Intelligent and Converged Networks ISSN 2708-6240
2022, 3(1): 33−63 DOI: 10.23919/ICN.2022.0007

Reconfigurable intelligent surface-aided wireless communications: An overview

(*invited paper*)

Muhammad Zain Siddiqi* and Talha Mir

Abstract: The reconfigurable intelligent surface (RIS) is an emerging technology, which will hopefully bring a new revolution in wireless communications. The RIS technology can be deployed in an indoor/outdoor environment to dynamically manipulate the propagation environment. The RIS consists of a large number of independently controllable passive elements, and these elements are involved in realizing high passive beamforming gain. Different from the conventional active phased antenna array, there is no dedicated radio-frequency (RF) chain installed at the RIS to perform complex signal processing operations. Therefore, it does not incur additional noise while retransmitting the incident wave, which is substantially a unique feature from the conventional wireless communication systems. Taking advantage of its working principle, RIS has been deployed in various practical scenarios. In this tutorial, at first we will review the latest advances in RIS, including the application scenarios such as the system and channel model, the information theoretic analysis, the physical realization and design, key signal processing techniques such as precoding and channel estimation, and prototyping. Finally, we discuss interesting future research problems for the RIS-aided communications.

Key words: reconfigurable intelligent surface (RIS); wireless communications; new paradigm

1 Introduction

The reconfigurable intelligent surface (RIS), also named as intelligent reflecting surface (IRS), large intelligent surface (LIS), programmable metasurface, etc., is an emerging technology which is envisioned to be integrated with high spectral band to achive high specral efficiency in the future wireless communications.

RIS is a physical-layer technology that can be deployed in an indoor and outdoor environment to manipulate the propagation of electromagnetic (EM) waves by varying the key electric and magnetic characteristics of an RIS surface. Besides having the potential to control the EM waves, RIS can be deployed to control the properties of the radio channel by adding an adequate sensing capabilities to it[1]. A reconfigurable metasurface is a rectangular array composed of a large number of controllable elements, where each controllable element can change the phase, amplitude, or polarization of the incident EM wave. The metasurface manipulates the incident EM wave received from the transmitter and then retransmits the EM wave towards the receiver. It is worth mentioning here that each RIS element controls the environment independently. Thus, by designing the coefficients of these RIS elements, we can realize passive beamforming with high array gain. Different from conventional active phased arrays, there are no dedicated transceivers installed at RIS to perform complex signal processing operations. As a result, no additional noise will be introduced by the RIS while redirecting the EM wave. These unique characteristics enable this technology to stand out from the

• Muhammad Zain Siddiqi is with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China. E-mail: siddiqimz10@mails.tsinghua.edu.cn.
• Talha Mir is with the Department of Electronic Engineering, Faculty of ICT, BUITEMS, Quetta 87300, Pakistan.
∗ To whom correspondence should be addressed.
Manuscript received: 2021-12-18; revised: 2022-02-24; accepted: 2022-03-22

conventional active phased arrays. Usually, RIS acts as a mediator between transmitter and receiver to assist communication. By deploying an RIS between transmitter and receiver, the wireless propagation channel can be controlled dynamically, which is substantially a unique feature compared to conventional wireless communication that focuses only on the design of transceivers[2].

The idea of controlling wireless propagation has recently gained attention with the advancement of metamaterials[3, 4]. Different from all-natural materials, metamaterials are artificially designed materials that exhibit unusual EM properties such as negative permittivity and permeability[5, 6]. The idea of intelligently controlling the wireless propagation environment in a wireless communication system originated back in 2012, where the intelligent walls were proposed[7] to change the indoor coverage by controlling the transparency of the intelligent walls according to different demands. Then, In 2014, Kaina et al. designed a binary phase-shift based electrically tunable metasurface that acted like a spatial microwave modulator[8]. In this work, the authors[8] proved that the received signal power at the user side can be effectively increased or suppressed using different phase shifts. Moreover, Cui et al. in Ref. [9] realized passive beamforming using digital coding metasurface. The digitally coded metasurface has the potential to manipulate the EM wave via different coding sequences of "0 s" and "1 s", which are controlled by a biased diode. Interestingly, all of the above-mentioned works have put more effort in the designing of metasurfaces, which opened a new era in the field of wireless communications in terms of controlling the propagation environment.

The metasurface/RIS brings many new opportunities and challenges to researchers in the fields of information theory, electromagnetics, wireless communications, and signal processing[10]. Since 2018, there are a steeply increasing number of papers about RIS, trying to answer the following important yet challenging questions. What is the system capacity? What are the performance limits? How can we model the channel of the RIS? How do we implement the RIS elements and arrays? How to optimize the tuning coefficients for wireless signal transmission? How to obtain the channel state information (CSI)? How can we combine the RIS with other existing technologies in wireless communications? Can we get some experimental results by building prototypes? What are the potential application scenarios of RIS in future communication systems?

In this paper, we provide a comprehensive survey on the emerging RIS technology and discuss its key aspects that have been discussed in the existing works to realize the objective of future 6G communications. For this purpose, we organize this paper based on the system and path loss modeling of RIS-aided communication, discuss potential research problems for RIS-aided systems, and also discuss the hardware implementations while highlighting the key prototypes of RIS-sided systems, as summarized in Fig. 1. At first, we comprehensively overview regarding the key RIS technology in Section 1 and then highlight the system model of RIS-aided communication in Section 2. The path loss models considering near- and far-fields as the key apsects are briefly summarized in Section 3. In Section 4, we categorize the literature that solely worked on the asymptotic performance limit of RIS-aided communication. Sections 5 and 6 mainly discuss the procoding and channel estimation problems for RIS-aided communication. We then investigate the impact of system performances by aiding RIS with other technologies, such as simultaneous wireless information and power transfer (SWIPT), non-orthogonal multiple access (NOMA), physical layer security, unmanned aerial vehicle (UAVs), and

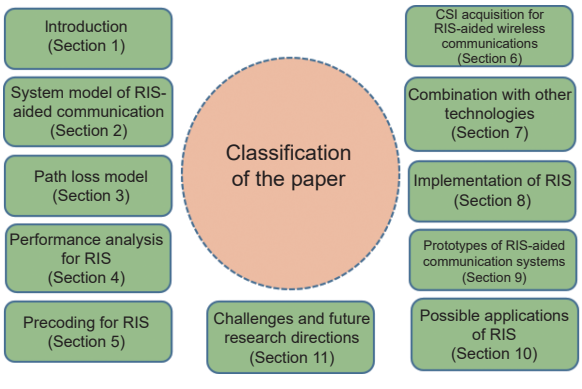

Fig. 1 Structure of the survey paper.

backscatter communications in Section 7. To realize practically, we then discuss RIS hardware implementation and its available prototypes in Sections 8 and 9, respectively. Section 10 illustrates the possible applications where RIS can be implemented to further improve the performance of wireless systems. We present the challenges and future research directions for RIS-aided communication in Section 11. Finally, we conclude our survey in Section 12.

2 System model of RIS-aided communication system

2.1 Signal model

We consider an RIS-aided wireless communication system as shown in Fig. 2. A massive multiple-input multiple-output (MIMO) base station (BS) with M antennas serves a single-antenna user with the aid of an N-element RIS. Considering the downlink transmission, the system model is given by

$$y=\left(\boldsymbol{h}^{\mathrm{T}}+\boldsymbol{f}^{\mathrm{T}}\boldsymbol{\Theta}\boldsymbol{G}\right)\boldsymbol{w}x+n \tag{1}$$

where $y\in\mathbb{C}$ is the received signal at the user, $\boldsymbol{h}\in\mathbb{C}^{M\times 1}$ is the direct channel between the BS and the user regardless of the RIS, $\boldsymbol{f}\in\mathbb{C}^{N\times 1}$ is the channel between the RIS and the user, $\boldsymbol{\Theta}\in\mathbb{C}^{N\times N}$ is the tuning coefficient matrix at the RIS, $\boldsymbol{G}\in\mathbb{C}^{N\times M}$ is the channel between the BS and the RIS, $\boldsymbol{w}\in\mathbb{C}^{M\times 1}$ is the precoding vector at the BS, $x\in\mathbb{C}$ is the transmitted signal at the BS, and $n\in\mathbb{C}$ is the additive noise, respectively. The tuning coefficient matrix $\boldsymbol{\Theta}=\mathrm{diag}(\boldsymbol{\theta})$, where $\boldsymbol{\theta}\in\mathbb{C}^{N\times 1}$ is the tuning coefficient vector of RIS, which subjects to certain constraints of the RIS hardware. For a typical RIS that only controls the phases but does not change the amplitudes of incident signals, the coefficients should have unit modulus, i.e., we can write $\boldsymbol{\theta}=\left[\mathrm{e}^{\mathrm{j}\theta_1},\mathrm{e}^{\mathrm{j}\theta_2},\ldots,\mathrm{e}^{\mathrm{j}\theta_N}\right]$, where $\theta_1,\theta_2,\ldots,\theta_N\in[0,2\pi)$ are the phase shifts.

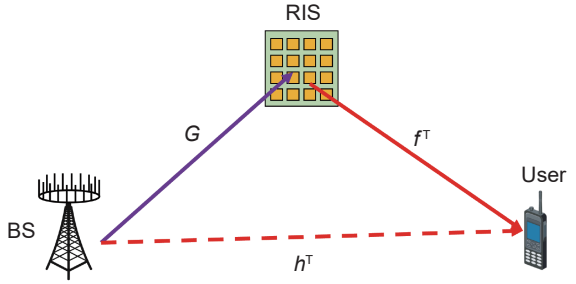

Fig. 2 System model of downlink transmission in RIS-aided single-user communication system.

When the user is equipped with K antennas, we can easily modify Eq. (1) and rewrite the equation as

$$\boldsymbol{y}=\left(\boldsymbol{H}^{\mathrm{T}}+\boldsymbol{F}^{\mathrm{T}}\boldsymbol{\Theta}\boldsymbol{G}\right)\boldsymbol{w}x+\boldsymbol{n} \tag{2}$$

where $\boldsymbol{y}\in\mathbb{C}^{K\times 1}$, $\boldsymbol{H}\in\mathbb{C}^{M\times K}$, $\boldsymbol{F}\in\mathbb{C}^{N\times K}$, and $\boldsymbol{n}\in\mathbb{C}^{K\times 1}$.

2.2 Channel model

(1) Rayleigh channel model: The simplest channel model is the Rayleigh-fading channel. In the Rayleigh-fading channel model, all coefficients of $\boldsymbol{h}$, $\boldsymbol{f}$, and $\boldsymbol{G}$ subject to independent and identical (i.i.d.) circularly symmetric complex Gaussian distribution, i.e., $\boldsymbol{h}\sim\mathcal{CN}\left(0,\sigma_h^2\boldsymbol{I}_{M\times M}\right)$, $\boldsymbol{f}\sim\mathcal{CN}\left(0,\sigma_f^2\boldsymbol{I}_{N\times N}\right)$, and $\mathrm{vec}(\boldsymbol{G})\sim\mathcal{CN}\left(0,\sigma_G^2\boldsymbol{I}_{MN\times MN}\right)$[11, 12]. The variances σ_h^2, σ_f^2, and σ_G^2 are related to the distance between BS and user end (UE), the distance between RIS and UE, and the distance between BS and RIS, respectively. Since these distances are usually different from one another, it is impossible to normalize these three channels at the same time. As a result, we need to set proper values for σ_h^2, σ_f^2, and σ_G^2 before generating a Rayleigh-fading channel for simulations.

(2) Multipath channel model: In addition to the Rayleigh-fading channels, the Saleh-Valenzuela multipath channel model is widely used[13–16]. Specifically, the channel $\boldsymbol{h}$ between the BS and the user, the channel $\boldsymbol{f}$ between the RIS and the user, and the channel $\boldsymbol{G}$ between the BS and the RIS can be respectively expressed by

$$\boldsymbol{h}=\sqrt{\rho_h}\sum_{l=1}^{L_h}\alpha_l^h\boldsymbol{b}\left(\vartheta_l^{rx},\psi_l^{rx}\right) \tag{3}$$

$$\boldsymbol{f}=\sqrt{\rho_f}\sum_{l=1}^{L_f}\alpha_l^f\boldsymbol{a}\left(\vartheta_l^{t},\psi_l^{t}\right) \tag{4}$$

$$\boldsymbol{G}=\sqrt{\rho_G}\sum_{l=1}^{L_G}\alpha_l^G\boldsymbol{a}\left(\vartheta_l^{r},\psi_l^{r}\right)\boldsymbol{b}\left(\vartheta_l^{tx},\psi_l^{tx}\right)^{\mathrm{T}} \tag{5}$$

where ρ_h, ρ_f, and ρ_G are the distance-dependent parameters; L_h, L_f, and L_G are the numbers of paths in the corresponding channels; $\{\alpha_l^h\}_{l=1}^{L_h}$, $\{\alpha_l^f\}_{l=1}^{L_f}$, and $\{\alpha_l^G\}_{l=1}^{L_G}$ represent the normalized complex gains for the

corresponding channels; $\{\vartheta_l^{rx}\}_{l=1}^{L_h}$ and $\{\psi_l^{rx}\}_{l=1}^{L_h}$ represent the elevation angles and the azimuth angles at the BS for $\boldsymbol{h}$, respectively; $\{\vartheta_l^{tx}\}_{l=1}^{L_G}$ and $\{\psi_l^{tx}\}_{l=1}^{L_G}$ represent the elevation angles and the azimuth angles at the BS for $\boldsymbol{G}$, respectively; $\{\vartheta_l^{t}\}_{l=1}^{L_f}$ and $\{\psi_l^{t}\}_{l=1}^{L_f}$ represent the elevation angles and the azimuth angles at the RIS for $\boldsymbol{f}$, respectively; $\{\vartheta_l^{r}\}_{l=1}^{L_G}$ and $\{\psi_l^{r}\}_{l=1}^{L_G}$ represent the elevation angles and the azimuth angles at the RIS for $\boldsymbol{G}$, respectively. $\boldsymbol{a}(\vartheta,\psi)\in\mathbb{C}^{N\times 1}$ and $\boldsymbol{b}(\vartheta,\psi)\in\mathbb{C}^{M\times 1}$ represent the normalized array steering vector associated to the RIS and the BS, respectively. For a typical $N_1\times N_2$ $(N=N_1\times N_2)$ uniform planar array (UPA), the $N\times 1$ array steering vector $\boldsymbol{a}(\vartheta,\psi)$ can be written by[17]

$$\boldsymbol{a}(\vartheta,\psi)=\frac{1}{\sqrt{N}}\left[\mathrm{e}^{-\mathrm{j}2\pi d\sin(\vartheta)\sin(\psi)\boldsymbol{n}_1/\lambda}\right]\otimes\left[\mathrm{e}^{-\mathrm{j}2\pi d\cos(\psi)\boldsymbol{n}_2/\lambda}\right] \tag{6}$$

where $\boldsymbol{n}_1=[0,1,\ldots,N_1-1]^{\mathrm{T}}$, $\boldsymbol{n}_2=[0,1,\ldots,N_2-1]^{\mathrm{T}}$; λ is the signal wavelength; d is the distance between adjacent elements usually satisfying $d=\frac{\lambda}{2}$[18]. Similarly, $\boldsymbol{b}(\vartheta,\psi)$ can be also represented like Eq. (6) by replacing N, $\boldsymbol{n}_1$, and $\boldsymbol{n}_2$ by M, $\boldsymbol{m}_1=[0,1,\ldots,M_1-1]^{\mathrm{T}}$, and $\boldsymbol{m}_2=[0,1,\ldots,M_2-1]^{\mathrm{T}}$ $(M=M_1\times M_2)$.

