a constant in the following which can be obtained by water-filling algorithm [3].

Use above notations, we can express the data rate for user  $k$  as

$$r_k = \mathbb{E}[\log_2(1 + \gamma_k)]. \quad (8)$$

## III. ANALYSIS ON ASYMPTOTIC CAPACITY

According to [7], the system capacity can be achieved by dirty paper coding, where

$$C = \mathbb{E} \left[ \sum_k \log_2 \left( 1 + \frac{P\Lambda_k}{K\sigma^2} [\mathbf{G}\mathbf{G}^H]_{k,k} \right) \right]. \quad (9)$$

In the following, we will provide a proposition to show the upper bound of the system capacity. Before this position, we first give a proposition on the channel matrix.

*Proposition 1:* When the number of reflective element  $N$  is large and the phase shifts of the RIS are given, i.e.,  $\theta_n$  is constant, we have<sup>3</sup>

$$\frac{\sum_{n=1}^N h_n^{km} \Gamma_n}{\Gamma \sqrt{N}} \sim \mathcal{CN}(0, 1). \quad (10)$$

*Proof:* Note that the fast fading for these channels is assumed to be independent. When the response of each reflective element is given, we have  $\mathbb{E}[\Gamma_n h_n^{km}] = 0$  since  $\mathbb{E}[h_n^{km}] = 0$ . Besides,  $\mathbb{E}[(\Gamma_n h_n^{km})^2] = \Gamma^2$  holds as  $\mathbb{E}[(h_n^{km})^2] = 1$  and  $|\Gamma_n| = \Gamma$ . Therefore, according to the central limit

<sup>3</sup>It is worthwhile to point out that this proposition holds with any distribution of  $h_n^{km}$ .

theorem [8], we have (10) when  $N$  is a sufficiently large number. ■

*Remark 1:* Proposition 1 implies that phase shifts of the RIS will not influence the distribution of the SNR when the number of reflective elements  $N$  is sufficiently large. This observation shows that the expectation of the SNR will keep unchanged even when we select phase shifts of the RIS randomly.

From this proposition, we can learn that the columns of  $\mathbf{H}$  are zero-mean independent Gaussian vectors. Based on this, we can have the following proposition.

*Proposition 2:* With the optimal power allocation, when the number of reflective elements is large but finite, the system capacity can be upper bounded by

$$C \leq \sum_k \log_2 \left( 1 + \frac{P\Lambda_k}{K\sigma^2} \bar{\beta}^k \Gamma^2 MN \right). \quad (11)$$

*Proof:* The proof is given in Appendix A. ■

*Remark 2:* The upper bound can be achieved when the eigenvalues of  $\mathbf{G}\mathbf{G}^H$  are the same. This can be obtained by optimizing the phase shifts of the RIS to maximize the effective rank of  $\mathbf{G}$  if the problem has a feasible solution [10].

It is worthwhile to point out that the system capacity will not increase infinitely. When the number of the reflective elements grows, the equivalent channel coefficient will decrease as well. In the following, we will give a proposition on the path loss  $\bar{\beta}^k$  at the extreme case, i.e.,  $N \rightarrow \infty$ .

*Proposition 3:* Without loss of generality, we assume that the RIS is on the plane with  $z = 0$ , and we assume that the coordinates of the BS and user  $k$  are  $(-x_0, 0, z_0)$  and  $(x_0, 0, z_k)$ , respectively, as illustrated in Fig. 2. When  $N \rightarrow \infty$ , we have  $\bar{\beta}_k N \approx \frac{2Az_0^3\lambda^2}{5\pi^2(z_0+z_k)^5}$ .

*Proof:* The detailed proof is given in Appendix B. ■

Based on the above two propositions, we have the following proposition.

