# Reconfigurable Refractive Surfaces: An Energy-Efficient Way to Holographic MIMO

Shuhao Zeng, Student Member, IEEE, Hongliang Zhang, Member, IEEE, Boya Di, Member, IEEE, Haichao Qin, Xin Su, and Lingyang Song, Fellow, IEEE

Abstract—Holographic Multiple Input Multiple Output (HMIMO), which integrates massive antenna elements into a compact space to achieve a spatially continuous aperture, plays an important role in future wireless networks. With numerous antenna elements, it is hard to implement the HMIMO via phased arrays due to unacceptable power consumption. To address this issue, reconfigurable refractive surfaces (RRS) is an energy efficient enabler of HMIMO since the surface is free of expensive phase shifters. Unlike traditional metasurfaces working as passive relays, the RRS is used as transmit antennas, where the far-field approximation does not hold anymore, urging a new performance analysis framework. In this letter, we first derive the data rate of an RRS-based single-user downlink system, and then compare its power consumption with the phased array. Simulation results verify our analysis and show that the RRS is an energy-efficient way to HMIMO.

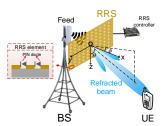
*Index Terms*—Reconfigurable refractive surfaces, holographic MIMO, energy efficient.

# I. INTRODUCTION

One of the key enablers of future wireless networks is Holographic Multiple Input Multiple Output (HMIMO), where a large number of tiny antennas or reconfigurable elements are integrated into a compact space to achieve a spatially continuous aperture [1]. Compared with conventional massive MIMO, the HMIMO has a larger dimension with smaller element spacing, and thus can achieve higher spectral efficiency and spatial resolution. However, it is hard for traditional phased arrays to realize the HMIMO, since hundreds of highresolution phase shifters are required, leading to unacceptable power consumption [2].

Recently, metasurface-based antennas, which are also referred to as reconfigurable refractive surfaces (RRS) [3], provide a promising solution for the implementation of HMIMO due to their high energy efficiency. An RRS is an ultra thin surface inlaid with a large number of sub-wavelength elements, each of which can refract incident electromagnetic (EM) waves and apply a tunable phase shift. By controlling the biased voltages applied to the diodes on the RRS elements, the refractive phase shifts can be reconfigured, so as to realize the desired beamforming. Different from the phased arrays, the RRSs do not contain any phase shifters, and the power consumption of each RRS element are ultra-low [4]. Thus, the RRSs are more energy efficient than the phased array.

Existing works on metasurface-based wireless communications mainly focus on utilizing the metasurface as a passive relay. For example, in [5], a metasurface is deployed between a base station (BS) and a user to extend cell coverage, and the location and orientation of the metasurface are jointly designed



1

Fig. 1. System model of a downlink single user network, where an RRS is used as the BS antenna.

to maximize the cell coverage. In [6], the authors investigate a multi-user MIMO system assisted by a metasurface, where the digital beamformer at the BS and the metasurface configuration are jointly optimized to maximize the sum rate. However, different from the existing works, the RRS acts as the transmit antenna in this letter.

The use of the RRS as the transmit antenna brings new challenges. Specifically, unlike the metasurface working as the passive relay, the feed of the transmitter cannot be assumed to locate in the far field of the RRS since the RRS is much closer to the feed. Therefore, the distances from different RRS elements to the feed cannot be assumed to be equal, and thus the analysis of the data rate and power consumption of the RRS-aided system is challenging. To cope with the above challenges, in this letter, we consider a fundamental downlink network with one BS and one user equipment (UE), where an RRS illuminated by a feed is deployed as the BS antenna. We first derive and analyze the data rate of the system. Then, the power consumption of the RRS is compared against the phased array under the same data rate requirement. Finally, simulation results validate our analysis and show that the RRS is an energy-efficient implementation of HMIMO.

#### **II. SYSTEM MODEL**

#### A. Scenario Description

As shown in Fig. 1, a narrow-band downlink network with one BS and one UE is considered. An RRS illuminated by a feed is utilized as the BS antenna to perform beamforming while an omin-directional antenna is adopted at the UE. For the sake of compactness, the feed is deployed in the near field of the surface. Besides, to point the transmitted wave of the feed towards the RRS, a directional antenna is used as the feed, whose radiation pattern can be given by [7]

$$G_F(\theta, \phi) = \begin{cases} 2(\alpha + 1)\cos^{\alpha}\theta, \ \theta \in [0, \frac{\pi}{2}],\\ 0, \ \text{otherwise}, \end{cases}$$
(1)

where  $2(\alpha+1)$  denotes the gain of the feed. For the simplicity of discussions, we assume that the direction of the main beam is vertical to the RRS.

To describe the topology of the network, we introduce a Cartesian coordinate as shown in Fig. 1, where the yoz plane

S. Zeng, B. Di, and L. Song are with Department of Electronics, Peking University, Beijing, China (email: {shuhao.zeng, boya.di, lingyang.song}@pku.edu.cn).