It must be pointed out that the channel models in this subsection only apply to far-field radiation of EM waves. When two components are close to each other, the channel between them has more complicated models. It is still an on-going research, which we will present in the next section.

3 Path loss model

In the recent years, the researchers are testing ultra-high frequencies such as mmWave and terahertz to achieve the performance gain of future 6G communications. However, these ultra-high frequencies suffer large penetration/path loss problems during transmission. Thus, by integrating the RIS technology with these frequencies we can not only achieve high spectral efficiency but also can minimize the large path loss problem. Since RIS is an emerging technology and is envisioned to be deployed in the future 6G communication, therefore, new path loss models are being investigated to analyze the performance limit of RIS-aided communications. The path model helps in estimating the link budget in advance to install RIS in any communication scenario. With an accurate path loss calculation, one can achieve high beamforming gain in different application scenarios, for instance, performing an accurate channel estimation or precoding. Thus, summing up, the path loss model provides the guidelines for the practical performance analysis, the system design, and key signal processing techniques in the RIS-aided wireless communication system. Since the researchers are adopting a conventional path loss model in most of the RIS-aided wireless scenarios, however, a more practical path loss model is required that could easily be tackled and adjusted according to the practical scenario.

Specifically, assuming that the signal is transmitted with the transmitting power P_{t}, and is then reflected by the RIS, finally is received with the receiving power P_{r}, the corresponding path loss refers to the ratio of the transmitting power and the receiving power $P_{\mathrm{t}}/P_{\mathrm{r}}$. As a matter of fact, the path loss model can vary with the EM field around the RIS. If the transmitter is far away from the RIS, the spherical wave transmitted by the antenna can be regarded as a plane wave when it arrives at the RIS. In this case, the transmitter is called to be in the far-field of the RIS. Otherwise, the transmitter is in the near-field of the RIS. In this case the EM wave arriving at the RIS is more complicated, since it is no longer a plane wave. Generally speaking, by denoting the large dimension D of the RIS and the wave length λ of the signal, the boundary between the far-field and the near-field of the RIS can be denoted by $B=\frac{2D^2}{\lambda}$. We can regard the system as the far-field one if the distance between the transmitter and the RIS is longer than B, while it is near-field when the distance is shorter than B. Similarly, the relative position between the RIS and the receiver can also be modeled as far-field or near field according to the distance.

In this section, we will introduce the literature about the path loss model in the RIS-aided system from two aspects: the far-field and the near-field. The basic diagram of near- and far-fields communication system is illustrated in Fig. 3.

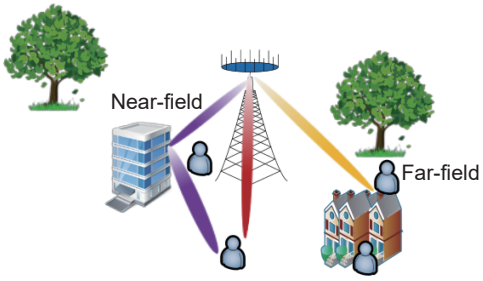

Fig. 3 Near- and far-fields communication system.

3.1 Far-field

Recently, few works have been reported which investigate the path loss model in the far-field of the RIS-aided communication[19–22]. For instance, Refs. [19, 20] explained that all elements of the finite-sized RIS can individually act as diffuse scatters when the transmitter and the receiver both lie in the far-field of the RIS. Similar to a phase shifter, each sub-wavelength-sized RIS element scatters the incoming signals with a unique phase shift. By denoting the incident angle ϑ_i from the transmitter to the RIS and assuming the scattered waves from the RIS can reach the maximum value in the observation angle ϑ_d leading to the receiver, Ref. [19] derived the far-field path loss model in terms of physical optics techniques:

$$PL_{\text{far-field}} = \frac{(4\pi)^2}{G_{\text{t}}G_{\text{r}}}\left(\frac{d_1 d_2}{ab\cos(\vartheta_i)}\right)^2 \tag{7}$$

where G_{t} and G_{r} are transmitting and receiving gains, respectively, $a\times b$ is the size of the RIS, d_1 represents the distance from the transmitter to the center of RIS, d_2 represents the distance from the center of RIS to the receiver. A more comprehensive far-field path loss model has been proposed in Ref. [21], taking the power radiation pattern and the gain of the RIS element into account. With the intelligent control of RIS elements, the signals received at the receiver are phase-aligned to achieve beamforming. Thus, the free-space path loss model for the far-field beamforming case can be represented by[21]

$$PL_{\text{far-field}}^{\text{beam}} = \frac{64\pi^3 (d_1 d_2)^2}{G_{\text{t}}G_{\text{r}}GN_x^2N_y^2 d_x d_y \lambda^2 F(\vartheta_{\text{t}},\psi_{\text{t}})F(\vartheta_{\text{r}},\psi_{\text{r}})A^2} \tag{8}$$

where G represents the gain of the RIS elements; N_x and N_y represent the number of rows and columns of RIS elements; $d_x\times d_y$ represents the size of an RIS element; ϑ_{t}, ψ_{t}, ϑ_{r}, and ψ_{r} represent the elevation angle and the azimuth angle from the center of the RIS to the transmitter, and the elevation angle and the azimuth angle from the center of the RIS to the receiver, respectively; $F(\vartheta,\psi)$ is the normalized power radiation pattern of the RIS element; A is the amplitude of the reflection coefficient for each RIS element.

Furthermore, the authors in Ref. [21] validated the performance of the proposed theoretical path loss models through extensive numerical simulations and experimental measurements. The prior studies[19–22] on RIS-aided communication have proved the far-field path loss model to be inversely proportional to $(d_1d_2)^2$, which is based on free-space. However, if we consider a more practical propagation environment by considering reflection, diffraction, and scattering, the far-field path loss model may be modified as inversely proportional to the higher power of d_1d_2.

3.2 Near-field

When the transmitter or the receiver is in the near-field of the RIS, EM wave receives at different RIS elements may have different angle of arrivals. The path loss modeling for near-field is more complicated. Reference [22] analyzed the near-field path loss model in two cases: near-field beamforming and near-field broadcasting. In the near-field beamforming case, the received signal power is only maximized for a specific user. The corresponding path loss model can be represented by

$$PL_{\text{near-field}}^{\text{beam}} = \frac{64\pi^3}{G_{\text{t}}G_{\text{r}}Gd_x d_y \lambda^2 A^2 \left|\sum_{x=1-u}^{u}\sum_{y=1-v}^{v}\frac{\sqrt{F_{\text{combine}}}}{r_{x,y}^t r_{x,y}^r}\right|^2} \tag{9}$$

where $u=\dfrac{N_x}{2}$, $v=\dfrac{N_y}{2}$, and $F_{\text{combine}} = F^{tx}(\vartheta_{x,y}^{tx},\psi_{x,y}^{tx})\times F(\vartheta_{x,y}^{t},\psi_{x,y}^{t})F(\vartheta_{x,y}^{r},\psi_{x,y}^{r})F^{rx}(\vartheta_{x,y}^{rx},\psi_{x,y}^{rx})$ represents the effect of the normalized power radiation patterns on the received signal power. Let $U_{x,y}$ denote the RIS element in the x-th row and y-th column, $\vartheta_{x,y}^{tx}$, $\psi_{x,y}^{tx}$, $\vartheta_{x,y}^{rx}$, and $\psi_{x,y}^{rx}$ represent the elevation angle and the azimuth angle from the transmitting antenna to $U_{x,y}$, and the elevation angle and the azimuth angle from the receiving antenna

to $U_{x,y}$, respectively. $\vartheta^t_{x,y}$, $\psi^t_{x,y}$, $\vartheta^r_{x,y}$, $\psi^r_{x,y}$, $r^t_{x,y}$, and $r^r_{x,y}$ are the elevation angle and the azimuth angle from $U_{x,y}$ to the transmitter, the elevation angle and the azimuth angle from $U_{x,y}$ to the receiver, the distance between the transmitter and $U_{x,y}$, the distance between the receiver and $U_{x,y}$, respectively. $F^{tx}(\vartheta,\psi)$ and $F^{rx}(\vartheta,\psi)$ denote the normalized power radiation patterns of the transmitter and the receiver, respectively.

Assuming that the RIS is large enough to cover all users in a specific area, Ref. [22] derived the near-field broadcasting path loss model. They advocated that the signal is transmitted from the mirror image of the transmitter by taking the RIS as the symmetric plane. On this condition, the propagation distance from the transmitter to the receiver is regarded as d_1+d_2. Therefore, the corresponding near-field path loss can be approximately expressed as

$$PL^{\text{broadcast}}_{\text{near-field}} \approx \frac{16\pi^2(d_1+d_2)^2}{G_{\text{t}}G_{\text{r}}A^2A^2} \tag{10}$$

Different from the above analysis of near-field path loss model, Refs. [23, 24] proposed the path loss between the two RISs from the perspective of microscopic EM field. In Ref. [23], the RIS is approximated as a continuous array of an infinite number of infinitesimal antennas, where any current distribution can be synthesized at the RIS. The source current of the transmitting RIS can generate the electric field satisfying the Helmholtz wave equation. The outgoing wave from any point source can be given by the tensor Green's function, which is the function of spatial coordinates. When calculating the path loss between the transmitting RIS and the receiving RIS, the double integral of the spatial position function is applied over the spatial area of the transmitting RIS and the receiving RIS.

4 Performance analysis for RIS

In this section, we will review the existing works that contributed to analyze the performance of RIS-aided wireless systems. Particularly, we will firstly introduce the asymptotic power scaling law, and then analyze the achievable rate for RIS-aided wireless systems.

4.1 Asymptotic power scaling law

A fundamental question that is concerned by most of the researchers is how much power gain can be achieved by using the RIS in future communication systems. To answer this question, some researchers have analyzed the asymptotic power scaling factor for RIS-aided wireless systems[13, 15, 25, 26]. They have studied the relation between the received power P_{r} at the user side and the number of the reflecting elements N at the RIS side such that when the number of RIS elements N tends to infinity, i.e., $N\to\infty$.

The asymptotic power scaling law considering ideal phase shifts is firstly characterized in Ref. [13], where they proved that the received power P_{r}, by carefully designing the reflecting coefficients, tends to increase quadratically with the increase in the number of reflecting elements N. To be more specific, Ref. [13] analyzed the received power of a single-antenna user supported by a single-antenna BS with the help of an RIS, assuming that the received power at the user side is dominated by the power of the reflected signal from the RIS[13, 25]. Neglecting the direct path between the BS and the user, the system model Eq. (1) reduces to $y=\boldsymbol{f}^{\text{T}}\boldsymbol{\Theta}\boldsymbol{g}x+n$, where $\boldsymbol{g}$ denotes the channel between the RIS and the BS. In this case, the received power P_{r} is determined by the transmit power and the channel gain, which can be expressed as

$$P_{\text{r}} = P_{\text{t}}\left|\boldsymbol{f}^{\text{T}}\boldsymbol{\Theta}\boldsymbol{g}\right|^2 = P_{\text{t}}\left|\sum_{n=1}^{N}|\boldsymbol{f}^{\text{T}}(n)||\boldsymbol{g}(n)|\text{e}^{\text{j}(\theta_n+\phi_n+\psi_n)}\right|^2 \tag{11}$$

where P_{t} is the transmit power at the BS, $\boldsymbol{f}\sim CN\left(0,\sigma_f^2\boldsymbol{I}\right)$ denotes the $N\times 1$ channel vector for the RIS-user link, $\boldsymbol{\Theta}=\text{diag}\left(\text{e}^{\text{j}\theta_1},\ldots,\text{e}^{\text{j}\theta_N}\right)$ is the diagonal matrix denoting the reflecting coefficients at the RIS, and $\boldsymbol{g}\sim CN(0,\sigma_g^2\boldsymbol{I})$ is the $N\times 1$ channel vector for the BS-RIS link. Note that for every element of the channel vectors $\boldsymbol{f}$ and $\boldsymbol{g}$, we assume that $\boldsymbol{f}^{\text{T}}(n)=\left|\boldsymbol{f}^{\text{T}}(n)\right|\text{e}^{\text{j}\phi_n}$ and $\boldsymbol{g}(n)=|\boldsymbol{g}(n)|\,\text{e}^{\text{j}\psi_n}$.

With the received power calculated in Eq. (11), the asymptotic power scaling law can be expressed by the following proposition.

Proposition 1. After optimizing the reflecting coefficients at the RIS, when $N\to\infty$, the received

power at the user side satisfies

$$P_{\mathrm{r}} \to N^2 \frac{P_{\mathrm{t}} \pi^2 \sigma_f^2 \sigma_g^2}{16} \tag{12}$$

Proof: In order to obtain maximum received power, the reflecting coefficients of the RIS should make the phase of each summation term $\mathrm{e}^{\mathrm{j}(\theta_n+\phi_n+\psi_n)}$ in Eq. (11) aligned with each other to achieve constructive summation, i.e., $\theta_n^* = -(\phi_n+\psi_n)$[13]. In this way, the maximum received power can be given by

$$P_{\mathrm{r}} = P_{\mathrm{t}} \left| \sum_{n=1}^{N} |\boldsymbol{f}^{\mathrm{T}}(n)||\boldsymbol{g}(n)| \right|^2 \tag{13}$$

where $|\boldsymbol{f}(n)|$ and $|\boldsymbol{g}(n)|$ for $n = 1, 2, \ldots, N$ are independent Rayleigh random variables with mean $\sqrt{\pi}\sigma_f/2$ and $\sqrt{\pi}\sigma_g/2$, respectively. According to central limit theorem, when $N \to \infty$, $\sum_{n=1}^{N} |\boldsymbol{f}^{\mathrm{T}}(n)||\boldsymbol{g}(n)|/N \to E(\boldsymbol{f}^{\mathrm{T}}(n)\boldsymbol{g}(n)) = \pi\sigma_f\sigma_g/4$. Thus, the asymptotic power received at the user when $N \to \infty$ can be written as

$$P_{\mathrm{r}} \to N^2 \frac{P_{\mathrm{t}} \pi^2 \sigma_f^2 \sigma_g^2}{16} \tag{14}$$

This completes the proof.

The quadratic relation in Formula (14) lies with the fact that the RIS not only has a receiving gain of order N by capturing a signal power proportional to N in the BS-RIS link but also achieves the transmit beamforming gain of order N, similar to an MIMO BS, in the RIS-user link[13]. The same work is also extended in the multi-RIS scenario[26] to realize a high gain. On the other hand, this quadratic relation cannot be achieved by scaling up the number of transmit antennas in the massive MIMO scenario because the transmit power does not scale up by the order of N. Interestingly, the power scaling law Formula (14) indicates that RIS enables us to trade space for transmit power. That is to say, by deploying a large number of low-cost reflecting elements at the RIS, we can scale down the transmit power of the BS by a factor of $1/N^2$ without compromising the received power at the user, which makes the system cost-effective and energy-efficient.