*Proposition 4:* When the number of reflective elements goes infinity, i.e.,  $N \rightarrow \infty$ , the system capacity cannot exceed

$$\tilde{C} = \sum_k \log_2 \left( 1 + \frac{P\Lambda_k \Gamma^2 M}{K\sigma^2} \frac{2Az_0^3\lambda^2}{5\pi^2(z_0+z_k)^5} \right). \quad (12)$$

#### IV. ANALYSIS ON THE NUMBER OF REFLECTIVE ELEMENTS

In the following, we will discuss the effect of the number of reflective elements on the system sum-rate. We first give a proposition on the data rate for each user, and based on this, we will try to investigate how many reflective elements will be needed so that the sum-rate reach  $\eta$  of the system capacity.

*Proposition 5:* The data rate for user  $k$  can be rewritten by

$$r_k = \mathbb{E} \left[ \log_2 \left( 1 + \frac{P\Lambda_k}{K\sigma^2 \left[ (\mathbf{G}\mathbf{G}^H)^{-1} \right]_{k,k}} \right) \right]. \quad (13)$$

*Proof:* The proof is given in Appendix C. ■

With this proposition, the sum rate can be written by  $R = \sum_k r_k$ . Define  $\epsilon$  as the ratio of the sum-rate to the system

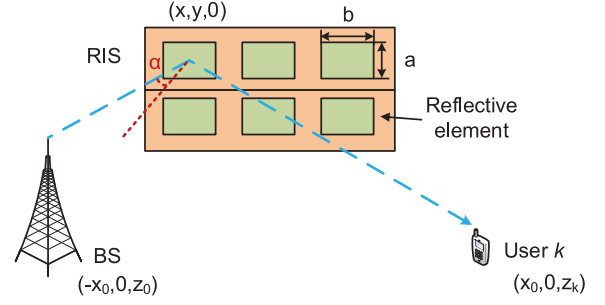


Fig. 2. Channel model on the RIS assisted downlink communications.

capacity with infinite reflective elements, i.e.,

$$\epsilon = R/\tilde{C}. \quad (14)$$

Therefore, the problem can be written by

$$\min_N N, \quad s.t. \quad \epsilon \geq \eta. \quad (15)$$

However, the aforementioned problem is hard to solve since the phase shifts of the reflective elements are not determined. In what follows, we first derive a lower bound of  $\epsilon$ , i.e.,  $\hat{\epsilon}$ , to guarantee  $\epsilon \geq \eta$ , and then transform the original problem into

$$\min_N N, \quad s.t. \quad \hat{\epsilon} \geq \eta. \quad (16)$$

#### A. Lower Bound of $\epsilon$

The lower bound of the ratio of the sum-rate to the system capacity  $\epsilon$  can be given by the proposition below.

*Proposition 6:*  $\epsilon$  can be lower bounded by

$$\hat{\epsilon} = \frac{\sum_k \log_2 \left( 1 + \frac{P}{\sigma^2} \Lambda_k \Gamma^2 \bar{\beta}^k N (\mu - 1) \right)}{\sum_k \log_2 \left( 1 + \frac{P}{\sigma^2} \Lambda_k \Gamma^2 \mu \frac{2Az_0^3\lambda^2}{5\pi^2(z_0+z_k)^5} \right)}, \quad (17)$$

where  $\mu = M/K$ .

*Proof:* According to the Jensen's inequality, we have

$$R \geq \sum_k \log_2 \left( 1 + \frac{P\Lambda_k}{K\sigma^2 \mathbb{E} \left[ \left[ (\mathbf{G}\mathbf{G}^H)^{-1} \right]_{k,k} \right]} \right). \quad (18)$$

Here,

$$\left[ (\mathbf{G}\mathbf{G}^H)^{-1} \right]_{k,k} = \frac{\mathbb{E} \left[ \text{Tr} \{ (\mathbf{H}\mathbf{H}^H)^{-1} \} \right]}{K\bar{\beta}^k}. \quad (19)$$

Note that  $\frac{1}{N\Gamma^2} \mathbf{H}\mathbf{H}^H$  is a central Wishart matrix with  $M$  degrees of freedom. Therefore, according to the results in [9], we have

$$\mathbb{E} \left[ \text{Tr} \{ (\mathbf{H}\mathbf{H}^H)^{-1} \} \right] = \frac{K}{N\Gamma^2(M-K)}. \quad (20)$$