H. Zhang is with Department of Electrical and Computer Engineering, Princeton University, USA (email: hz16@princeton.edu).

H. Qin and X. Su are with CICT Mobile Communication Technology Co., Ltd. (email: {qinhaichao,suxin}@cictmobile.cn).

coincides with the RRS, and the x-axis is vertical to the surface. Define  $r_F$  as the distance from the feed to the center of the RRS. By assuming that the feed is on the x-axis, the coordinate of the feed is given by  $q_F = [-r_F, 0, 0]^T$ . For the UE, let  $r_U$  represent its distance to the center of the RRS, and  $\theta_U$  and  $\phi_U$  denote the zenith and azimuth angles, respectively.

### B. Reconfigurable Refractive Surfaces

A reconfigurable refractive surface (RRS) contains  $M_R \times N_R$  sub-wavelength elements, each with the size of  $s_{M,R} \times s_{N,R}$ . Different from the traditional reflective metasurface element such as that in [6], the RRS element does not contain the metallic ground in order to enable refraction. Each RRS element has several diodes onboard [8], which can be positive-intrinsic-negative (PIN) or varactor diodes. An *RRS controller* is utilized to apply different biased voltages to the diodes (i.e., *ON* and *OFF* states), which can influence the refraction phase shift of the element. Denote the refraction amplitude and phase shift of the (m, n)-th RRS element by A and  $\varphi^{(m,n)}$ , respectively. Therefore, the transmission coefficient of the (m, n)-th RRS element is written by  $\Gamma^{(m,n)} = A \exp(j\varphi^{(m,n)})$ .

Benefited from the reconfigurability, the RRS can generate controllable refractive beams. Specifically, when the signal transmitted by the feed impinges upon the RRS elements, it will be refracted by the elements with specific phase shifts. By controlling the applied phase shifts through configuring the elements' biased voltages, the radiated beam of the RRS can be directed accurately towards the UE with a high gain.

## C. Signal Model

1) Refracted signals by the RRS elements: Define  $G_F^{(m,n)}$  as the feed gain towards the direction of the (m, n)-th RRS element. Besides, let  $D^{(m,n)}$  represent the distance between the feed and the (m, n)-th RRS element, and denote the wavelength corresponding to the carrier frequency as  $\lambda$ . Since the feed is in the near field of the RRS, the received signal at the (m, n)-th element from the feed can be given by [9],

$$y_R^{(m,n)} = \left(\sqrt{\frac{G_F^{(m,n)}A_F^{(m,n)}}{4\pi (D^{(m,n)})^2}} \exp\left(-j\frac{2\pi}{\lambda}D^{(m,n)}\right)\right)x, \quad (2)$$

where x is the transmitted signal of the feed, and  $A_F^{(m,n)}$  is the projected aperture of the (m, n)-th RRS element towards the direction of the feed. Here, projected aperture  $A_F^{(m,n)}$  can be further expressed as  $A_F^{(m,n)} = (-u_x)^{\mathrm{T}}(q_F - q^{(m,n)})s_{M,R}s_{N,R}/D^{(m,n)}$ , where  $\hat{u}_x$  represents the unit vector in the x direction, and  $q^{(m,n)}$  is the location of the (m, n)-th RRS element. Afterwards, the RRS elements will refract the received signals, where the refracted signal of the (m, n)-th RRS element can be given by

$$\widetilde{y}_{R}^{(m,n)} = y_{R}^{(m,n)} \Gamma^{(m,n)}.$$
(3)

2) Channels from the RRS elements to the UE: To simplify analysis, we assume that the channels only contain pathloss. Since the size of the RRS can be extremely large, the UE can locate in either the near field or the far field of the RRS. When the UE is in the near field of the RRS, new characteristics are introduced into the channels compared with the far field case [9]. Specifically, the wavefront corresponding to the UE can be more accurately described by the spherical wavefront. Besides, the variation of the received signal strength from different RRS elements cannot be ignored, and the projected aperture of different RRS elements is also unequal. To show these characteristics, we model the channel by

$$h_{R}^{(m,n)} = \sqrt{\frac{G_{U}A_{U}^{(m,n)}}{4\pi(d^{(m,n)})^{2}}} \exp\left(-j\frac{2\pi}{\lambda}d^{(m,n)}\right), \quad (4)$$

where  $G_U$  is the antenna gain of the UE,  $d^{(m,n)}$  represents the distance between the UE and the (m, n)-th RRS element, and  $A_U^{(m,n)}$  denotes the projected aperture of the (m, n)-th RRS element in the direction of the UE. By defining  $q_U$  as the location of the UE,  $A_U^{(m,n)}$  can be expressed as  $A_U^{(m,n)} = (u_x)^T (q_U - q^{(m,n)}) s_{M,R} s_{N,R} / d^{(m,n)}$ .