For the analysis mentioned above, the authors have assumed that the reflecting elements at the RIS can adjust their phase shifts with arbitrary precision, which may not be practical in real systems. Manufacturing each reflecting element of RIS with large quantization levels of phase shifts brings a higher hardware cost, not to mention the number of elements for RIS is usually very large[25]. Therefore, the asymptotic power scaling law under a more practical scenario, i.e., employing discrete phase shifts for the RIS elements, was investigated in Ref. [25]. Fortunately, the authors in Ref. [25] found that RIS with discrete phase shifts achieves the same squared power gain by deploying asymptotically large number of reflecting elements. Specifically, the b-bit phase shifts are assumed in Ref. [25], where the number of phase-shift levels L can be calculated as $L = 2^b$. The discrete phase-shift values are obtained by uniformly quantizing the interval $[0, 2\pi)$. Thus, the candidate set of discrete phase-shift values at each element is given by

$$\mathcal{F} = \{0, \Delta\theta, 2\Delta\theta, \ldots, (L-1)\Delta\theta\} \tag{15}$$

where $\Delta\theta = 2\pi/L$[25]. Other than the phase shifts, the considered RIS-aided communication scenario is the same as that with the ideal phase shifts. Therefore, the received power $P_{\mathrm{r}}(b)$ is given by

$$P_{\mathrm{r}}(b) = P_{\mathrm{t}} \left| \boldsymbol{f}^{\mathrm{T}} \boldsymbol{\Theta} \boldsymbol{g} \right|^2 = P_{\mathrm{t}} \left| \sum_{n=1}^{N} |\boldsymbol{f}^{\mathrm{T}}(n)||\boldsymbol{g}(n)| \mathrm{e}^{\mathrm{j}(\theta_n+\phi_n+\psi_n)} \right|^2 \tag{16}$$

where $\theta_n \in \mathcal{F}$. Then, the asymptotic power scaling law with discrete phase shifts can be formulated by the following proposition.

Proposition 2. After optimizing the reflection coefficients at the RIS, when $N \to \infty$, the received power at the user side with discrete phase shifts satisfies

$$P_{\mathrm{r}}(b) \to N\sigma_f^2\sigma_g^2 + N(N-1)\frac{\pi^2\sigma_f^2\sigma_g^2}{16}\left(\frac{2^b}{\pi}\sin\left(\frac{\pi}{2^b}\right)\right)^2 \tag{17}$$

Proof: According to Proposition 1, the optimal phase shifts can be calculated as $\theta_n^* = (-)(\phi_n+\psi_n)$[13]. Thus, it can be verified that for discrete phase shifts, the maximum received power can be achieved by setting $\theta_n = \arg\min_{\theta\in\mathcal{F}} \left|\theta - \theta_n^*\right|$[25]. Let $\bar{\theta}_n = \theta_n + \phi_n + \psi_n = \theta_n - (-)(\phi_n+\psi_n) = \theta_n - \theta_n^*$ denote the quantization error of the discrete phase shifts with respect to the optimized phase shifts, we have

$$\begin{aligned} &P_{\mathrm{r}}(b)= \\ &P_{\mathrm{t}}\left|\sum_{n=1}^{N}|\boldsymbol{f}^{\mathrm{T}}(n)\|\boldsymbol{g}(n)|\mathrm{e}^{\mathrm{j}\bar{\theta}_n}\right|^2= \\ &P_{\mathrm{t}}\left(\sum_{n=1}^{N}|\boldsymbol{f}^{\mathrm{T}}(n)|^2|\boldsymbol{g}(n)|^2+\right. \\ &\left.\sum_{n=1}^{N}\sum_{i\neq n}^{N}|\boldsymbol{f}^{\mathrm{T}}(n)\|\boldsymbol{g}(n)\|\boldsymbol{f}^{\mathrm{T}}(i)\|\boldsymbol{g}(i)|\mathrm{e}^{\mathrm{j}\bar{\theta}_n-\mathrm{j}\bar{\theta}_i}\right) \end{aligned} \tag{18}$$

It is worth pointing out that $\boldsymbol{f}^{\mathrm{T}}(n)$, $\boldsymbol{g}(n)$, and $\mathrm{e}^{\mathrm{j}\bar{\theta}_n}$ are independent random variables. Moreover, as $\{\theta_n\}_{n=1}^{N}$ is uniformly distributed in $\mathcal{F}$, the quantization error $\{\bar{\theta}_n\}$ is also uniformly distributed in $[-\pi/2^b,\pi/2^b]$, and thus $\mathcal{E}(\mathrm{e}^{\mathrm{j}\bar{\theta}_n})=\mathcal{E}(\mathrm{e}^{-\mathrm{j}\bar{\theta}_i})=2^b/\pi\sin(\pi/(2^b))$. Therefore, according to the central limit theorem, when $N\to\infty$, we have $\sum_{n=1}^{N}|\boldsymbol{f}^{\mathrm{T}}(n)|^2|\boldsymbol{g}(n)|^2\to N\sigma_f^2\sigma_g^2$, and $\sum_{n=1}^{N}\sum_{i\neq n}^{N}|\boldsymbol{f}^{\mathrm{T}}(n)\|\boldsymbol{g}(n)\|\boldsymbol{f}^{\mathrm{T}}(i)\|\boldsymbol{g}(i)|\mathrm{e}^{\mathrm{j}\bar{\theta}_n-\mathrm{j}\bar{\theta}_i}\to N(N-1)\frac{\pi^2\sigma_f^2\sigma_g^2}{16}\left(\frac{2^b}{\pi}\sin\left(\frac{\pi}{2^b}\right)\right)^2$, which completes the proof.

Proposition 2 shows that the asymptotic quadratic power gain of $O(N^2)$ can also be achieved by using discrete phase shifts at the RIS. Thus, the more cost-effective RIS with discrete phase shifts can be deployed without compromising the performance in the large-N regime.

4.2 Achievable rate analysis

We have discussed the asymptotic power scaling law in the previous subsection, we now present the achievable rate analysis for RIS-aided wireless communication systems. Particularly, the analysis introduced in Subsection 4.1 assumes an asymptotically large N at the RIS. However, the number of reflecting elements N is usually finite and limited in the real-time system. But how large N is required to achieve a superior achievable rate compared with traditional point-to-point communication systems remains unknown. Therefore, in this section, we mainly focus on answering the question "how large N is needed for RIS-aided wireless communication system to surpass the achievable spectral performance of traditional systems".

To answer this question, Ref. [27] compared the achievable rate of three systems, namely the traditional single-input single-output (SISO) system, the half-duplex decode-and-forward (DF) relay-aided communication system, and the RIS-aided communication system. To be more specific, the relay is equipped with one antenna, and the RIS is manufactured with N reflecting elements. The channels for the BS-relay link, the relay-user link, and the BS-user direct link are denoted by h_{sr}, h_{rd}, and h_{sd}, respectively, with $|h_{\mathrm{sr}}|=\sqrt{\beta_{\mathrm{sr}}}$, $|h_{\mathrm{rd}}|=\sqrt{\beta_{\mathrm{rd}}}$, and $|h_{\mathrm{sd}}|=\sqrt{\beta_{\mathrm{sd}}}$. Besides, the channel vectors of size $N\times 1$ for the BS-RIS link and the RIS-user link are denoted by $\boldsymbol{g}$ and $\boldsymbol{f}$, respectively. To ensure fair comparison, the RIS and the DF relay are deployed at the same location. As a result, all elements in $\boldsymbol{g}$ have the same magnitude as h_{sr}, and all elements in $\boldsymbol{f}$ have the same magnitude as h_{rd}. With $\boldsymbol{\Theta}=\mathrm{diag}(\theta_1,\theta_2,\ldots,\theta_N)$ denoting the diagonal matrix for the reflection coefficients of the RIS, the achievable rate for the RIS-aided system can be given by

$$\begin{aligned} &R_{\mathrm{RIS}}(N)= \\ &\max_{\theta_1,\ldots,\theta_N}\log_2\left(1+\frac{P\left|h_{\mathrm{sd}}+A\boldsymbol{f}^{\mathrm{T}}\boldsymbol{\Theta}\boldsymbol{g}\right|^2}{\sigma^2}\right)= \\ &\log_2\left(1+\frac{P\left(|h_{\mathrm{sd}}|+A\sum_{n=1}^{N}|\boldsymbol{f}(n)\|\boldsymbol{g}(n)|\right)^2}{\sigma^2}\right)= \\ &\log_2\left(1+\frac{P\left(\sqrt{\beta_{\mathrm{sd}}}+NA\sqrt{\beta_{\mathrm{RIS}}}\right)^2}{\sigma^2}\right) \end{aligned} \tag{19}$$

where P is the transmit power at the BS, $A\in[0,1]$ denotes the fixed amplitude reflection coefficient, σ^2 is the noise power, and $\sqrt{\beta_{\mathrm{RIS}}}=\sum_{n=1}^{N}|\boldsymbol{f}(n)\|\boldsymbol{g}(n)|/N=\sqrt{\beta_{\mathrm{sr}}\beta_{\mathrm{rd}}}$. Note that this expression can be derived using the same method as that in Proposition 1, i.e., the phase shifts should be aligned to enable constructive summation[13, 27]. Besides, we can readily give the achievable rates for the traditional SISO system and the DF relay-aided system according to Ref. [28] as

$$R_{\mathrm{SISO}}=\log_2\left(1+\frac{P\beta_{\mathrm{sd}}}{\sigma^2}\right) \tag{20}$$

$$R_{\mathrm{DF}}=\frac{1}{2}\log_2\left(1+\min\left(\frac{P_1\beta_{\mathrm{sr}}}{\sigma^2},\frac{P_1\beta_{\mathrm{sd}}}{\sigma^2}+\frac{P_2\beta_{\mathrm{rd}}}{\sigma^2}\right)\right) \tag{21}$$

where P_1 and P_2 are transmit powers of the BS and DF relay, respectively.

With the achievable rates obtained above, the achievable rate comparison of the traditional SISO

system and the DF relay-aided system is characterized in the following proposition[27].

Proposition 3. Assume that P_1 and P_2 are selected under the constraint $P=(P_1+P_2)/2$. If $\beta_{sd}>\beta_{sr}$, it holds that $R_{SISO}>R_{DF}$ for any P_1, P_2. If $\beta_{sd}\leqslant\beta_{sr}$, the achievable rate of the DF relay-aided system is maximized by $P_1=2P\beta_{rd}/(\beta_{sr}+\beta_{rd}-\beta_{sd})$ and $P_2=2P\cdot(\beta_{sr}-\beta_{sd})/(\beta_{sr}+\beta_{rd}-\beta_{sd})$, leading to

$$R_{DF}=\frac{1}{2}\log_2\left(1+\frac{2P\beta_{rd}\beta_{sr}}{(\beta_{sr}+\beta_{rd}-\beta_{sd})\sigma^2}\right) \tag{22}$$

Proof: As stated in Ref. [27], if $\beta_{sd}>\beta_{sr}$, $\min\left(\frac{P_1\beta_{sr}}{\sigma^2},\frac{P_1\beta_{sd}}{\sigma^2}+\frac{P_2\beta_{rd}}{\sigma^2}\right)=\frac{P_1\beta_{sr}}{\sigma^2}$, which is maximized by $P_1=2P$ and $P_2=0$. In this case, the DF relay is actually not used and thus $R_{SISO}>R_{DF}$. If $\beta_{sd}\leqslant\beta_{sr}$, R_{DF} is maximized by ensuring $\frac{P_1\beta_{sr}}{\sigma^2}=\frac{P_1\beta_{sd}}{\sigma^2}+\frac{P_2\beta_{rd}}{\sigma^2}$ under the constraint $P=(P_1+P_2)/2$, and this gives the expression Eq. (22) after some mathematical manipulations.

Proposition 3 introduced above lays a foundation for the analysis of the RIS-aided wireless communication systems. Based on this proposition, the following proposition is given in Ref. [27] to calculate the required number of reflecting elements at the RIS which ensures a superior performance in terms of the achievable rate.

Proposition 4. Compared with the traditional SISO system and the DF relay-aided communication system, the RIS-aided communication system provides the highest achievable rate if $\beta_{sd}>\beta_{sr}$. Else if $\beta_{sd}\leqslant\beta_{sr}$, it provides the highest rate if and only if

$$N>\frac{\sqrt{\left(\sqrt{1+\frac{2P\beta_{rd}\beta_{sr}}{(\beta_{sr}+\beta_{rd}-\beta_{sd})\sigma^2}}-1\right)\frac{\sigma^2}{P}}-\sqrt{\beta_{sd}}}{A\sqrt{\beta_{RIS}}} \tag{23}$$

Proof: As shown in Eqs. (19) and (20), the achievable rate of the RIS-aided system is always larger than that of the traditional SISO system, i.e., $R_{RIS}(N)>R_{SISO}$, since $N\geqslant 1$[27]. Therefore, RIS provides the highest achievable rate if and only if $R_{RIS}(N)>R_{DF}$. When $\beta_{sd}>\beta_{sr}$, we have $R_{SISO}>R_{DF}$ as stated in Proposition 3, and thus $R_{RIS}(N)>R_{SISO}>R_{DF}$. When $\beta_{sd}\leqslant\beta_{sr}$, the inequality $R_{RIS}(N)>R_{DF}$ can be simplified to Formula (23) by using Eqs. (19) and (22).

Now we interpret the result given by Proposition 4. Although the RIS-aided system yields the best performance in terms of the achievable rate for any N when $\beta_{sd}>\beta_{sr}$, its performance gain compared with the traditional SISO system is not that obvious in this case. This is because β_{rd} is usually a small number in practice[27], and thus $\sqrt{\beta_{sd}}>>NA\sqrt{\beta_{RIS}}=NA\sqrt{\beta_{sr}\beta_{rd}}$ in this case. As a result, the achievable rate for RIS-aided system R_{RIS} in Eq. (19) is actually dominated by $\log_2\left(1+\frac{P\beta_{sd}}{\sigma^2}\right)$. When $\beta_{sd}\leqslant\beta_{sr}$, RIS will provide the highest achievable rate if and only if N satisfies Formula (23). It is worth pointing out that the right-hand side of Formula (23) approaches $-\frac{\sqrt{\beta_{sd}}}{A\sqrt{\beta_{sr}\beta_{rd}}}<0$ when $P\to\infty$, which implies that RIS can provide the largest rate at high signal-to-noise ratio (SNR) for any $N>0$. Besides, the inequality in Formula (23) becomes

$$N>\frac{\sqrt{\frac{1}{\beta_{sr}+\beta_{rd}-\beta_{sd}}}-\frac{\sqrt{\beta_{sd}}}{\sqrt{\beta_{sr}\beta_{rd}}}}{A} \tag{24}$$

for $P\to 0$, which can be quite large. For instance, Formula (24) becomes $N>963$ for $A=1$, $\beta_{sd}=-110$ dB, $\beta_{sr}=-80$ dB, and $\beta_{rd}=-60$ dB[27]. To conclude, compared with the traditional SISO system and the single-antenna DF relay-aided system, RIS can provide higher achievable rate as long as it is equipped with sufficient reflecting elements, which is mainly due to the receiving gain and beamforming gain introduced by large array.