Based on these, we have

$$\begin{aligned} \epsilon &\geq \frac{\sum_k \log_2 \left( 1 + \frac{P\Lambda_k}{K\sigma^2} \Gamma^2 \bar{\beta}^k N (M-K) \right)}{\sum_k \log_2 \left( 1 + \frac{P\Lambda_k \Gamma^2 M}{K\sigma^2} \frac{2Az_0^3\lambda^2}{5\pi^2(z_0+z_k)^5} \right)}, \\ &= \frac{\sum_k \log_2 \left( 1 + \frac{P}{\sigma^2} \Lambda_k \Gamma^2 \bar{\beta}^k N (\mu - 1) \right)}{\sum_k \log_2 \left( 1 + \frac{P}{\sigma^2} \Lambda_k \Gamma^2 \mu \frac{2Az_0^3\lambda^2}{5\pi^2(z_0+z_k)^5} \right)}. \end{aligned} \quad (21)$$

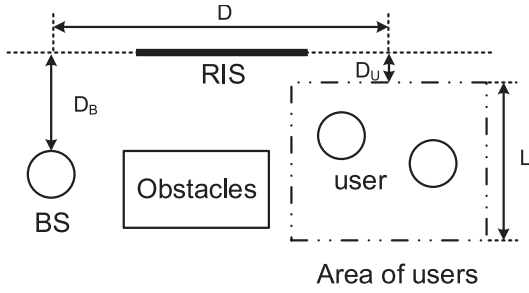


Fig. 3. Simulation layout for the RIS-based multi-user cellular network (top view).

which ends the proof. ■

*Remark 3:* From the lower bound of  $\epsilon$ , we can infer that a moderate number of antennas are required at the BS side to achieve an acceptable performance. Especially, when the number of antennas at the BS is equal to the number of users, we have  $\mu = 1$ , which leads to the lower bound of  $\epsilon$  being 0.

### B. Solution of Problem (16)

To satisfy the constraint in (16), we have

$$\prod_k \frac{K + \frac{P}{\sigma^2} \Lambda_k \Gamma^2 \bar{\beta}^k N (\mu - 1)}{\left( K + \frac{P}{\sigma^2} \Lambda_k \Gamma^2 \mu \frac{2A z_0^3 \lambda^2}{5\pi^2 (z_0 + z_k)^5} \right)^\eta} \geq 1. \quad (22)$$

We assume that the received SNR is high, i.e.,  $\frac{P}{\sigma^2} \Lambda_k \Gamma^2 \bar{\beta}^k N (\mu - 1) \gg K$  and  $\frac{P}{\sigma^2} \Lambda_k \Gamma^2 \mu \gg K$ . Thus, we have

$$\prod_k \frac{\frac{P}{\sigma^2} \Lambda_k \Gamma^2 \bar{\beta}^k N (\mu - 1)}{\left( \frac{P}{\sigma^2} \Lambda_k \Gamma^2 \mu \frac{2A z_0^3 \lambda^2}{5\pi^2 (z_0 + z_k)^5} \right)^\eta} \geq 1. \quad (23)$$

In other words,

$$N \geq \left( \prod_k \frac{\left( \frac{P}{\sigma^2} \Lambda_k \Gamma^2 \mu \frac{2A z_0^3 \lambda^2}{5\pi^2 (z_0 + z_k)^5} \right)^\eta}{\frac{P}{\sigma^2} \Lambda_k \Gamma^2 \bar{\beta}^k (\mu - 1)} \right)^{1/K}. \quad (24)$$