When the refracted signals transmit through the channels from the RRS elements to the UE, the received signals at the UE can be written by

$$y = \sum_{m,n} h_R^{(m,n)} \tilde{y}_R^{(m,n)} + n,$$
 (5)

where *n* represents the additive white Gaussian noise (AWGN) at the UE with zero mean and  $\sigma^2$  as variance. Therefore, the received signal-to-noise ratio (SNR) at the UE is given by

$$\gamma_R = \left|\sum_{m,n} h_R^{(m,n)} \widetilde{y}_R^{(m,n)} \right|^2 / \sigma^2.$$
(6)

# III. DATA RATE ANALYSIS

# A. Derivation of Data Rate

Based on the SNR in (6), the system data rate can be expressed as  $C_R = \log_2(1 + \frac{|\sum_{m,n} h_R^{(m,n)} \tilde{y}_R^{(m,n)}|^2}{\sigma^2})$ . Define  $\Psi_U = \sin \theta_U \cos \phi_U$ ,  $\Phi_U = \sin \theta_U \sin \phi_U$ , and  $\Omega_U = \cos \theta_U$ . Besides, denote the transmit power by *P*. Since the data rate is correlated with the phase shifts of the RRS, it can be maximized by optimizing the phase shifts.

Theorem 1. The maximized data rate for the RRS is

$$C_{R} = \log_{2}(1 + L_{R} \left( \iint_{\mathcal{S}_{R}} f_{R}(y, z) dy dz \right)^{2}), \quad (7)$$
where  $L_{R} = \frac{PA^{2}G_{U}G_{0}\Psi_{U}r_{U}^{2}}{\sigma^{2}(4\pi)^{2}r_{F}^{2}}, \quad \mathcal{S}_{R} = \left[ -\frac{M_{R}s_{M,R}}{2r_{U}}, \frac{M_{R}s_{M,R}}{2r_{U}} \right] \times \left[ -\frac{N_{R}s_{N,R}}{2r_{U}}, \frac{N_{R}s_{N,R}}{2r_{U}} \right], \text{ and } f_{R}(y, z) = \left( 1 + \frac{r_{U}^{2}}{r_{F}^{2}} (y^{2} + z^{2}) \right)^{-\frac{\alpha+3}{4}} \left( 1 - 2\Phi_{U}y - 2\Omega_{U}z + y^{2} + z^{2} \right)^{-\frac{3}{4}}$ 

*Proof.* When the phases of different RRS-based channels are aligned, the data rate can be maximized as

$$C_{R} = \log_{2} \left( 1 + \left( \sum_{m,n} |h_{R}^{(m,n)} \tilde{y}_{R}^{(m,n)}| \right)^{2} / \sigma^{2} \right),$$
(8)

which can be further written as  $C_R = \log_2(1 + L_0(\sum_{m,n} f_R(\frac{m_{SM,R}}{r_U}, \frac{n_{SN,R}}{r_U}))^2)$ . Afterwards, following the method presented in Appendix A of [9], the maximum data rate can be transformed into the form in (7).

**Remark 1.** Theorem 1 indicates that the RRS can influence the data rate via the number  $M_R N_R$  of the RRS elements, the refraction amplitude A of the elements, the gain  $\alpha$  of the feed, and the distance  $r_F$  between the feed and the RRS.

Consider a circle  $\mathcal{O}_R$  with the origin as center and  $\min\left(\frac{M_Rs_{M,R}}{2r_U}, \frac{N_Rs_{N,R}}{2r_U}\right)$  as radius, which is found to be a part of region  $\mathcal{S}_R$ . Since  $f_R(y, z) \geq 0$ , a lower bound for  $C_R$  can be derived by replacing the integral region  $\mathcal{S}_R$  with  $\mathcal{O}_R$ , as shown in the following remark, which will be used for analyzing the power consumption of the RRS.

**Remark 2.** Data rate 
$$C_R$$
 can be lower bounded by  
 $C_R \ge \log_2 \left( 1 + L_R \left( \iint_{\mathcal{O}_R} f_R(y, z) dy dz \right)^2 \right) \triangleq C_R^{lb}.$  (9)

**Remark 3.** The lower bound  $C_R^{lb}$  in (9) is achievable when the number  $M_R N_R$  of the RRS elements is sufficiently large while the gap to  $C_R^{lb}$  is small when  $M_R N_R$  is relatively small.

For comparison, the data rate for the phased array is also derived. Assume that each antenna element within the phased array is omni-directional, whose antenna gain is denoted by  $G_E$ . Besides, use  $M_PN_P$  and  $(s_{M,P}, s_{N,P})$  to represent the number of the phased array elements and the separations among the elements, respectively.