Apart from Ref. [27], Ref. [29] compared the achievable rate of multiple RIS-aided wireless communication systems and a single-antenna full-duplex DF relay-aided communication system, both of which consider a single-cell downlink scenario with multiple users served by one BS. The multiple RISs, as well as the full duplex DF relays, are randomly deployed, and the achievable rates of the users are averaged over the random locations as well as the wireless channel fading, which are defined as spatial throughput. With the theoretical analysis and Monte-Carlo simulations, Ref. [29] proved that the multiple RIS-aided systems outperform the full-duplex DF relay-aided system in terms of the spatial throughput with sufficient reflecting elements deployed at each

RIS (e.g., $N = 2000$), which again verifies the result given in Proposition 4.

5 Precoding for RIS

As a fundamental technique in MIMO systems, precoding makes the signals sent by the multi-antenna base station (BS) more directional instead of radiating around, which enhances the possibility of the signal reception at users and increases the system capacity. In RIS-aided communication systems, the novel precoding technique called joint precoding is the key guarantee for RIS to improve the system performance. Different from the conventional system in which precoding is done only at the BS, in the RIS-aided communication system, we need to jointly optimize the active beamforming at the BS and the reflection coefficients at the RIS. Up to now, numerous joint precoding design problems with different design goals in different application scenarios have been widely investigated in existing works. In the following subsections, we summarize some typical RIS application scenarios and the corresponding design goals.

5.1 RIS-aided single-user transmission scenario

(1) Capacity enhancement problem formulation: Based on the system model in Eq. (1), we can further write the SNR of the user as

$$\gamma = \frac{\left|\left(\boldsymbol{h}^{\mathrm{T}} + \boldsymbol{f}^{\mathrm{T}}\boldsymbol{\Theta}\boldsymbol{G}\right)\boldsymbol{w}\right|^2}{\sigma^2} \tag{25}$$

Note that we assume the transmitted symbol s has normalized power, i.e., $\mathcal{E}\{|s|^2\} = 1$. Then, we can consider the most typical joint precoding problem of SNR maximization. This typical problem has been formulated and discussed in Refs. [30−32], which can be written as

$$\begin{aligned} \max_{\boldsymbol{w},\boldsymbol{\Theta}} \quad & \left|\left(\boldsymbol{h}^{\mathrm{T}} + \boldsymbol{f}^{\mathrm{T}}\boldsymbol{\Theta}\boldsymbol{G}\right)\boldsymbol{w}\right|^2, \\ \text{s.t.} \quad & \mathrm{C1}: \ \|\boldsymbol{w}\|^2 \leqslant P_{\max}, \\ & \mathrm{C2}: \ \theta_n \in \mathcal{F},\ n \in \{1,\ldots,N\} \end{aligned} \tag{26}$$

where $P_{\max}$ denotes the maximal transmit power of the BS, and $\mathcal{F}$ denotes the feasible set of the reflection coefficient at the RIS, which depends on the RIS hardware. For example, if the RIS element is able to continuously and independently control the phase shift of the incident signals, we have $\mathcal{F} = \left\{\theta_n \middle| 0 \leqslant \theta_n < 2\pi\right\}$. While if the RIS elements are 2-bit phase-shifters, we have $\mathcal{F} = \{0, \pi/2, \pi, 3\pi/2\}$. As shown in Formula (26), to maximize the SNR, the goal is to jointly design the active precoding vector $\boldsymbol{w}$ and the phase shift matrix $\boldsymbol{\Theta}$. To solve this problem, the authors in Ref. [30] proposed two optimization methods which are based on semidefinite relaxation (SDR) and alternating optimization, and the authors in Ref. [31] solved it by combining alternating maximization with the majorization-minimization method. According to their results, the user's SNR can be significantly increased with the aid of the RIS.

(2) Other extension cases: In RIS-aided communication systems, the applications and design goals for joint precoding vary in different scenarios. For example, for the applications in millimeter-wave (mmWave) systems, by exploiting the rank-one structure of the mmWave channel, the authors in Ref. [26] have put forward a closed-form solution to maximize the SNR. Then, the model is also extended to orthogonal frequency division multiplexing (OFDM) based wideband system in Ref. [33], where the researchers considered maximizing the sum-rate of multiple subcarriers for a single user based on their proposed alternative optimization method. To further reduce the symbol error rate (SER), the authors in Ref. [34] extended the model to the MIMO system and proposed a joint precoding method based on the semidefinite programming (SDP) and minimum Euclidean distance (MED).

5.2 RIS-aided multi-user transmission scenario

The above joint precoding designs are straightforward, since the considered single-user case is easy to analyze. In this case, many classical precoding design problems are almost equivalent to SNR maximization, e.g., transmit power minimization. However, the multi-user scenario is more realistic in practical communication systems. In multi-user scenarios, numerous design goals for joint precoding such as sum-rate/energy efficiency maximization[35] are much more challenging, which also have been discussed in the

existing works.

The multi-user case can be extended from the single-user case. Figure 4 shows the typical scenario, where one multi-antenna BS aided by one RIS serves U single-antenna users. Here we denote the channel responses from the BS to the u-th user and from the RIS to the u-th user by $\boldsymbol{h}_u$ and $\boldsymbol{f}_u$, $u=1,2,\ldots,U$, respectively. The transmitted signal at the BS as $\boldsymbol{x}=\sum_{u=1}^{U}\boldsymbol{w}_u s_u$, where s_u and $\boldsymbol{w}_u$ denote the transmitted symbol and the precoding vector of the u-th user, respectively. Therefore, the received signal-to-interference-plus-noise ratio (SINR) of user u can be written as

$$\gamma_u=\frac{\left|\left(\boldsymbol{h}_u^{\mathrm{T}}+\boldsymbol{f}_u^{\mathrm{T}}\boldsymbol{\Theta}\boldsymbol{G}\right)\boldsymbol{w}_u\right|^2}{\sum\limits_{j=1,j\neq u}^{U}\left|\left(\boldsymbol{h}_u^{\mathrm{T}}+\boldsymbol{f}_u^{\mathrm{T}}\boldsymbol{\Theta}\boldsymbol{G}\right)\boldsymbol{w}_j\right|^2+\sigma^2} \tag{27}$$

Then, based on this expression, various multi-user precoding design goals have been investigated in the literature. Here we introduce some typical optimization goal in the following parts, including capacity enhancement, energy saving, and user fairness.

(1) Capacity enhancement: Capacity enhancement plays the central role in the researches about precoding. Similar to Formula (26), we can formulate the multi-user sum-rate maximization problem as

$$\begin{aligned}&\max_{\boldsymbol{w},\boldsymbol{\Theta}}\ \sum_{u=1}^{U}\log_2(1+\gamma_u),\\ &\text{s.t.}\quad \mathrm{C1}:\ \sum_{u=1}^{U}\|\boldsymbol{w}_u\|^2\leqslant P_{\max},\\ &\qquad \mathrm{C2}:\ \theta_n\in\mathcal{F},\ n\in\{1,\ldots,N\}\end{aligned} \tag{28}$$

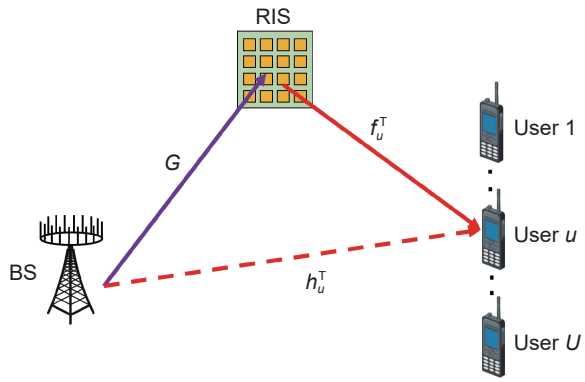

Fig. 4 System model of downlink transmission in RIS-aided communication multi-user system.

Without considering the direct link $\boldsymbol{h}$, the authors in Ref. [31] first solved this problem by combining alternating maximization with the majorization-minimization method. While, the alternate optimization method[13] can find a sub-optimal solution. Then, based on this problem, the achievable sum-rate of this problem is analyzed in Ref. [36], and many extended problems have been further considered in Refs. [11, 37−41]. To consider more general optimization objective, the authors in Ref. [11] have extended the sum-rate maximization problem into weighted sum-rate (WSR), while some researchers extended it to effective sum-rate (ESR) maximization[37]. To further reduce power consumption in mmWave system, the sum-rate maximization problem while applying hybrid precoding has been studied in Ref. [38]. In addition, to apply in wideband scenario, the same problem is also extended to the OFDM-based system[39]. Particularly, the authors in Ref. [41] have proposed a joint precoding framework as a general solution to maximize the WSR in RIS-aided communication system, since their proposed RIS-aided cell-free network has taken multi-BS, multi-RIS, and multi-user in MIMO-OFDM system into consideration simultaneously[42]. The proposed method can subsume the above papers as its special cases.

(2) Energy saving: In addition to enhance the capacity, researchers have also tried to save the transmit power with the aid of RIS. They have formulated the transmit power minimization problem as

$$\begin{aligned}&\min_{\boldsymbol{w},\boldsymbol{\Theta}}\ \sum_{u=1}^{U}\|\boldsymbol{w}_u\|^2,\\ &\text{s.t.}\quad \mathrm{C1}:\ \gamma_u>\gamma_{\min},\\ &\qquad \mathrm{C2}:\ \theta_n\in\mathcal{F},\ n\in\{1,\ldots,N\}\end{aligned} \tag{29}$$

where $\gamma_{\min}$ denotes the minimal quality of service (QoS) guarantee of the user. This problem has been analyzed in Ref. [13], where the authors proposed a sub-optimal alternating optimization method to solve it. Then, this problem is also extended to a scenario where RIS has limited phase-shift bits[12, 25]. In addition, apart from setting the transmit power as the design goal, some researchers also considered the energy efficiency

as the optimization objective[43], which can be written as

$$\begin{aligned}\max_{\boldsymbol{w},\boldsymbol{\Theta}}\ & \frac{\sum_{u=1}^{U}\log_2(1+\gamma_u)}{\xi\sum_{u=1}^{U}\|\boldsymbol{w}_u\|+P_{\mathrm{s}}},\\ \text{s.t.}\quad & \mathrm{C1}:\ \sum_{u=1}^{U}\|\boldsymbol{w}_u\|^2\leqslant P_{\max},\\ & \mathrm{C2}:\ \theta_n\in\mathcal{F},\ n\in\{1,\ldots,N\}\end{aligned}\tag{30}$$

where ξ denotes the efficiency of the transmit power amplifier, while P_{s} denotes the static power consumption at BS, RIS, and users. By using sequential fractional programming and conjugate gradient search, this problem can be alternatively solved.

(3) User fairness: Apart from capacity enhancement and energy saving, the fairness among multiple users is also a practically important problem since it can directly affect the user experience. Mathematically, it is a max-min SINR problem[44]

$$\begin{aligned}\max_{\boldsymbol{w},\boldsymbol{\Theta}}\ & \min_{u}\{\gamma_u\},\\ \text{s.t.}\quad & \mathrm{C1}:\ \sum_{u=1}^{U}\|\boldsymbol{w}_u\|^2\leqslant P_{\max},\\ & \mathrm{C2}:\ \theta_n\in\mathcal{F},\ n\in\{1,\ldots,N\},\\ & \mathrm{C3}:\ \gamma_u>\gamma_{\min}\end{aligned}\tag{31}$$

This problem can be solved by the method based on the deterministic approximations proposed in Ref. [44]. Then, to analyze this design goal in more general scenario, the authors in Ref. [45] have extended this problem to the multi-RIS multiple-input single-output (MISO) system. Then, this problem has also been extended to a cellular network scenario in Ref. [46], in which RISs can be exploited to reduce the inter-cell interference when the users are near the cell boundary.

(4) Physical layer security: During the signal transmission between the authorized users, an unknown transceiver can intercept the wireless channel and may disrupt the communication between the legitimate users that causes a serious threat to an on-going communication. Therefore, safeguarding the wireless communication system against an eavesdropper/s is equally important. The system model of the RIS-aided downlink wireless communication system considering physical layer security scenario is illustrated in Fig. 5.

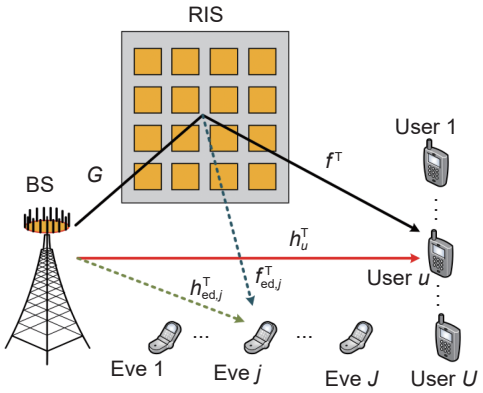

Fig. 5 System model of RIS-aided communication under security threat.

Mathematically, the received signal for the legtimale u-th user can be given by using Eq. (1), while an extra received signal for the j-th eavesdropper (eve) can be given by

$$y_{\mathrm{eve}}=\left(\boldsymbol{h}_{\mathrm{ed},j}^{\mathrm{T}}+\boldsymbol{f}_{\mathrm{ed},j}^{\mathrm{T}}\boldsymbol{\Theta}\boldsymbol{G}\right)\boldsymbol{w}_u x+n_{\mathrm{evs},j}\tag{32}$$

Thus, the SINR for the leg u-th user can be written as in Eq. (27), while the SINR for the j-th eve can be written as

$$\gamma_{u,j}=\frac{\left|\left(\boldsymbol{h}_{\mathrm{ed},j}^{\mathrm{T}}+\boldsymbol{f}_{\mathrm{ed},j}^{\mathrm{T}}\boldsymbol{\Theta}\boldsymbol{G}\right)\boldsymbol{w}_u\right|^2}{\left|\left(\boldsymbol{h}_{\mathrm{ed},j}^{\mathrm{T}}+\boldsymbol{f}_{\mathrm{ed},j}^{\mathrm{T}}\boldsymbol{\Theta}\boldsymbol{G}\right)\boldsymbol{w}_j\right|^2+\sigma_{\mathrm{evs},j}^2},\forall u\in U,j\in J\tag{33}$$

Thus, the sum-rate for the authorized u-th user can then be computed as $R_{\mathrm{au},u}=\{\log_2(1+\gamma_u)-\max_{j^*}[\log_2(1+\gamma_{u,j^*})]\}$. Then, the sum-rate maximization problem can be written as

$$\begin{aligned}\max_{\boldsymbol{w},\boldsymbol{\Theta}}\ & R_{\mathrm{au},u},\\ \text{s.t.}\quad & \mathrm{C1}:\ \|\boldsymbol{w}_u\|^2\leqslant\rho_u,\\ & \mathrm{C2}:\ |\theta_n|=1,\ n\in\{1,\ldots,N\}\end{aligned}\tag{34}$$

where ρ_u represents the maximum allowable transmit power for the u-th user. The formulated problem can be reduced to the single user case by setting $U=1$[47]. The authors in Ref. [47] adopted an alternating optimization method to solve the problem. To this end, they at first disintegrated the original problem into multiple subproblems by fixing one after the other, and solved them alternatively. Moreover, a self-sustainable IRS-aided wireless communication system to enhance the security of the wireless system was discussed in Ref. [48]. The authors proposed the idea that an RIS

may have the potential to simultaneously act as a reflector and harvest energy from the incoming EM wave. Benefitting from a large number of reflecting elements employed at RIS, the author proposed that some of the elements can be used to reflect the signals while the rest of them can be exploited to harvest the energy from the received signal.