## V. SIMULATION RESULTS

In this section, we verify the derivation of the achievable asymptotic capacity and evaluate the impact of the number of reflective elements on the data rate in the RIS assisted multi-user communications, the layout of which is given in Fig. 3. The parameters are selected according to 3GPP standard [11] and existing works [5]. The height of the BS is 25m and the distance between the RIS and the BS  $D_B = 100$ m. We set the number of users  $K = 5$ . The users are uniformly located in a square area whose side length is set as  $L = 100$  m. The distance between the RIS and the closest side of the square area is  $D_u = 10$  m, and the horizontal distance between the BS and the center of the square area is  $D = 100$  m. The center of the RIS is located at the middle between the BS and the square area with the height being 25m. The working frequency of the RIS is set as  $f = 5.9$  GHz and reflection amplitude is assumed to be  $\Gamma = 1$ . The length and width of a reflective element are set the same, i.e.,  $a = b = 0.02$ m. Transmit power is set as  $P = 46$  dBm, noise power is set as  $\sigma^2 = -96$  dBm, and

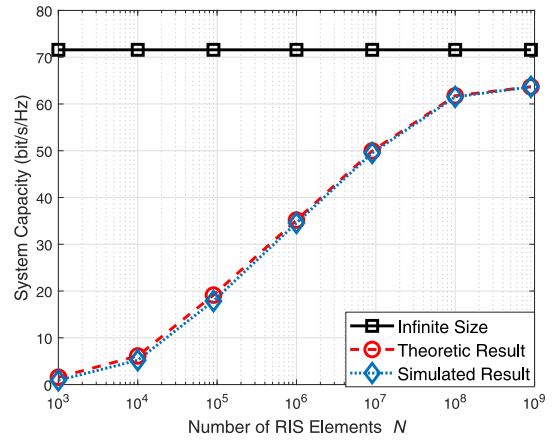


Fig. 4. System capacity vs. Number of reflective elements  $N$  with the number of antennas  $M = 10$ .

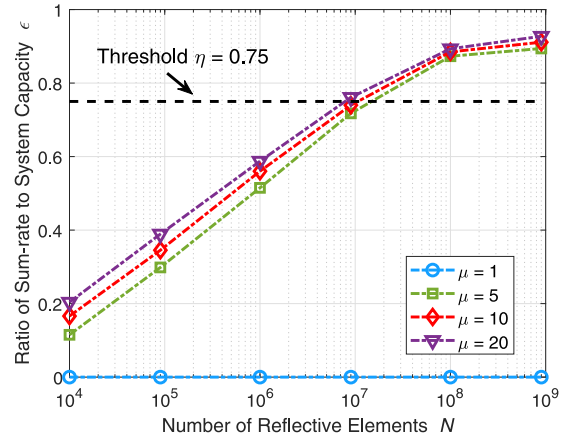


Fig. 5. Ratio of sum-rate to system capacity  $\hat{\epsilon}$  vs. Number of reflective Elements  $N$ .

antenna gain is  $A = 0$  dB. All numeral results are obtained by 100 Monte Carlo simulations.

In Fig. 4, we plot the system capacity vs. the number of reflective elements  $N$  with the number of antennas  $M = 10$ . From this figure, we can observe that the theoretic result is a tight approximation of the simulated one, and the gap will decrease as the number of reflective elements increases because it can make the singular values of the channel matrix more equal. Moreover, we can observe that the capacity per user will increase as the number of reflective elements  $N$  increases first and then become saturated, as suggested by Proposition 4. We can observe that the system capacity with an infinite size of the RIS can reach 71.5 bit/s/Hz for 5 users.

In Fig. 5, we plot the ratio of sum-rate to system capacity  $\hat{\epsilon}$  vs. the number of reflective elements  $N$  with threshold  $\eta = 0.75$  for different values of  $\mu$ . From this figure, we can observe that the ratio  $\hat{\epsilon}$  first increases linearly with  $\log N$  and then saturates, which suggests that it is more cost-effective to operate the RIS on the linear growth zone, i.e.,  $N < 10^7$ . We can also observe that it requires the number of reflective elements  $N = 8 \times 10^6$  to achieve 75% of the system capacity with  $\mu = 20$ . In other words, the side length of the RIS should be 56.7m. Besides, we can observe that the required number of reflective elements will decrease with a higher value of  $\mu$ . This implies that the size of the RIS can be reduced with

more antennas at the BS. Therefore, there exists an trade-off between the number of reflective elements and the number of antennas at the BS, and we can reduce the total cost by optimizing the size of the RIS. Moreover, we can find that the ratio remains 0 when  $\mu = 1$ , i.e., the number of antennas at the BS is the same as the number of users. This implies that the antennas at the BS should be larger than the number of users, which is consistent with remark 3.