**Theorem 2.** The maximized data rate for the phased array under optimized phase shifters can be given by,

$$C_P = \log_2(1 + L_P(\iint_{\mathcal{S}_P} f_P(y, z) dy dz)^2)$$
(10)

where  $L_P = \frac{P\lambda^2 G_E G_U r_U^2}{(4\pi)^2 \sigma^2 s_{M,P}^2 s_{N,P}^2}, \quad S_P = \left[-\frac{M_{P} s_{M,P}}{2r_U}, \frac{M_{P} s_{M,P}}{2r_U}\right] \times \left[-\frac{N_{P} s_{N,P}}{2r_U}, \frac{N_{P} s_{N,P}}{2r_U}\right], \text{ and } f_P(y,z) = \frac{1}{\sqrt{M_P N_P}} \frac{1}{\sqrt{1+y^2+z^2-2\Phi_U y-2\Omega_U z^2}}.$ 

## B. Discussions on Data Rate

**Lemma 1.** According to the derived data rate in (7), for an infinitely large RRS, the data rate does not increase unbounded, which is consistent with practical results.

# Proof. See Appendix B.

Compared with the phased array, the RRS has several new degrees of freedom for design, i.e., the distance  $r_F$  between the feed and the RRS, and the gain  $\alpha$  of the feed. Therefore, the impacts of  $r_F$  and  $\alpha$  on the data rate are discussed.

**Lemma 2.** When the number of the RRS elements is sufficiently large<sup>1</sup>, the data rate is positively correlated with the distance  $r_F$  between the feed and the RRS, and is negatively correlated with the gain  $\alpha$  of the feed.

*Proof.* Since the data rate  $C_R$  monotonically increases with the received SNR  $\gamma_R$  at the UE, we will show the influence of  $r_F$  and  $\alpha$  on  $\gamma_R$  as follows.

Based on (8), the received SNR can be rewritten as  $\gamma_R = (\sum_{m,n} |h_R^{(m,n)} \Gamma^{(m,n)}| |y_R^{(m,n)}|)^2 / \sigma^2$ . Then, by applying the Cauchy inequality, an upper bound for  $\gamma_R$  can be acquired:

$$\gamma_R \le \frac{1}{\sigma^2} (\sum_{m,n} |h_R^{(m,n)} \Gamma^{(m,n)}|^2) (\sum_{m,n} |y_R^{(m,n)}|^2).$$
(11)

Then, we show that when  $r_F$  increases or  $\alpha$  decreases, the gap between  $\gamma_R$  and the derived upper bound in (11) becomes smaller. Specifically, the equality in (11) holds only when

$$\frac{|y_R^{(m_1,n_1)}|}{|h_R^{(m_1,n_1)}\Gamma^{(m_1,n_1)}|} = \frac{|y_R^{(m_2,n_2)}|}{|h_R^{(m_2,n_2)}\Gamma^{(m_2,n_2)}|}.$$
 (12)

Denote the index of the center element of the RRS by (0,0). Since the feed directs more power towards the center of the RRS than the edge of the RRS, and the feed is closer to the RRS as compared to the distance between the UE and the RRS, we have  $\frac{|y_R^{(0,0)}|}{|h_R^{(0,0)}\Gamma^{(0,0)}|} > \frac{|y_R^{(m,n)}|}{|h_R^{(m,n)}\Gamma^{(m,n)}|}$  for  $(m,n) \succ (0,0)$  based on (2) and (4). Therefore, there is gap between  $\gamma_R$  and its upper bound in (11). However, according to (1), when  $\alpha$  becomes smaller, the radiation of the feed is more evenly distributed in different directions, and thus the gap can be narrowed. Besides, the gap can also be reduced by increasing the distance  $r_F$  between the feed and the RRS. Given the upper bound in (11) which is not related to  $r_F$ and  $\alpha$ , we find that when  $r_F$  increases or  $\alpha$  decreases, the SNR  $\gamma_R$  becomes larger, which verifies the remark.

### IV. POWER CONSUMPTION COMPARISON

In this part, we will compare the power consumption of the phased array and the RRS. For fairness, we assume that the required data rate for the RRS is equal to that for the phased array<sup>2</sup>, which are represented by C. Thus, the data rate  $C_R^{lb}$  for the RRS and  $C_P$  for the phased array satisfy  $C_R^{lb} = C_P = C$ . To guarantee the quality-of-service (QoS), we also set a minimum required data rate  $C^{min}$ , i.e.,  $C \ge C^{min}$ . Moreover, it is assumed that the power consumption of the BS antenna cannot exceed  $P^{max}$ .

#### A. Power Consumption Model

In general, the power consumption of the BS antenna contains two parts, i.e., transmit power P and the power consumed to support the operation of the antenna, denoted by  $P_R$  for the RRS and  $P_P$  for the phased array. Since the dynamic transmit power for the RRS and that for the phased array are set equal, we only discuss  $P_R$  and  $P_P$  in the following.