(5) Satisfying latency requirement: Besides studying the spectral enhancement, energy-saving, and ensuring the security for the wireless channel for RIS-aided communication system, it is also important to achieve the minimum latency requirement even in bad channel conditions to enjoy the perks of 6G communications. For instance, millimeter-wave (mmWave) can provide a huge bandwidth resource, nevertheless, it is highly susceptible to the blockage problem that causes a serious issue to the low-latency services. Therefore, the researchers in Ref. [49] integrated RIS with mmWave to minimize the blockage problem associated with mmWave. Mathematically, they have formulated the total transmitted power minimization problem while ensuring the latency requirement as

$$\begin{aligned} \min_{\boldsymbol{\theta},\boldsymbol{P}} \quad & \sum_{u=1}^{U} p_u, \\ \text{s.t.} \quad & \text{C1}: \ |\theta_n| \in [0, 2\pi),\ n \in \{1,\ldots,N\}, \\ & \text{C2}: \ p_u \geqslant 0,\ \forall u \in U, \\ & \text{C3}: \ \frac{D_u}{\alpha \log(1+\gamma_k)} \leqslant T,\ \forall u \in U \end{aligned} \tag{35}$$

where α and D_u represent the weighing factor and data size, respectively. The constraint C1 denotes that the range of the RIS elements can lie anywhere between 0 to 2π. C2 shows that optimized power allocated to the u-th user will always be positive, while C3 ensures the low-latency requirement even in bad channel conditions.

6 CSI acquisition for RIS-aided wireless communications

In this section, we summarize several CSI acquisition schemes in RIS-aided wireless communications. To improve the performance of the joint precoding at the BS and RIS, it is important to acquire CSI by performing channel estimation in the RIS-aided wireless communications. A number of channel estimation schemes have been proposed in Refs. [50−52] for the conventional MIMO systems. However, these schemes can not be directly applied in the RIS-aided systems, since the CSI estimation problem for RIS-aided systems is different from the conventional MIMO systems. In the RIS-aided systems, the channel between the BS and the UE consists of two parts: the indirect link (i.e., the BS-RIS-UE cascaded channel), and the direct link (i.e., the direct BS-UE channel).

Since the RIS has the potential to control the radio channel, the real-time equivalent channel depends on the RIS tuning coefficients $\boldsymbol{\theta}$ at the moment. The equivalent channel $\tilde{\boldsymbol{H}}(\boldsymbol{\theta}) \in \mathbb{C}^{M\times K}$ is given by

$$\tilde{\boldsymbol{H}}(\boldsymbol{\theta}) = \boldsymbol{H} + \boldsymbol{G}^{\mathrm{T}} \operatorname{diag}(\boldsymbol{\theta}) \boldsymbol{F} \tag{36}$$

where $\boldsymbol{H}$, $\boldsymbol{F}$, $\boldsymbol{\theta}$, and $\boldsymbol{G}$ are defined in Eq. (2). Then, Eq. (2) can be simplified as $\boldsymbol{y} = \tilde{\boldsymbol{H}}(\boldsymbol{\theta})^{\mathrm{T}} \boldsymbol{w} x + \boldsymbol{n}$. If the UE has only one antenna (i.e., $K = 1$), the Eq. (36) can be rewritten as

$$\tilde{\boldsymbol{h}}(\boldsymbol{\Theta}) = \boldsymbol{h} + \boldsymbol{G}^{\mathrm{T}} \operatorname{diag}(\boldsymbol{f}) \boldsymbol{\theta} = \boldsymbol{h} + \boldsymbol{H}_c^{\mathrm{T}} \boldsymbol{\theta} \tag{37}$$

where $\boldsymbol{f} \in \mathbb{C}^{N\times 1}$, when $U = 1$ in Eq. (37). It is worth noting that $\boldsymbol{H}_c \triangleq \operatorname{diag}(\boldsymbol{f}) \boldsymbol{G}$ necessarily and sufficiently describes the signal dependencies of the reflecting link. Henceforth, we call $\boldsymbol{H}_c$ the BS-RIS-UE cascaded channel, and we estimate $\boldsymbol{H}_c$ without the distraction brought by the RIS tuning coefficients $\boldsymbol{\theta}$.

The CSI acquisition for the BS-RIS-UE cascaded channel is challenging due to the following reasons. Firstly, the BS-RIS-UE cascaded channel has the linear relationship between the received signal, the transmitted signal, and the RIS tuning coefficients, which means that the system model is more complex than the conventional MIMO systems. Secondly, the RIS has no radio frequency (RF) chains or signal processing modules. It can neither transmit nor receive pilots (reference signals), thus, we can only indirectly acquire the CSI related to the RIS based on the pilot transmission between the BS and the UE. Thirdly, the BS-RIS-UE cascaded channel has a high dimension due to the large number of BS antennas and RIS elements. Hence, the large dimensional channel brings

more overhead while computing CSI in the RIS-aided wireless communications. This important yet challenging problem has attracted much attention of the researchers in the recent years[48, 53−74].

It is worth noting that there are some differences between the CSI acquisition problem for single-antenna users and that for multi-antenna users. Specifically, for single-antenna users, we can estimate the BS-RIS-UE cascaded channel $\boldsymbol{H}_c$ based on the signal model in Eq. (37). However, we can not express the BS-RIS-UE cascaded channel as a matrix in the scenario of multi-antenna users. Hence, when there are multiple antennas at the users, it is necessary to design different schemes for channel estimation, since we can not estimate the BS-RIS-UE cascaded channel directly unless introducing extremely high complexity. In this paper, we divide existing channel estimation schemes in the RIS-aided wireless communications into two categories: channel estimation for single-antenna users and that for multi-antenna users.

6.1 Channel estimation for single-antenna users

The authors in Refs. [53−64] considered channel estimation for a single-antenna user in the RIS-aided wireless communications. Since there are direct and indirect channel links between BS and users, i.e., the direct path is BS-UE while the indirect path is BS-RIS-UE, the channel available at both of the links is estimated successively. In Ref. [53], a channel estimation scheme based on the minimum mean squared error (MMSE) with several sub-phases was proposed. For the first sub-phase, the elements of RIS are set to "OFF" state (i.e., the value of elements in tuning coefficient vector are zero) to estimate the direct BS-UE channel. On the other hand, the channels available at the indirect path BS-RIS-UE are estimated in N sub-phases, where every RIS element is turned "ON" (i.e., the value of tuning coefficient for the RIS element is non-zero) in different sub-phases. With the increase of the number of RIS elements, the number of sub-phases will increase equally, which imports more pilot overhead for channel estimation, since the total pilot overhead for channel estimation is proportional to the number of sub-phases.

To reduce the CSI estimation overhead, Jensen and Carvalho[54] proposed an optimal channel estimation (CE) scheme, where all RIS elements are set to "ON" during the channel estimation. The channel estimation scheme was designed according to the Cremer-Rao lower bound model. Since the channel estimation usually relies on the RIS elements, the authors presented the channel estimation scheme where RIS elements follow a sequence of activation patterns that mimics the discrete Fourier transform (DFT). Till now, we only have discussed the literature that considered a single user based system model. However, in the wireless communication systems, generally the number of users is very large, and hence it is crucial to investigate the CSI of the individual user in a multi-user MIMO communication systems. Considering the angular channel sparsity, the authors in Ref. [55] considered the channel estimation problem in the RIS-aided multi-user mmWave massive MIMO system. Specifically, they have proposed a framework for downlink pilot transmission, and then based on known line-of-sight (LoS) dominated BS-RIS channel, the channels between BS to user and RIS-user are estimated using a compressive sensing algorithm. Additionally, in the recent past, deep learning (DL) has gained enormous attention from researchers due to its excellent performance in solving nonlinear problems. DL has been used in wireless communications in solving precoding as well as channel estimation problems, for instance, a DL-based channel estimation technique was proposed in Ref. [56]. In the proposed scheme, each user uses a deep neural network (DNN) where the input of the DNN is the received pilots while the output of the DNN is to estimate the channels of the user at the direct and indirect links. Simulation results show that the DL-based channel estimation scheme in this paper provides a lower normalized mean square error (NMSE) compared to the traditional least-squares algorithm.

In the multi-user RIS-aided communication, it is worth mentioning here that different users share the same BS-RIS channel, while a different channel vector for the different users is available in the link at RIS-user in the overall channel link BS-RIS-UE.

Benefitting from this, Wang et al.[57] tried to reduce the huge pilot overhead caused by a large number of BS antennas and RIS elements while acquiring the CSI. Usually, a single cell consisting of a BS with M antennas, U single-antenna users, and an RIS with N reflecting elements requires $UMN+UM$ channel coefficients to be estimated. To accurately estimate such a large number of channel coefficients within a short interval of time, Ref. [57] investigated an efficient pilot based channel estimation method in RIS-aided communications. The direct BS-UE channel is estimated with a conventional scheme and then, a user is selected randomly from the available number of users to estimate the BS-RIS and RIS-UE channels. Similarly, the BS-RIS-UE cascaded channel (consists of many coefficients) is estimated for the other users with a few coefficients. Thus, by ignoring receiver noise, the method[57] is useful to recover all the channel coefficients with the usage of $U+N+\max(U-1, \lceil(U-1)N/M\rceil)$ pilot symbols.

In addition to utilizing the channel property to guide the channel estimation, some hardware structural changes have also been proposed for RIS to benefit channel estimation. Different from most of the research considering that all elements of RIS are passive, Taha et al. added a few active elements in the designing of the RIS hardware[59]. The active elements are connected to the baseband signal processors to receive pilots. Based on the active RIS elements, two renowned strategies such as compressive sensing and DL are developed to jointly optimize the channel estimation and precoding at RIS. Since these active RIS elements are connected with the baseband signal processing, they introduce additional noise into the wireless system[61]. The pilot overhead for multi-user RIS-aided communication can be reduced by grouping the number of RIS elements into multiple large-size groups, where each large element consists of several adjacent RIS elements. Since pilot overhead is proportional to the number of RIS elements, hence by grouping RIS elements the overhead reduces to the number of groups. However, this strategy can cause some errors in computing the accurate phase shifts of RIS elements, since we compute the RIS phase shifters in groups rather than individual elements. Therefore, Ref. [61] proposed a solution that created a trade-off between pilot overhead and accuracy. The authors considered a quasi-static block fading channel to analyze the performance of RIS-aided wireless communications. Furthermore, to set the RIS groups' state, the authors in Ref. [61] employed the "ON/OFF" technique similar to that in Ref. [53]. Then, after receiving the pilot information of all users in a coherence interval, the least square (LS) estimation algorithm was employed to estimate the indirect BS-RIS-UE cascaded channel and direct BS-UE channel. On the other hand, Ref. [64] realized that by keeping the group at "ON" mode, error in computing the accurate reflection coefficients for RIS can be minimized. However, these solutions still require a large training overhead, which is unaffordable to realize practically in most of the cases. Thus, Ref. [48] proposed a semi-blind channel estimation solution to minimize the overall training overhead of the RIS-aided system. Particularly, they formulated semi-blind problem as a trilinear estimation problem where a received signal is a combination of user-RIS and RIS-BS channel matrix, respectively. To solve the trilinear problem, a low-complexity Bayesian MMSE iterative solution was adopted. Nevertheless, algorithm to converge fast, all the applicable matrices should be sparse. The channel estimation problem for RIS-aided mmWave was studied in Ref. [65] where the authors exploited the sparsity property of mmWave to obtain the relevant channels. Additionally, the sparsity can also be achieved by exploiting a sparse RIS phase shifting matrix. However, these strategies may not achieve the optimal solution of channel estimation and these solutions may be applicable for some specific scenarios. Therefore, to minimize these issues, a unitary approximate message passing (UAMP) algorithm was proposed in Ref. [66] to deal with linear reverse problem at a low-complexity. In Ref. [66], the authors first proposed a new signal model, and then transformed it to a structured signal recovery problem to be solved with UAMP algorithm.

It is worth mentioning here that by adopting the channel reciprocity technique in the time-division

duplexing (TDD), the CSI of the downlink path can easily be acquired if the CSI at the uplink is known at the transmitter in the RIS-aided communication. In contrast to this, the lack of channel reciprocity mechanism in frequency division duplexing (FDD) makes it less intelligent, as TDD is, to acquire the downlink CSI according to the uplink CSI. Hence, designing the codebook becomes essential for the RIS-aided system to provide the feedback of the downlink CSI to the BS. However, the only problem that arises with the designing of the codebook is that its size grows exponentially with the increase of the number of BS antennas and RIS elements, which is intolerable in practical RIS-aided wireless communication systems. To solve this problem, by utilizing the sparsity of angular channel, a dimension reduced channel feedback scheme with a small codebook size and a low overhead rate was proposed in Ref. [67].

6.2 Channel estimation for multi-antenna users

In MIMO systems, the user is equipped with multiple antennas, and thus the signal model for the multiple antenna user can be expressed as in Eq. (37). The channel estimation schemes that are proposed for single-antenna users cannot be applied directly to multi-antenna users, especially for the BS-RIS-UE cascaded channel estimation. Although the multi-antenna user can be equivalent to several single-antenna users, however, applying a single-user channel estimation strategy to the multi-antenna user can bring unbearable pilot overhead and complexity. Thus, the channel estimation schemes need to be revised, therefore, different from the conventional schemes, the authors in Refs. [68−74] have studied channel estimation schemes for multi-antenna users. In Ref. [68], the researchers proposed a two-stage algorithm to compute the CSI of the downlink block-fading channel. In the first stage, a sparse matrix factorization method is used to obtain the RIS-UE channel information. Then, in the second stage, the BS-RIS channel is computed using the matrix completion method. Similarly, Mirza and Ali[69] proposed a two-stage algorithm in which the direct BS-UE channel is estimated by utilizing the conventional TDD based channel estimation technique, keeping the elements of RIS to "OFF" state. Then, the bilinear adaptive vector approximate message passing (BAdVAMP) algorithm was proposed to estimate the BS-RIS-UE cascaded channel by modeling the estimation problem as an ill-conditioned dictionary learning problem. In Ref. [71], two-channel estimation schemes were proposed, where the time-domain pattern of pilots and RIS phase shifts are aligned such that the RIS phase shift vector is constant during the T time slots in the k-th block and varies from block to block. Moreover, these pilot signal vectors are repeated over the K blocks. It is worth noting that the received signal follows a parallel factor (PARAFAC) tensor model with canonical polyadic decomposition[75], which can be exploited to estimate the involved communication channels in closed-form or iteratively. Based on this property, the least-squares Khatri-Rao factorization (LSKRF) algorithm and bilinear alternating least squares (BALS) are designed to solve the problem.