## VI. CONCLUSION

In this letter, we have investigated RIS assisted downlink multi-user communications and have derived the corresponding asymptotic capacity. Based on the derived capacity, we also have analyzed the impact of the number of the reflective elements on the system sum-rate. Particularly, we have obtained a requirement on the RIS size to achieve an acceptable data rate. From the analysis and simulation, given the location of the RIS, we can draw the following conclusions: 1) The system capacity cannot increase indefinitely as the RIS size grows, but rather is upper bounded; and 2) To achieve a data rate threshold, the required number of reflective elements will decrease with more antennas at the BS.

### APPENDIX A

#### PROOF OF PROPOSITION 2

According to Jensen's inequality, we have

$$\begin{aligned} C &\leq \sum_k \log_2 \left( 1 + \frac{P\Lambda_k}{K\sigma^2} \mathbb{E} \left[ [\mathbf{G}\mathbf{G}^H]_{k,k} \right] \right) \\ &= \sum_k \log_2 \left( 1 + \frac{P\Lambda_k}{K\sigma^2} \frac{\bar{\beta}^k}{K} \mathbb{E}[\text{Tr}\{\mathbf{H}\mathbf{H}^H\}] \right). \end{aligned} \quad (25)$$

Based on Proposition 1,  $\frac{1}{N\Gamma^2} \mathbf{H}\mathbf{H}^H \sim \mathcal{W}_K(M, \mathbf{I}_M)$  is a central Wishart matrix with  $M$  degrees of freedom where  $M \geq K$ . According to the results in [9], we have

$$\mathbb{E}[\text{Tr}\{\mathbf{H}\mathbf{H}^H\}] = MN\Gamma^2. \quad (26)$$

Therefore, we can have (11).

### APPENDIX B

#### PROOF OF PROPOSITION 3

Based on the coordinates given in Fig. 2, we have

$$\beta_n^k = \frac{Aabz_0^3\lambda^2}{64\pi^3((x-x_0)^2 + y^2 + z_k^2)((x+x_0)^2 + y^2 + z_0^2)^{5/2}}. \quad (27)$$

Therefore, we have  $\bar{\beta}^k N = \sum_{n=1}^N \beta_n^k$ .

When  $N \rightarrow \infty$ , in the most area of the RIS, the distance from the reflective element and the center between the BS and user  $k$  is much longer than that between the BS and user  $k$ . Therefore, we can assume that the distances from the reflective element to the BS and the user are almost the same, and thus have the approximations below:

$$\begin{aligned} ((x-x_0)^2 + y^2 + z_k^2) &\approx ((x+x_0)^2 + y^2 + z_0^2) \\ &\approx (x^2 + y^2 + (z_0 + z_k)^2/4). \end{aligned} \quad (28)$$

Define the area of the RIS as  $\Phi$ . The received energy of the signals reflected by reflective elements in this area can be expressed by

$$\begin{aligned} \sum_{(x,y) \in \Phi} \beta_n^k &\approx \int_{(x,y) \in \Phi} \frac{Az_0^3\lambda^2}{64\pi^3(x^2 + y^2 + (z_0 + z_k)^2/4)^{7/2}} dx dy \\ &= \frac{Az_0^3\lambda^2}{32\pi^2} \int_0^\infty \frac{1}{(\rho^2 + (z_0 + z_k)^2/4)^{7/2}} \rho d\rho = \frac{2Az_0^3\lambda^2}{5\pi^2(z_0 + z_k)^5}, \end{aligned} \quad (29)$$

where  $\rho^2 = x^2 + y^2$ .