1) Power consumption model for the RRS: Based on the structure of the RRS-based antenna, we can find that power consumption  $P_R$  mainly comes from the diodes on the RRS elements and the controller [8]. Note that the controller is composed of an FPGA and several voltage converters [11], and both of them are active components. Therefore, we have

$$P_{R} = M_{R} N_{R} L_{R}^{(D)} P_{R}^{(D)} + \frac{M_{R} N_{R}}{Q} P_{R}^{(V)} + P_{R}^{(F)}, \quad (13)$$

where  $P_R^{(D)}$ ,  $P_R^{(V)}$ , and  $P_R^{(F)}$  are the power consumption of one diode, one voltage converter, and the FPGA, respectively,  $L_R^{(D)}$  is the number of diodes within an RRS element, and Q is the number of RRS elements within one group<sup>3</sup>.

2) Power consumption model for the phased array: For comparison, a model for the power consumption  $P_P$  of the phased array is also required. The phased array generally contains one power divider, where each output of the power divider connects to one phase shifter followed by one antenna element. Besides, an FPGA is utilized to configure the phase shifters. Since the power divider and the antenna elements are passive, the total power consumption is

$$P_{P} = M_{P} N_{P} P_{P}^{(S)} + P_{P}^{(F)}, \qquad (14)$$

where  $P_P^{(S)}$  and  $P_P^{(F)}$  are the power consumption of one phase shifter and the FPGA, respectively. We assume that the power consumption of the FPGA in the RRS is equal to that in the phased array, i.e.,  $P_R^{(F)} = P_P^{(F)}$ .

# B. Comparison of Power Consumptions between the RRS and the Phased array

Our aim is to derive the range of the required data rate Cunder which the RRS is more energy efficient. We would like to point out that different C can be achieved by changing the size of the RRS and the phase array. Define  $l_{pw}$  as the ratio of

 $<sup>^{1}\</sup>mathrm{A}$  sufficiently large RRS is one that captures most power radiated by the feed.

<sup>&</sup>lt;sup>2</sup>For fairness, the dynamic transmit power for the RRS and that for the phased array are set equal, as well.

<sup>&</sup>lt;sup>3</sup>For the ease of control, the metasurface is generally divided into several groups [10]. Each group contains a couple of elements sharing one voltage converter, and thus, the biased voltages for these elements are the same.

power consumptions between one phased array element and one RRS element, i.e.,

$$l_{pw} = P_P^{(S)} / \left( L_R^{(D)} P_R^{(D)} + P_R^{(V)} / Q \right).$$
(15)  
According to (13) and (14), we have the following remark.

**Remark 4.** When  $\frac{M_R N_R}{M_P N_P} \leq l_{pw}$ , the RRS consumes less power than the phased array. Otherwise, the phased array has a smaller power consumption.

It can be found that the required data rate C has an influence on  $\frac{M_R N_R}{M_P N_P}$  while  $l_{pw}$  is not correlated with C. Define g as a function indicating the relation between  $\frac{M_R N_R}{M_P N_P}$  and C, i.e.,  $g(C) = \frac{M_R N_R}{M_P N_P}$ . The closed-form expression of  $g(\cdot)$  is available only when the UE locates in the far field of BS antenna. This is because in the far field case, the distances from different antenna elements to the UE are approximately the same, and thus the expressions for the data rate in (9) and (10) can be rewritten in a closed form, from which the closed-form expression of  $g(\cdot)$  can be acquired. Therefore, the following discussions are separated into two parts, i.e., 1) the UE is in the far field of the RRS and the phased array, and 2) the UE locates in the near field of at least one antenna.

1) The UE locates in the far field: The far field assumption indicates that the distance from the UE to the RRS (the phased array) is larger than the boundary of the far field of the RRS (the phased array), i.e.,  $r_U > 2((M_R s_{M,R})^2 + (N_R s_{N,R})^2)/\lambda$ , and  $r_U > 2((M_P s_{M,P})^2 + (N_P s_{N,P})^2)/\lambda$ . Therefore, the number of the RRS elements and that of the phased array elements are limited by  $M_R N_R < k_R r_U \lambda / [2(k_R^2 s_{M,R}^2 + s_{N,R}^2)] \triangleq (M_R N_R)^{(thr)}$ , and  $M_P N_P < k_P r_U \lambda / [2(k_P^2 s_{M,P}^2 + s_{N,P}^2)] \triangleq (M_P N_P)^{(thr)}$ , respectively. Besides, according to (9) and (10), the data rate  $C_R^{(lb)}$  for the RRS and the data rate  $C_P$  for the phased array are positively correlated with  $M_R N_R$  and  $M_P N_P$ , respectively. Therefore,  $C_R^{(lb)}$  and  $C_P$  are smaller than their values under  $(M_R N_R)^{(thr)}$  and  $(M_P N_P)^{(thr)}$ , respectively. Due to the equal data rate requirement between the RRS and the phased array, denoted by C, we have  $C_R^{lb} = C_P \triangleq C$ . Therefore, C is upper bounded by  $C^{(thr)}_R (M_P N_P)^{(thr)}$ ). (16)