In Ref. [72], channel estimation was discussed in the terahertz (THz) communication scenario, where a cooperative channel estimation procedure for RIS-aided wireless communication system via beam training was proposed based on hybrid beamforming (HB) architectures. They had adopted 3-tree hierarchical search schemes to realize the beam training with low search complexity.

7 Combination with other technologies

Since RIS has gained huge attention by the researchers in these years, RIS has been integrated with various technologies, which are discussed as follows.

(1) Simultaneous wireless information and power transfer (SWIPT): SWIPT is a new paradigm shift in wireless communications in which energy is transferred to the destination device through intermediate wireless devices. These wireless devices operate by supplying limited energy source/s, called batteries. These batteries have a limited lifetime, therefore the lifetime can be maximized either by recharging the wireless device or replacing the battery source with the new one. Thus, in both of the scenarios, the operational cost of the system can be too high,

which hinders its application practically. Hence, SWIPT has emerged as a cost-effective technique to maximize the lifetime of wireless devices[74]. It is a promising solution that ensures simultaneous delivery of data and energy for numerous low-power devices in future wireless powered Internet-of-Things (IoT)[76]. However, due to high path loss, the transmission efficiency and range of SWIPT are expected to be unsatisfactory for long-distance transmission[77].

Fortunately, RIS can be exploited to mitigate the transmission loss problem caused by long-distance communication. Integrating RIS with the SWIPT technology brings two major benefits, it improves the transmission efficiency of low-powered devices and extends the transmission range of wireless devices[2]. Since the access point (AP) in the SWIPT system aims at sending both the information and energy signals to information decoding (ID) or energy harvesting (EH) IoT devices, thus it is important to design a joint scheme for the transmission and reflection in an RIS-aided SWIPT system. Therefore, the most of existing works focused on jointly optimizing the transmit precoding matrix at the AP and the passive phase shifting matrix at the RIS with different optimization targets.

For example, the authors in Ref. [78] jointly optimized the transmit precoding matrix of the BS and the passive beamforming matrix at the RIS for RIS-aided SWIPT MIMO system to maximize the WSR of the wireless system, while satisfying the EH requirement of the energy signal receivers. To this end, an alternating iterative algorithm is developed to optimize the transmit precoding matrix at the BS and the passive beamforming matrix at the RIS. At a given transmit precoding matrix, the authors formulated the phase shift optimization problem as a nonconvex quadratically constrained quadratic programming (QCQP) and solved it with the developed iterative algorithm. Their simulation results demonstrated that aiding RIS to the SWIPT system can further improve the spectral efficiency of the wireless system. Additionally, the researchers in Ref. [79] proposed a joint precoding scheme for an RIS-aided SWIPT system to enhance the wireless power transfer (WPT) efficiency of the wireless system. The proposed model shows a significant performance gain over the SWIPT system that does not incorporate an RIS. On the other hand, a novel channel estimation protocol and the corresponding precoding design were studied in Ref. [80].

(2) Non-orthogonal multiple access (NOMA): With the exponential growth of mobile devices that support high data rate IoT applications, future wireless communications need to support the massive connectivity of users. NOMA is a promising technique that can accommodate a large number of users via non-orthogonal resource allocation. Specifically, NOMA, which allows multiple users to share the same time and frequency resources in the same spatial layer via power domain multiplexing and code domain multiplexing, realizes high spectrum efficiency and low latency[81]. This enables NOMA to be considered as a favorable technology for future wireless communication systems[82]. In the NOMA system, the directions of users' channel vectors are the same, and manipulating the directions of users' channel vectors is an intractable problem because the users' channels are determined by the fixed propagation environments.

As RIS is capable of intelligently adjusting the phase shifts of its passive elements to manipulate the EM waves, it is envisioned as an innovative technology to facilitate the problem of NOMA. Moreover, RIS may also facilitate in contructing new channel links to enhance the performance of the NOMA system. The research on the RIS-aided NOMA is still in the initial stage, some of the latest works have been summarized as follows: In Ref. [83], the authors proposed a framework of the low-cost RIS-aided NOMA system for downlink transmission to ensure that additional cell-edge users could be served while their original channels are not aligned. Specifically, the cell-edge users' effective channel vectors are aligned with the prearranged spatial directions by tuning the reflecting coefficients of phase shifts and amplitude reflection coefficients of RIS.

Yang et. al[84] formulated the problem to ensure the minimum SINR requirement for a user in the RIS-aided downlink NOMA system. They proposed a joint

optimizing algorithm to optimize the power allocation at the BS and the phase shifts at RIS to maximize the minimum SINR of a user. On the other hand, besides the designing of joint optimization of power allocation at BS and phase shifts of RIS, Ref. [85] also described where to deploy the RIS to realize high energy efficiency of RIS-aided NOMA system. The deep reinforcement learning (DRL) approach was developed to solve the proposed problem, which proved to be a better tradeoff between prediction accuracy and computational complexity.

(3) Physical layer security: Unlike the network layer security, physical layer security is acknowledged as a key technique that provides secure wireless transmissions between source and destination, since wireless transmissions are inherently vulnerable to security[86]. However, one of the critical challenges in physical layer security is to ensure a high secrecy rate (SR) for the legitimate information receiver (IR), while avoiding eavesdroppers (Eve) getting access to the wireless transmission[87]. The SR can further be improved by improving the channel quality link from BS to IR and discrediting the channel link from BS to Eve. This can be done by transmitting the artificial noise from the BS to contaminate Eve's signal. Though transmitting artificial noise from the BS may increase the SR for IR, it also dissipates an extra amount of power which creates a bottleneck to use it in an energy-limited system. Thus, RIS can be deployed to combat the power consumption problem but also strengthen the security of the wireless system.

Considering RIS to enhance the security of the physical layer could be a critical problem in future wireless communications, and hence attracting much attention in industry and academia. In Ref. [88], the authors used RIS in the system model to control a wireless environment to ensure the security of a wireless network. To this end, they jointly optimized the beamforming of the BS and reflecting coefficient at the RIS, eliminating the interference of multiple Eves, to maximize the SR of the system. In Ref. [89], the authors considered the scenario in which the legitimate IRs lack the direct LoS communication links with the BS due to the presence of multiple potential Eves. To counter the strong effect of eavesdropping channels, a dummy signal called artificial noise is transmitted from the BS deliberately. Then, the beamforming matrix and covariance matrix of artificial noise at the BS are jointly optimized with the phase shifters of RIS to maximize the sum-rate of the system under the consideration of leakage information to the Eves. To solve the problem, the authors presented a novel DRL-based beamforming approach that ensures high secrecy among the IRs. Additionally, the same problem has also been considered in Refs. [90−92] for RIS-aided wireless systems. Therefore, benefitting from the RIS, we can conclude that RIS could be an effective technology in future 6G communications to ensure high secrecy and privacy if we properly optimize RIS phase shifters.

(4) Unmanned aerial vehicle (UAV): UAV is deployed to the scenarios where the direct path between the BS (transmitter) and user (receiver) is unavailable. Thus, by introducing a UAV between the transmitter and receiver, an extra LoS link can be developed to extend the communication range between source and destination devices. UAV acts as a relay between the transmitter and receiver, and all the complex signal processing operations are performed while retransmitting the signals to the destinations. As the radio conditions are random and vary continually with time, it is very likely that the LoS link between the transmitter and receiver may be blocked which degrades the performance of the communication. Thus, to improve the radio link between the transmitter and receiver, an RIS can be deployed to control the harsh propagation environment[93, 94]. Thus, by installing the RIS phased antenna array in the vicinity of UAV, we can save the hardware cost and power consumption while achieving the same target, i.e., energy efficiency/spectral efficiency. Taking advantage of the RIS, in Ref. [93], the authors proposed RIS-aided UAV communication framework to maximize the spectral efficiency of wireless system. They jointly investigated the UAV trajectory and passive beamforming for an RIS to maximize the spectral efficiency. To solve the problem, a successive convex approximation (SCA) technique was proposed. Similarly, aiding RIS in UAV

communication model, the authors in Ref. [94] minimized the outage probability while ensuring a good SINR of the considered system.

(5) Backscatter communication: Another possible scenario where RIS can be a good choice is RIS-aided backscatter communication (Backcom). Backcom is a valuable energy-efficient technology to be deployed in IoT networks. Despite its enormous advantage of being a simple technology, its short-range operation is still creating trouble for the researchers to be deployed at a large scale. Thus to extend the communication between the pair of devices to continue communication, RIS could be a favorable choice. The second problem that lies in Backcom is self-interference. Hence, taking advantage of the RIS phase shifting matrix to direct the incoming beam to a specific direction may reduce the inter-user interfernace which improves the detection performance of the Backcom system[95, 96].

8 Implementation of RIS

Implementing an RIS practically is one of the most challenging tasks to do. Indeed, one cannot enjoy the theoretical benefits of the RIS-aided communication system without good RIS hardware. The hardware realization of an RIS has become one of the main interests of researchers in the field of electromagnetics and communication engineering. Since RIS is an emerging technology, the implementation of RIS is evolving rapidly in recent years. Several different types of RIS prototypes have been proposed. In this section, we will categorize the implementation of RIS in two parts: RIS elements and RIS arrays. To introduce the implementation of RIS elements, we will present several kinds of RIS elements made of different materials. Their characteristics will also be compared. Then, we will introduce the design of RIS arrays by discussing the geometric structure, the architecture, and the degree of freedom.

8.1 RIS elements

RIS is proposed to adjust the phase shift of the incident EM wave to change its propagation. Different from natural EM materials which control the behavior of the incident EM waves based on classical refraction, reflection, absorption, and diffraction, RIS benefits from replacing traditional materials with artificial tunable metasurfaces/metamaterials. Metasurfaces are artificial planar metamaterials which consist of periodic subwavelength metal/dielectric structures[97, 98]. Due to the subwavelength structures and integrated functional materials, metasurfaces exhibit some unusual properties such as anomalous reflection and refraction[98]. Based on tuning mechanisms, there are two major categories of metasurfaces: stimuli-responsive materials and voltage-driven elements. The former depends on the physical properties of materials and the latter relies on their circuit structure.

The physical properties of stimuli-responsive materials can be changed by applying different external stimuli to the metasurface, thus resulting in desired EM wave response. There are different types of materials, e.g., electric tuning materials, magnetic tuning materials, and light tuning materials. Liquid crystal and graphene are two typical electric-sensitive materials to implement controllable reflection. In Ref. [99], authors utilized the characteristic of liquid crystal that liquid crystal has a voltage-controlled dielectric constant. Based on this characteristic, a reconfigurable metasurface reflector is designed, whose phase shifts are adjusted by the direct current voltage to achieve controllable reflection. Moreover, liquid crystal may also be used in tunable absorbers, as shown in Fig. 6a[100]. To realize tunable reflection, the Fermi

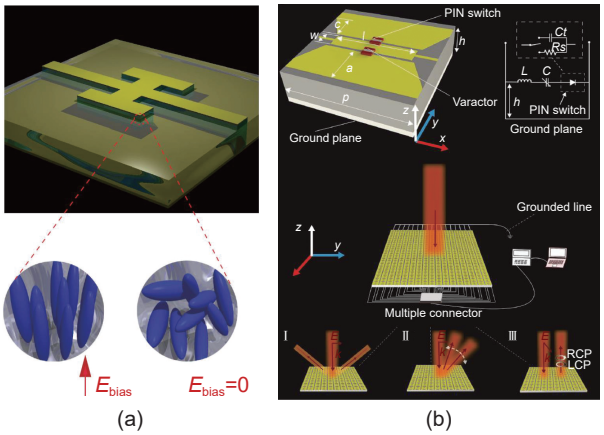

Fig. 6 Metasurfaces based on two typical tuning mechanisms. (a) Stimuli-responsive: liquid crystal based tunable metamaterial absorber[100]. (b) Voltage-driven: varactor-based reconfigurable metasurface with multiple EM functionalities[101].

level of graphene has to change by subjecting it to an electrostatic field. However, the impedance incurred by graphene is usually very large compared to other surrounding materials, resulting in low tunability. This impedance can be reduced by combining graphene with metallic blocks, which effectively improves the phase adjustment[102, 103]. In Refs. [104, 105], several kinds of metamaterials sensitive to the magnetic field, light, and temperature were also introduced, for the reference of interested readers.

However, to control the properties of each reflecting element independently, we need to apply delicate external excitation to each element. The stimuli-responsive materials used to design reflecting elements of the metasurface are not practical due to the large unit size, especially in the high-frequency bands. As a result, the voltage-driven mechanism is a more viable way. Therefore, Ref. [106] proposed a switching diodes based scheme to control the switching of total absorption and total reflection in which the state of diodes is controlled by excitation voltage. A similar concept of the smart wall was proposed in Ref. [7] to realize transmission or reflection by applying a bias voltage to positive intrinsic-negative (PIN) diodes embedded in the metasurfaces. In Refs. [101, 107], varactor diodes were controlled by excitation electric field, which further improves frequency tunability and high-precision continuous phase tunability, as shown in Fig. 6b. In Ref. [108], the elements of the reconfigurable reflection array were designed based on a variable-volume resonator, where the resonant frequency and phase shifts are controlled by the bias voltage applied on the resonator.

8.2 RIS arrays

The existing works on RIS arrays can be categorized into three aspects: geometry, architecture, and the control degree of freedom. The geometry of an RIS array aims at encapsulating a large number of reflecting antenna elements in a given surface area of m^2. For the same, Ref. [109] designed the RIS array as a hexagonal lattice structure. The hexagonal structure minimizes the array area under the constraint of the independent signal dimension. In particular, the hexagonal lattice saves up to 23% of the surface area compared to the rectangular lattice considering the same number of independent signal dimensions. Another constraint of the dense lattice was introduced in Ref. [110]. Given the independent signal dimensions provided by the unit area of the sampling lattice, the minimum number of antennas are deployed on each unit.

The RIS is available in two different architectures, including passive architecture and sparse sensor based architecture. The tunable RIS reflecting elements are considered passive that bring a lot of benefits such as low hardware cost and power consumption. In passive RIS architecture, each element is involved to acquire the channel state information (CSI)[58]. Nevertheless, a large number of antennas in RIS provide excessive training overhead to acquire exact CSI. On the other hand, a novel RIS architecture with sparse active sensors was propped in Ref. [59]. This sparse active sensors based architecture is composed of passive reflecting elements, which are connected to the baseband controller. The active sensors are responsible to compute CSI which we already have discussed in Section 6. The RIS elements furnished over any of the above-mentioned architectures should be tuned and optimized as per the radio environment. These reflecting elements are designed based on either stimuli-material or voltage-driven to realize high hardware gain. To this end, the researchers in Ref. [101] presented the column-controlled varactors method to steer the beam in a particular direction. This is done by varying the voltage for each reflecting element in one direction while keeping it unchanged for the other. Reference [111] further realized more flexible phase shifts with good precision by controlling each PIN diode independently, thus achieving a two-dimension controlled degree of freedom.