### APPENDIX C

#### PROOF OF PROPOSITION 5

According to the definition in (7), since  $\frac{P\Lambda_k}{K}$  is a constant, we can derive  $(\mathbf{g}_k \mathbf{w}_k)(\mathbf{g}_k \mathbf{w}_k)^H$  as follows:

$$(\mathbf{g}_k \mathbf{w}_k)(\mathbf{g}_k \mathbf{w}_k)^H = \mathbf{g}_k \frac{\mathbf{v}_k}{\|\mathbf{v}_k\|_F} \frac{\mathbf{v}_k^H}{\|\mathbf{v}_k\|_F} \mathbf{g}_k^H \stackrel{(a)}{=} \frac{1}{\|\mathbf{v}_k\|_F^2}. \quad (30)$$

Here, (a) can be achieved by

$$\begin{aligned} \mathbf{g}_k \mathbf{v}_k \mathbf{v}_k^H \mathbf{g}_k^H &= [\mathbf{G}\mathbf{G}^H(\mathbf{G}\mathbf{G}^H)^{-1}]_{k,k} [\mathbf{G}\mathbf{G}^H(\mathbf{G}\mathbf{G}^H)^{-1}]_{k,k}^H \\ &= 1. \end{aligned} \quad (31)$$

Note that  $\|\mathbf{v}_k\|_F^2 = \mathbf{v}_k^H \mathbf{v}_k$ , we have

$$\begin{aligned} \mathbf{v}_k^H \mathbf{v}_k &= [\mathbf{V}^H \mathbf{V}]_{k,k} = [(\mathbf{G}\mathbf{G}^H)^{-1} \mathbf{G}\mathbf{G}^H (\mathbf{G}\mathbf{G}^H)^{-1}]_{k,k} \\ &= [(\mathbf{G}\mathbf{G}^H)^{-1}]_{k,k}. \end{aligned} \quad (32)$$

### REFERENCES

- [1] M. D. Renzo *et al.*, "Smart radio environments empowered by reconfigurable AI meta-surfaces: An idea whose time has come," *EURASIP J. Wireless Commun. Netw.*, vol. 2019, no. 129, pp. 1–20, May 2019.
- [2] M. A. Elmoallamy, H. Zhang, L. Song, K. G. Seddik, Z. Han, and G. Y. Li, "Reconfigurable intelligent surfaces for wireless communications: Principles, challenges, and opportunities," *IEEE Trans. Cogn. Commun. Netw.*, vol. 6, no. 3, pp. 990–1002, Sep. 2020.
- [3] B. Di, H. Zhang, L. Song, Y. Li, Z. Han, and H. V. Poor, "Hybrid beamforming for reconfigurable intelligent surface based multi-user communications: Achievable rates with limited discrete phase shifts," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1809–1822, Aug. 2020.
- [4] Y. Cao, C. Yong, Z. Xiong, D. Niyato, Y. Xiao, and J. Zhao, "Reconfigurable intelligent surface for MISO systems with proportional rate constraints," in *Proc. IEEE ICC*, London, U.K., Jun. 2020, pp. 1–7.
- [5] H. Zhang, B. Di, L. Song, and Z. Han, "Reconfigurable intelligent surfaces assisted communications with limited phase shifts: How many phase shifts are enough?" *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 4498–4502, Apr. 2020.
- [6] W. Tang *et al.*, "Wireless communications with reconfigurable intelligent surface: Path loss modeling and experimental measurement," *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 421–439, Jan. 2021.
- [7] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
- [8] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.
- [9] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Foundations and Trends in Communications and Information Theory*, vol. 1. Hanover, MA, USA: Now Publishers, Jun. 2004, pp. 1–182.
- [10] M. A. Elmoallamy *et al.*, "On spatial multiplexing using reconfigurable intelligent surfaces," *IEEE Wireless Commun. Lett.*, vol. 10, no. 2, pp. 226–230, Feb. 2021.
- [11] "Study on channel model for frequencies from 0.5 to 100 GHz (release 14)," 3GPP, Sophia Antipolis, France, Rep. TR 38.901, Jan. 2018.