 $C^{thr} \triangleq \min\{C_R^{lb}((M_R N_R)^{(thr)}), C_P((M_P N_P)^{(thr)})\}.$  (16) Recall that a minimum required data rate  $C^{min}$  is set in order to guarantee the QoS. As a result, we have  $C \in [C^{min}, C^{thr})$ . It is worthwhile noting that when the UE is in the near field of the BS antenna, the required data rate C is larger than  $C^{thr}$ , and thus  $C^{thr}$  can be regarded as a threshold between the required data rate corresponding to the far field and that corresponding to the near field.

Without loss of generality, we assume that the dimension of the RRS in the y direction is larger than that in the z direction, i.e.,  $M_R s_{M,R} > N_R s_{N,R}$ . Under the far field assumption, the data rate for the RRS in (9) and the data rate for the phased array in (10) can be simplified into

$$C_{R}^{lb} = \log_{2} \left( 1 + \left( \frac{4\pi L_{R}^{2} r_{F}^{2}}{(\alpha - 1) r_{U}^{2}} \left( 1 - \left( 1 + \frac{M_{R} N_{R} k_{R} s_{N,R}^{2}}{4 r_{F}^{2}} \right)^{\frac{1 - \alpha}{4}} \right) \right)^{2} \right).$$

$$C_P = \log_2 \left( 1 + (M_P N_P s_{M,P}^2 s_{N,P}^2 L_P) / r_U^2 \right), \tag{18}$$

respectively. Recall that  $C_R^{lb} = C_P \triangleq C$ . Thus, from (17) and (18), we can derive the expression for the function  $g(\cdot)$  indicating the relation between  $\frac{M_R N_R}{M_P N_P}$  and C as shown below,

$$g(C) = \frac{4s_{M,P}^2 s_{N,P}^2 L_P r_F^2}{k_R s_{N,R}^2 r_U^4 (2^C - 1)} \Big[ (1 - \frac{(\alpha - 1)r_U^2}{4\pi L_R^{0.5} r_F^2} \sqrt{2^C - 1})^{-\frac{4}{\alpha - 1}} - 1 \Big]$$

**Remark 5.** g(C) first decreases and then increases with C. Besides, under sufficiently small or sufficiently large data rate, g(C) is larger than  $l_{pw}$  defined in (15).

We can infer from Remark 5 that there are only two data rate satisfying  $g(C) = l_{pw}$ , denoted by  $C^{(e,1)}$  and  $C^{(e,2)}$ , respectively<sup>4</sup>. Without loss of generality, we assume that  $C^{(e,1)} < C^{(e,2)}$ . Recall that the RRS consumes less power only when  $\frac{M_R N_R}{M_P N_P} = g(C) < l_{pw}$  as indicated by Remark 4, and that we only consider the required data rate within  $[C^{min}, C^{thr}]$ . Therefore, we have:

**Theorem 3.** When  $\max\{C^{min}, C^{(e,1)}\} \leq C \leq \min\{C^{thr}, C^{(e,2)}\}$ , the RRS consumes less power. Otherwise, the phased array is more power-saving.

**Remark 6.** Since the maximum of  $g(C) = \frac{M_R N_R}{M_P N_P}$  within  $[C^{min}, C^{thr})$  is achieved at  $C^{min}$  or  $C^{thr}$ , the RRS always consumes less power than the phased array when  $l_{pw} > \max\{g(C^{min}), g(C^{thr})\}$  according to Remark 4.

2) The UE is in the near field: Without the far field assumption, the closed-form expression of g(C) is not available, and thus only qualitative analysis is presented in this part.

**Theorem 4.** Given sufficiently large required data rate C, by moving the feed away from the RRS or using a feed with a smaller gain  $\alpha$ , more power can be saved by the RRS compared with the phased array.

*Proof.* Since the required data rate C is given, by increasing the distance  $r_F$  between the RRS and the feed or decreasing  $\alpha$ , the size of the RRS under which the required data rate can be achieved becomes smaller according to Lemma 2. However,  $r_F$  and  $\alpha$  have no influence on the size of the phased array, which ends the proof.