RIS arrays are available in various forms for indoor and outdoor communication. For example, an indoor metasurface based on a software programmable model was proposed in Ref. [112]. This metasurface is composed of several metasurfaces that are stacked one after the other are called hypersurface, controlled by a central environment configuration server. The central

server controls the reflecting elements through the software-defined interface to adjust EM response, beamforming, reflection, absorption, and polarization.

9 Prototypes of RIS-aided communication systems

In this section, we will introduce prototypes of RIS-aided communication systems that have been studied in the literature, including the RF chain-free transmitter, the space-down-conversion receiver, the power-efficient communication system, and the transparent dynamic metasurface. Based on the performance, these cases have verified the possibility that the RIS can be deployed in future wireless networks.

9.1 RF chain-free transmitter

(1) Principle: The massive MIMO is a key technology to meet the growing demand of high data traffic. In a conventional massive MIMO transmitter, antenna elements are connected to RF chains, where each RF chain is equipped with a power amplifier, two mixers and several filters[113], which increases the hardware cost. To reduce the hardware cost, the directional modulation (DM) is proposed to realize the massive MIMO transmitter without RF chains. DM utilizes far-filed characters of RIS to modulate the signals rather than performing baseband modulation[114, 115]. Specifically, a single-tone generator transmits an EM wave, the EM wave is then reflected by a series of reflection coefficients of RIS, where the reflection coefficients have four alternative values, which is consistent with the four constellation points of quadrature phase-shift keying (QPSK). To control the RIS, the data are converted to a sequence of baseband signals and then passed to a digital-to-analog converter (DAC) to produce the bias voltage to control the RIS elements. Finally, the users hear the modulated signal after the controlled reflection of the RIS.

(2) Prototype: Based on this new architecture, researchers built some prototypes, where the transmitter consists of baseband signal processors, DACs, single-tone generators, and an RIS[113, 116]. A QPSK DM prototype was develpoed in Ref. [116]. The prototype is designed at the 7 GHz band, where the number of RIS elements is set to 32. However, due to the limited signal processing capability of the computer, a bit rate of only 200 kbps is achieved. To test the performance of the proposed prototype model, the researchers also compared the bit error rate (BER) performance of the proposed RF chain-free transmitter with the conventional RF chain based transmitter at different receiver directions. The results show that the DM transmitter can generate a narrower beam than the traditional transmitter. However, the model required higher transmit power to hold the same BER as that of the traditional RF chain based transmitter. Another field-programmable gate array (FPGA) based signal processor prototype was implemented in Ref. [113]. This prototype also uses QPSK modulation but operates at 4 GHz, and the number of RIS elements is 128. Thanks to the powerful processor, the data rate increases to 2.048 Mbps. Finally, this new architecture also compared the BER performance with the conventional transceiver, and the results show that the transmit power should be increased by 5 dB for the new transmitter to maintain the same BER as the traditional one, due to the spectral leakage of the reflected EM wave.

(3) Discussions: The RF chain-free idea is proposed to get the conventional architecture “free” of the energy-intensive RF chain. There is no doubt that this is a novel, brave, and interesting idea. More importantly, this idea has been successfully tested in a real-time application. However, there are still some unsolved issues that need to address in further studies as stated in Ref. [113]. For example, the single tone wave cannot support wideband communications. It is practically important to study how to realize the architecture in a multi-carrier communication system in the future.

9.2 Space-down-conversion receiver

(1) Principle: In a conventional receiver, the antenna receives the signal and passes it to the other modules such as low-noise amplifiers, filters, mixers, analog-to-digital converter (ADC), and baseband signal processor. These modules have a high cost and consume a significant amount of energy while

operating. To address this problem, the authors in Ref. [117] gave an idea that the received signal can be processed and converted into a low-frequency signal in space using an RIS rather than processing it in a traditional down-converter module. Specifically, when the EM wave is transmitted from the transmitter, it is directed towards the RIS, where the controller of the RIS changes the phase shift of the elements linearly with time and redirects the incident EM wave towards the receiver. During this process, the incident wave is rescaled from high to low frequency. In other words, we say that the RIS acts as a down-converter, and the procedure of down-conversion is done in space rather than by a circuit.

(2) Prototype: A 5 MHz down-conversion receiver was developed by Ref. [117]. In this prototype, the incident wave is down-converted and reflected by an RIS which is then processed by other low-frequency modules. This system works at the 4.25 GHz band, where the 256 number of reflecting elements is considered at the RIS. The authors used 16-quadrature amplitude modulation (16QAM) constellation in their experiments and the experimental results show that the shape of the received 16QAM constellation is clear and distinguishable, which shows the significance of the RIS in future systems.

(3) Discussions: Different from the RF chain-free transmitter, this space-down-conversion model is applied in the receiver. In this prototype, if a high-speed phase shifting metasurface such as RIS is available, the rest of the modules at the receiver act as a baseband. This prototype can be envisioned as a new paradigm for the user equipment. However, since the current phase shifting speed in the RIS is limited, this prototype achieves only 5 MHz down-conversion. Therefore, the available designs are not enough to achieve the goals of future communication systems, and thus the study is continued to design efficient hardware systems that could work even at higher frequency bands.

9.3 Power-efficient communication system

(1) Principle: The RF chain-free transmitter and the space-down-conversion receiver are proposed to reduce the hardware cost and power consumption of a conventional massive MIMO system. To this end, a novel structure of an RIS with 2-bit phase-shift reflective elements was proposed in Ref. [118]. This new structure only needs 5 PIN diodes and 2 control bits to manipulate the phase shift of a reflective element. Thanks to the spatial feeding mechanism of an RIS, the excessive power loss caused by the bulky feeding networks of phase array is avoided, and thus the power consumption and hardware cost of RIS are reduced.

(2) Prototype: In Ref. [118], the researchers presented prototypes of an RIS for both transmitter and receiver to realize power-efficient communication. At the transmitter side, they considered only one RF chain that generates EM wave to be directed at the RIS which has 2-bit phase shift elements. The RIS is used to realize passive beamforming by controlling the phases on all the elements. A high-level modulation scheme such as 16QAM, 64QAM is used to modulate the bit sequence on the RF chain. The transmitter is composed of the same size 256 antenna elements, as shown in Fig. 7a , that operate at a 2.3 GHz band. The researchers claimed that it can achieve 21.7 dBi beamforming gain in comparison to a single-antenna transmitter. Furthermore, the researchers also investigated the power consumption analysis of the

(a) Transmitter

(b) Receiver

Fig. 7 RIS-based power-efficient communication system operating at 2.3 GHz[118].

proposed RIS model in their work, which indicates that RIS requires 153 W, while the conventional phase array requires 370 W to achieve the same effective isotropic radiated power (EIRP). The second prototype is proposed for the receiver, which is shown in Fig. 7b, that achieves a 19.1 dBi beamforming gain compared to the single antenna receiver. Both the models (transmitter and receiver) have passed over the air (OTA) test, which means these models have a strong potential to transport high-resolution video frames in no time.

(3) Discussions: The results provided by the authors in Ref. [118] indicate that the 2-bit phase quantization of the reflective elements at an RIS is practically enough to realize high beamforming gain to construct the massive MIMO transmitter. Furthermore, the prototype has also been tested at the 28.5 GHz band, where the OTA test shows a strong potential to exploit RIS in future 6G communications. In the future, the authors will extend this work to a multi-user communication system and a higher frequency band like sub-THz.

9.4 Transparent dynamic metasurface

(1) Principle: Different from the prototypes discussed in the previous subsections, the transparent dynamic metasurface is developed by two companies, i.e., DOCOMO and AGC recently. The metasurface is composed of a common glass[119], where many sub-wavelength unit cells called elements are spread over the entire surface covered by a glass substrate. The scientists claimed that the newly designed RIS model can be naturally deployed in our living environment, as it does not affect humans. Thanks to the transparent character of the glass surface, we can decorate the window of a house, a window of a car, even the furniture in the house with it.

(2) Prototype: Transparent dynamic metasurface was presented in DOCOMO Open House in Tokyo on January 23, 2020. The researchers have claimed that the glass-based RIS model can achieve 1.3 Gbps in downlink mode within a 100 m range[120] while working at 28 GHz with 400 MHz bandwidth.

(3) Discussions: Compared with other RIS prototypes, this new metasurface could be a good choice in future wireless communications as it can be installed anywhere in the surrounding environment without affecting humans. However, it is still an open research direction. The researchers can work in this research direction to further improve its performance and design to realize the objective of future 6G and beyond communications.

10 Possible applications of RIS

In this section, we discuss four different scenarios where RIS could bring a significant improvement to overall systems' performance. In these scenarios, the RIS is exploited to ensure, high data rate, energy efficiency, a good quality signal, and less interference among the users.

10.1 NLoS communications

The first scenario where RIS can play an important role is non-line-of-sight (NLoS) communications, for instance, there are some cases where the direct communication link or the quality of the received signal is compromised due to the scattering and high path loss problems. This is a common problem in an urban environment where the high-rise buildings or trees may weaken the quality of the signal received at the destination/receiver device. Here, RIS can be an effective technology to create a line-of-sight (LoS) link between transmitter and receiver, which improves the performance of the wireless system. The NLoS channel is a more common condition in mobile networks than LoS channel[121]. For example, mmWave massive MIMO systems mainly take advantage of multiple NLoS links to transmit information, which means that the stronger NLoS link integrated with RISs may lead to better channel quality, which ensures a high system capacity.

10.2 Energy-efficient MIMO transmitters

As we know that the conventional massive MIMO systems can realize a high spectral efficiency[122]. In massive MIMO systems, each antenna element is linked with the energy-intensive RF chain, which intensifies the energy consumption of the entire

system. For example, a 64 antenna elements based massive MIMO system requires 16 W to operate[123, 124]. The power consumption further increases with the increase in the number of antennas, and this huge power consumption creates a bottleneck in realizing the massive MIMO technique practically in future wireless systems. Hence, RIS can be the best substitute for the conventional high-cost energy-intensive massive MIMO system. Since RIS contains low-cost passive reflecting elements, it does not require any dedicated RF chain and any active power sources for transmission, which ensures high energy efficiency[125].

10.3 Service for cell-edge users and interference suppression

In mobile communications, the users are mobile and are randomly located in the spherical region. The users located at the edge of a cell are called cell-edge users. These users always experience a low quality SINR from BS, since the distance between BS and the cell-edge user/s is too large that deteriorates the link quality due to path loss problem. Moreover, the cell-edge users also experience a large interference from the neighboring cells, hence degrading the performance of a communication system. To avoid this situation, a conventional solution is to deploy a large number of BSs to extend the coverage. Although this method provides better coverage, however, this method increases high hardware cost and also brings more EM pollution in the space. Thus, RIS can be a good choice in extending the coverage of a cell at low transmit power.

10.4 Massive machine-type communications (mMTC)

The mMTC will support many applications such as smart cities, environmental monitoring, intelligent agriculture, forest fire-prevention, and other sensing and data acquisition tasks in the next-generation networks. In mMTC, the data collected by the devices are transmitted to the core server for further processing. Since the devices are typically battery-powered and require long-term autonomy, the communication process should be power-efficient. However, the power constraint limits the capacity of data upload. This problem can be avoided if the devices are embedded with an RIS. Instead of transmitting a new signal, the RIS collects the signal transmitted by a BS and then the data are embedded into the reflected signals[126]. Since the passive elements of RIS consume limited power, the IoT devices can transmit their data at a low power, which increases the lifetime of the devices.

11 Challenges and future research directions

We believe that RIS technology will become an integral part of future wireless communications. RIS has gained enormous attention in recent years by researchers of different communities, and therefore, several research papers have been published in this regard. However, there are still some important open problems that need to be solved. In this section, we will introduce the major challenges and future research directions for RIS-aided communication.

11.1 Channel modeling

Up to now, the channel model of an RIS is still controversial, especially for the near-field scenario. The reason is that some analytical results in existing papers have not been verified by experiments yet, therefore we are not sure, among the existing available models, which model will be the best choice for practical scenarios, such as the clock drift model[127], statistical CSI model[128], and Nakagami fading model[129]. In the future, it will add more value to the RIS research field to do more experiments and compare the existing models. The existing path loss model for near-field RIS is obtained by considering the RIS as a whole surface. However, in the future, it will be an interesting topic to study the element-wise channel gains for the near-field RIS.

11.2 Performance analysis

The performance analysis for RIS-aided systems is still in its early stage, and it is important to study these issues to answer a few important questions. For instance, the fundamental differences between RIS and relay have not been clearly stated out, which may be an important research direction in the future. Moreover, a

performance comparison between RIS and relay in many of the existing works considers single-antenna relay for simplicity[27, 29], while the more general case with multi-antenna relay has not been fully investigated. Besides, although the received power with discrete phase shifts deployed at RIS has been analyzed[25], the loss of system capacity due to the non-ideal phase shifts remains an open problem. In addition, to characterize the ultimate performance of RIS from the physical layer, the new concept of electromagnetic information theory can be introduced to analyze the RIS-aided communication systems through using random field principle[130, 131].

11.3 RIS implementations

To gain intelligent control over the environment, external excitation should be applied to each antenna element. With the limited size of the excitation source, integrating different metamaterials and excitation devices, and integrating low-cost controller networks in metasurfaces are still challenging problems[97]. To further study the implementation of RIS, it is necessary to deeply explore the properties of metamaterials from the perspective of physics. In addition, software-defined metasurface is a potential direction, including the joint design of geometric structure and EM response mechanism. Whether it is valuable to adopt active elements in RIS and its influence are also valuable issues to be explored[132−134].

11.4 Precoding for RIS

There have been many research papers on precoding design for RIS-aided communication[135]. Most of the proposed algorithms have considred the case in which CSI is known at the transmitter, which is not true. There are some works in which authors have considered the imperfect CSI, however, the cost of the precoding algorithms is too high which is hard to realize in practical systems. Hence we need to investigate more practical precoding algorithms that could easily be realizable in practical systems. Moreover, the precoding with non-ideal CSI can be another problem. For example, when we have only a part of the CSI rather than full CSI, or when we have only biased estimated CSI rather than accurate CSI, we need a new precoding scheme to realize robust transmission[136, 137].

11.5 CSI acquisition

The CSI acquisition for RIS-aided communication is a difficult problem, which should be reinvestigated due to the following reasons. Firstly, the existence of RIS makes the pilot overhead for downlink channel estimation unaffordable. Since the overhead is the bottleneck in acquiring CSI, it is important to find new methods to reduce the pilot overhead. Secondly, the pilot transmission suffers from severe path loss for the RIS-aided reflection link. The received power of pilots can be negligible, which will reduce the channel estimation accuracy. Thus, another valuable research direction is to design a set of beam patterns for pilot transmission to maximize the received SNR.

12 Conclusion

In this article, we have summarized the potential applications for RIS technology in future 6G communications. At first, we briefly introduced the key idea of RIS, and then discussed the application scenarios of RIS such as system and path loss model, asymptotic analysis, precoding, CSI acquisition, prototypes, and its combination with other technologies. To have a deep insight on the individual topic, we have given some examples to show the importance of RIS technology in future wireless systems. Finally, we highlighted the key challenges and provided some opportunities to guide the readers for future research trends in the field of RIS-aided communication. It is expected that RIS will be an important research topic in the years to come as it has the potential to create a new horizon in the field of wireless communications.