### V. SIMULATION RESULTS

In this section, we verify the analytical results and evaluate how much power can be saved by the RRS through simulations. The parameters are based on existing works [5]. Specifically, the center frequency is set as 26 GHz, with a wavelength of  $\lambda = 1.15$  cm. The location of the UE is given by  $(\theta_U, \phi_U, r_U) = (\frac{\pi}{6}, \frac{\pi}{4}, 50 \text{ m})$ . The transmit power and noise variance are P = 43 dBm and  $\sigma^2 = -96$  dBm, respectively. We set  $k_R = k_P = 1$ , i.e.,  $\frac{N_R}{M_R} = \frac{N_P}{M_P} = 1$ . The gain of the phased array antenna element is  $G_E = 1$  and that of the UE antenna is  $G_U = 1$ . The size of the RRS elements and the element separation for the phased array are selected as  $s_{M,R} = s_{N,R} = \frac{\lambda}{6}$ , and  $s_{M,P} = s_{N,P} = \frac{\lambda}{2}$ , respectively. The refraction amplitude of the RRS elements is given by A = 0.8. Each RRS element contains  $L_R^{(D)} = 1$  varactor diode of the power consumption  $P_R^{(D)} = 5 \times 10^{-6}$  W. The power consumption of each phase shifter in the phased array, of one FPGA, and of one voltage converter are set as  $P_P^{(S)} = 0.1$  W,  $P_R^{(F)} = P_P^{(F)} = 5$  W, and  $P_R^{(V)} = 5 \times 10^{-4}$  W, respectively. The maximum power constraint is  $P^{max} = 250$  W, and the minimum required data rate is  $C^{min} = 20$  bit/s/Hz.

Fig. 2(a) compares the data rate  $C_R$  of the RRS against its lower bound  $C_R^{lb}$ . We can find that  $C_R^{lb}$  is achievable when the number of the RRS elements is sufficiently large, while the gap between  $C_R$  and  $C_R^{lb}$  is small when the number of the RRS elements is relatively small, which verifies Remark 3.

<sup>&</sup>lt;sup>4</sup>In general, there exists data rate under which the RRS consumes less power than the phased array, i.e.,  $\exists C, s.t.g(C) < l_{pw}$ .

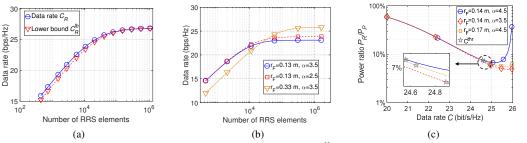


Fig. 2. (a) Comparison between the exact data rate  $C_R$  for the RRS and its lower bound  $C_R^{lb}$  defined in (9), with with  $r_F = 0.15$  m and  $\alpha = 5$ . (b) Data rate versus the number of the RRS elements. (c) Power consumption ratio  $\frac{P_R}{P_P}$  versus the required data rate C, where  $C \leq C^{thr}$  and  $C > C^{thr}$  correspond to the far field case and the near field case, respectively.

Fig. 2(b) depicts the data rate versus the number  $M_R N_R$ of the RRS elements. According to Fig. 2(b), the data rate will increase with  $M_R N_R$  when  $M_R N_R$  is relatively small. However, the data rate will converge when  $M_R N_R$  continues to grow, which verifies Lemma 1. From Fig. 2(b), we can also find that when  $M_R N_R$  is sufficiently large,  $r_F$  and  $\alpha$ have a positive impact and a negative impact on the data rate, respectively, which consists with Lemma 2.

In Fig. 2(c), we plot power consumption ratio  $\frac{P_R}{P_P}$  versus the required data rate C. We can find that when  $C < C^{thr}$ , the RRS can always save energy compared with the phased array, which is consistent with Remark 6 given the fact that  $l_{pw}$ is larger than max $\{g(C^{min}), g(C^{thr})\}$ . In addition, Fig. 2(c) shows that given sufficiently large data rate, by reducing the gain  $\alpha$  of the feed or increasing the distance  $r_F$  between the feed and the RRS, more power can be saved by the RRS compared with the phased array, which is consistent with Theorem 4. Fig. 2(c) also indicates that an RRS has the potential to significantly reduce power consumption compared with the phased array for any required data rate, and thus the RRS can serve as a practical enabler of HMIMO.

# VI. CONCLUSION

In this letter, we have considered a downlink network with one BS and one UE, where an RRS has been used as the BS antenna. The data rate of the RRS-aided system has been derived and analyzed. Then, we have compared the power consumption of the RRS against the phased array under the same data rate requirement. From the analytical and simulations results, we can conclude that: 1) When the RRS contains a sufficiently large number of elements, the data rate is positively correlated with the distance between the feed and the RRS, while it is negatively related with the gain of the feed. 2) We derive the range of the required data rate, by achieving which the RRS consumes less power than the phased array when the UE is in the far field of the BS antenna. 3) By designing the gain of the feed and its distance from the RRS, the RRS can significantly reduce the power consumption compared with the phased array for any required data rate, and thus the RRS is an energy-efficient way to holographic MIMO. APPENDIX A

# **PROOF OF THEOREM 2**

For simplicity, we assume that the channels from the phased array elements to the UE only contains pathloss, which is described by the free space pathloss model. Therefore, the channel from the (m,n)-th element to the UE can be given by  $h_P^{(m,n)} = \frac{\lambda\sqrt{G_EG_U}}{4\pi d^{(m,n)}} \exp(-j\frac{2\pi d^{(m,n)}}{\lambda})$  [12]. Besides, the received signal at the UE can be given by  $y = \sum_{m,n} h_P^{(m,n)} \exp(j\varphi_P^{(m,n)}) \frac{1}{\sqrt{M_PN_P}}x + n$ , where x is

the transmitted signal,  $\varphi_P^{(m,n)}$  denotes the phase shift induced by the (m,n)-th phase shifter, and n represents the received AWGN. Therefore, the data rate can be expressed as  $C_P = \log_2(1 + \frac{P\lambda^2 G_E G_U}{M_P N_P (4\pi)^2 \sigma^2} |\sum_{m,n} (d^{(m,n)})^{-1} \exp(j(-\frac{2\pi d^{(m,n)}}{\lambda} + \varphi_P^{(m,n)}))|^2)$ . By aligning the phases of the received signals from different array elements,  $C_P$  is maximized. Then, we can prove (10) following the methods to prove Theorem 1.

#### APPENDIX B Proof of Lemma 1

Note that  $f_R(y, z)$  decreases with the antenna gain  $\alpha$ , and  $f_R(y, z) > 0$ . Thus, the data rate  $C_R$  for the RRS is maximized when  $\alpha = 0$ . Besides, based on the triangular inequality, we have  $\left(1 + \frac{r_U^2}{r_F^2}(y^2 + z^2)\right)^{\frac{1}{2}} \ge |\sqrt{(\frac{r_U}{r_F})^2(y^2 + z^2)} - 1|$ , and  $(1 - 2\Phi_U y - 2\Omega_U z + y^2 + z^2)^{\frac{1}{2}} \ge |\sqrt{y^2 + z^2} - 1|$ . Thus, we have  $C_R \le \log_2(1 + L_R(\iint_{S_R} f_R^{ub}(y, z) dy dz)^2) \triangleq C_R^{ub}$ , (19) where  $f_R^{ub}(y, z) = |\sqrt{(\frac{r_U}{r_F})^2(y^2 + z^2)} - 1|^{-3/2}|\sqrt{y^2 + z^2} - 1|^{-3/2}$ . We can easily find out that  $C_R^{ub}$  is limited when the

 $1|^{-3/2}$ . We can easily find out that  $C_R^{ub}$  is limited when the number of the RRS elements is infinite. Therefore,  $C_R$  does not increase infinitely with the number of the RRS elements.

# REFERENCES

- A. Pizzo, T. L. Marzetta, and L. Sanguinetti, "Spatially-stationary model for holographic MIMO small-scale fading," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 9, pp. 1964-1979, Sept. 2020.
- [2] L. Dai, et. al., "Reconfigurable intelligent surface-based wireless communications: Antenna design, prototyping, and experimental results," *IEEE Access*, vol. 8, no. 1, pp. 45913-45923, Mar. 2020.
- [3] S. Zeng, et. al., "Reconfigurable intelligent surfaces in 6G: Reflective, transmissive, or both," *IEEE Commun. Lett.*, vol. 25, no. 6, pp. 2063-2067, Feb. 2021.
- [4] M. Di Renzo, et. al., "Reconfigurable intelligent surfaces vs. relaying: Differences, similarities, and performance comparison," *IEEE Open J. Commun. Soc.*, vol. 1, no. 1, pp. 798-807, June 2020.
- [5] S. Zeng, et. al., "Reconfigurable intelligent surface (RIS) assisted wireless coverage extension: RIS orientation and location optimization," *IEEE Commun. Lett.*, vol. 25, no. 1, pp. 269-273, Jan. 2021.
- [6] B. Di, et. al., "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.
- [7] W. L. Stutzman and G. A. Thiele, Antenna Theory Design, 3rd ed. New York, NY, USA: Wiley, 2012.
- [8] W. Gao, et. al., "1-bit reconfigurable transmitarray with low loss and wide bandwidth," *New J. Phys.*, vol. 1, no. 1, pp. 1-9, June 2021.
- [9] H. Lu, and Y. Zeng, "Communicating with extremely large-scale array/surface: Unified modelling and performance analysis," *IEEE Trans. Wireless Commun.*, to be published.
- [10] H. Zhang, et. al., "Intelligent omni-surfaces for full-dimensional wireless communications: Principles, technology, and implementation," *IEEE Commun. Mag.*, vol. 60, no. 2, pp. 39-45, Feb. 2022.
- [11] J. Dai, et. al., "Wireless communications through a simplified architecture based on time-domain digital coding metasurface," *Adv. Mater. Technol.*, vol. 1, no. 1, pp. 1-8, Feb. 2019.
- [12] A. Goldsmith, Wireless Communications, New York, NY, USA: Cambridge University Press, 2005.