References

[1] M. D. Renzo, A. Zappone, M. Debbah, M. -S. Alouini, C. Yuen, J. D. Rosny, and S. Tretyakov, Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead, *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2450–2525, 2020.

[2] Q. Wu and R. Zhang, Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless

network, *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, 2020.

[3] N. Engheta and R. W. Ziolkowski, *Electromagnetic Metamaterials: Physics and Engineering Explorations*. Piscataway, NJ, USA: Wiley-IEEE Press, 2006.

[4] T. J. Cui, D. R. Smith, and R. Liu, *Metamaterials: Theory, Design, and Applications.* New York, NY, USA: Springer Science & Business Media, 2009.

[5] D. R. Smith, W. J. Padilla, D. C. Vier, S. C. Nemat-Nasser, and S. Schultz, Composite medium with simultaneously negative permeability and permittivity, *Phys. Rev. Lett.*, vol. 84, no. 18, pp. 4184–4187, 2000.

[6] D. R. Smith, J. B. Pendry, and M. C. K. Wiltshire, Metamaterials and negative refractive index, *Science*, vol. 305, no. 5685, pp. 788–792, 2004.

[7] L. Subrt and P. Pechac, Controlling propagation environments using intelligent walls, in *Proc. 2012 6th European Conf. Antennas Propag. (EUCAP)*, Prague, Czech Republic, 2012, pp. 1–5.

[8] N. Kaina, M. Dupré, G. Lerosey, and M. Fink, Shaping complex microwave fields in reverberating media with binary tunable metasurfaces, *Scientific Reports*, vol. 4, no. 1, p. 6693, 2014.

[9] T. J. Cui, M. Qi, X. Wan, J. Zhao, and Q. Cheng, Coding metamaterials, digital metamaterials and programmable metamaterials, *Light: Science & Applications*, vol. 3, p. e218, 2014.

[10] Z. Zhang and L. Dai, Continuous-aperture MIMO for electromagnetic information theory, arXiv preprint arXiv: 2111.08630, 2021.

[11] H. Guo, Y. -C. Liang, J. Chen, and E. G. Larsson, Weighted sum-rate optimization for intelligent reflecting surface enhanced wireless networks, arXiv preprint arXiv: 1905.07920, 2019.

[12] S. Abeywickrama, R. Zhang, Q. Wu, and C. Yuen, Intelligent reflecting surface: Practical phase shift model and beamforming optimization, *IEEE Trans. Commun.*, vol. 68, no. 9, pp. 5849–5863, 2020.

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

[14] Q. Wu and R. Zhang, Beamforming optimization for intelligent reflecting surface with discrete phase shifts, in *Proc. 2019 IEEE Int. Conf. Acoust. Speech and Signal Process. (ICASSP)*, Brighton, UK, 2019, pp. 7830–7833.

[15] P. Wang, J. Fang, and H. Li, Joint beamforming for intelligent reflecting surface-assisted millimeter wave communications, arXiv preprint arXiv: 1910.08541, 2019.

[16] Y. Cao, T. Lv, and W. Ni, Intelligent reflecting surface aided multi-user millimeter-wave communications for coverage enhancement, in *Proc. 2020 IEEE 31st Int. Symp. Personal, Indoor and Mobile Radio Commun.*, London, UK, 2020, pp. 1–6.

[17] X. Gao, L. Dai, S. Han, C. -L. I, and X. Wang, Reliable beamspace channel estimation for millimeter-wave massive MIMO systems with lens antenna array, *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 6010–6021, 2017.

[18] S. Han, C. -L. I, Z. Xu, and C. Rowell, Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G, *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 186–194, 2015.

[19] O. Özdoğan, E. Björnson, and E. G. Larsson, Intelligent reflecting surfaces: Physics, propagation, and pathloss modeling, *IEEE Wireless Commun. Lett.*, vol. 9, no. 5, pp. 581–585, 2019.

[20] M. D. Renzo, F. H. Danufane, X. Xi, J. D. Rosny, and S. Tretyakov, Analytical modeling of the path-loss for reconfigurable intelligent surfaces-anomalous mirror or scatterer? in *Proc. 2020 IEEE 21st Int. WkSp. Sig. Process. Advances in Wireless Commun. (SPAWC)*, Atlanta, GA, USA, 2020, pp. 1–5.

[21] W. Tang, M. Z. Chen, X. Chen, J. Y. Dai, Y. Han, M. D. Renzo, Y. Zeng, S. Jin, Q. Cheng, and T. J. Cui, Wireless communications with reconfigurable intelligent surface: Path loss modeling and experimental measurement, *IEEE Trans. Wireless Commun.*, vol. 20, no. 1, pp. 421–439, 2020.

[22] S. W. Ellingson, Path loss in reconfigurable intelligent surface-enabled channels, in *Proc. 2021 IEEE 32nd Annual Int. Sym. on Personal, Indoor and Mobile Radio Commun. (PIMRC)*, Helsinki, Finland, 2021, pp. 829–835.

[23] D. Dardari, Communicating with large intelligent surfaces: Fundamental limits and models, *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2526–2537, 2020.

[24] E. Björnson and L. Sanguinetti, Power scaling laws and near-field behaviors of massive MIMO and intelligent reflecting surfaces, *IEEE Open J. Commun. Society*, vol. 1, pp. 1306–1324, 2020.

[25] Q. Wu and R. Zhang, Beamforming optimization for wireless network aided by intelligent reflecting surface with discrete phase shifts, *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1838–1851, 2020.

[26] P. Wang, J. Fang, X. Yuan, Z. Chen, H. Duan, and H. Li, Intelligent reflecting surface-assisted millimeter wave communications: Joint active and passive precoding design, *IEEE Trans. Veh. Tech.*, vol. 69, no. 12, pp. 14960–14973, 2020.

[27] E. Björnson, O. Özdogan, and E. G. Larsson, Intelligent

reflecting surface versus decode-and-forward: How large surfaces are needed to beat relaying, *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 244–248, 2020.

[28] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, Cooperative diversity in wireless networks: Efficient protocols and outage behavior, *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, 2004.

[29] J. Lyu and R. Zhang, Spatial throughput characterization for intelligent reflecting surface aided multiuser system, *IEEE Wireless Commun. Lett.*, vol. 9, no. 6, pp. 834–838, 2020.

[30] Q. Wu and R. Zhang, Intelligent reflecting surface enhanced wireless network: Joint active and passive beamforming design, in *Proc. 2018 IEEE Glob. Commun. Conf. (GLOBECOM)*, Abu Dhabi, United Arab Emirates, 2018, pp. 1–6.

[31] C. Huang, A. Zappone, M. Debbah, and C. Yuen, Achievable rate maximization by passive intelligent mirrors, in *Proc. 2018 IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Calgary, Canada, 2018, pp. 3714–3718.

[32] K. Liu, Z. Zhang, L. Dai, and L. Hanzo, Compact user-specific reconfigurable intelligent surfaces for uplink transmission, *IEEE Trans. Commun.*, vol. 70, no. 1, pp. 680–692, 2021.

[33] Y. Yang, S. Zhang, and R. Zhang, IRS-enhanced OFDM: Power allocation and passive array optimization, in *Proc. 2019 IEEE Global Commun. Conf. (GLOBECOM)*, Waikoloa, HI, USA, 2019, pp. 1–6.

[34] J. Ye, S. Guo, and M. -S. Alouini, Joint reflecting and precoding designs for SER minimization in reconfigurable intelligent surfaces assisted MIMO systems, *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5561–5574, 2020.

[35] M. Z. Siddiqi, T. Mir, M. Hao, and R. MacKenzie, Low-complexity joint active and passive beamforming for ris-aided mimo systems, *IET Elect. Lett.*, vol. 57, no. 9, pp. 384–386, 2021.

[36] M. Jung, W. Saad, M. Debbah, and C. S. Hong, On the optimality of reconfigurable intelligent surfaces (RISs): Passive beamforming, modulation, and resource allocation, *IEEE Trans. Wireless Commun.*, vol. 20, no. 7, pp. 4347–4363, 2021.

[37] Z. Zhou, N. Ge, W. Liu, and Z. Wang, RIS-aided offshore communications with adaptive beamforming and service time allocation, in *Proc. 2020 IEEE Int. Conf. Commun. (ICC)*, Dublin, Ireland, 2020, pp. 1–6.

[38] 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. S. Areas Commun.*, vol. 38, no. 8, pp. 1809–1822, 2020.

[39] H. Li, R. Liu, M. Liy, Q. Liu, and X. Li, IRS-enhanced wideband MU-MISO-OFDM communication systems, in *Proc. 2020 IEEE Wireless Commun. Net. Conf. (WCNC)*, Seoul, Republic of Korea, 2020, pp. 1–6.

[40] M. Z. Siddiqi, R. Mackenzie, M. Hao, and T. Mir, On energy efficiency of wideband RIS-aided cell-free network, *IEEE Access*, vol. 10, pp. 19742–19752, 2022.

[41] Z. Zhang and L. Dai, A joint precoding framework for wideband reconfigurable intelligent surface-aided cell-free network, *IEEE Trans. Signal Process.*, vol. 69, pp. 4085–4101, 2021.

[42] Z. Zhang and L. Dai, Capacity improvement in wideband reconfigurable intelligent surface-aided cell-free network, in *Proc. 2020 IEEE 21st Int. WkSp. Signal Process. Adv. Wireless Commun. (SPAWC)*, Atlanta, GA, USA, 2020, pp. 1–5.

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

[44] Q. -U. -A. Nadeem, A. Kammoun, A. Chaaban, M. Debbah, and M. -S. Alouini, Asymptotic max-min SINR analysis of reconfigurable intelligent surface assisted MISO communication, *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 7748–7764, 2020.

[45] X. Li, J. Fang, F. Gao, and H. Li, Joint active and passive beamforming for intelligent reflecting surface-assisted massive MIMO systems, arXiv preprint arXiv: 1912.00728, 2019.

[46] H. Xie, J. Xu, and Y. -F. Liu, Max-min fairness in IRS-aided multi-cell MISO systems with joint transmit and reflective beamforming, *IEEE Trans. Wireless Commun.*, vol. 20, no. 2, pp. 1379–1393, 2020.

[47] H. Shen, W. Xu, S. Gong, Z. He, and C. Zhao, Secrecy rate maximization for intelligent reflecting surface assisted multi-antenna communications, *IEEE Commun. Lett.*, vol. 23, no. 9, pp. 1488–1492, 2019.

[48] S. Hu, Z. Wei, Y. Cai, C. Liu, D. W. K. Ng, and J. Yuan, Robust and secure sum-rate maximization for multiuser MISO downlink systems with self-sustainable IRS, *IEEE Trans. Commun.*, vol. 69, no. 10, pp. 7032–7049, 2021.

[49] Y. Cao, T. Lv, Z. Lin, and W. Ni, Delay-constrained joint power control, user detection and passive beamforming in intelligent reflecting surface-assisted uplink mmWave system, *IEEE Trans. Cog. Commun. Net.*, vol. 7, no. 2, pp. 482–495, 2021.

[50] C. Hu, X. Wang, L. Dai, and J. Ma, Partially coherent

compressive phase retrieval for millimeter-wave massive MIMO channel estimation, *IEEE Trans. Signal Process.*, vol. 68, pp. 1673–1687, 2020.

[51] X. Gao, L. Dai, S. Zhou, A. M. Sayeed, and L. Hanzo, Wideband beamspace channel estimation for millimeter-wave MIMO systems relying on lens antenna arrays, *IEEE Trans. Signal Process.*, vol. 67, no. 18, pp. 4809–4824, 2019.

[52] M. Cui and L. Dai, Channel estimation for extremely large-scale MIMO: Far-field or near-field? arXiv preprint arXiv: 2108.07581, 2022.

[53] Q. -U. -A. Nadeem, A. Kammoun, A. Chaaban, M. Debbah, and M. -S. Alouini, Intelligent reflecting surface assisted wireless communication: Modeling and channel estimation, arXiv preprint arXiv: 1906.02360v2, 2019.

[54] T. L. Jensen and E. D. Carvalho, An optimal channel estimation scheme for intelligent reflecting surfaces based on a minimum variance unbiased estimator, in *Proc. 2020 IEEE Int. Conf. Acous., Sp. Signal Process. (ICASSP)*, Barcelona, Spain, 2020, pp. 5000–5004.

[55] Z. Wan, Z. Gao, and M. -S. Alouini, Broadband channel estimation for intelligent reflecting surface aided mmWave massive MIMO systems, in *Proc. 2020 IEEE Int. Conf. Commun. (ICC)*, Dublin, Ireland, 2020, pp. 1–6.

[56] A. M. Elbir, A. Papazafeiropoulos, P. Kourtessis, and S. Chatzinotas, Deep channel learning for large intelligent surfaces aided mm-Wave massive MIMO systems, *IEEE Wireless Commun. Lett.*, vol. 9, no. 9, pp. 1447–1451, 2020.

[57] Z. Wang, L. Liu, and S. Cui, Channel estimation for intelligent reflecting surface assisted multiuser communications, in *Proc. 2020 IEEE Wireless Commun. Net. Conf. (WCNC)*, Seoul, Republic of Korea, 2020, pp. 1–6.

[58] J. Lin, G. Wang, R. Fan, T. A. Tsiftsis, and C. Tellambura, Channel estimation for wireless communication systems assisted by large intelligent surfaces, arXiv preprint arXiv: 1911.02158v1, 2019.

[59] A. Taha, M. Alrabeiah, and A. Alkhateeb, Enabling large intelligent surfaces with compressive sensing and deep learning, *IEEE Access*, vol. 9, no. 2, pp. 44304–44321, 2021.

[60] L. Wei, C. Huang, G. C. Alexandropoulos, C. Yuen, Z. Zhang, and M. Debbah, Channel estimation for RIS-empowered multi-user MISO wireless communications, *IEEE Trans. Commun.*, vol. 69, no. 6, pp. 4144–4157, 2021.

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

[62] G. Zhou, C. Pan, H. Ren, K. Wang, and A. Nallanathan, A framework of robust transmission design for IRS-aided MISO communications with imperfect cascaded channels, *IEEE Trans. Signal Process.*, vol. 68, pp. 5092–5106, 2020.

[63] Y. Cui and H. Yin, An efficient CSI acquisition method for intelligent reflecting surface-assisted mmWave networks, arXiv preprint arXiv: 1912.12076v1, 2019.

[64] B. Zheng and R. Zhang, Intelligent reflecting surface-enhanced OFDM: Channel estimation and reflection optimization, *IEEE Wireless Commun. Lett.*, vol. 9, no. 4, pp. 518–522, 2019.

[65] P. Wang, J. Fang, H. Duan, and H. Li, Compressed channel estimation for intelligent reflecting surface-assisted millimeter wave systems, *IEEE Signal Process. Lett.*, vol. 27, pp. 905–909, 2020.

[66] Y. Guo, P. Sun, Z. Yuan, C. Huang, Q. Guo, Z. Wang, and C. Yuen, Efficient channel estimation for RIS-aided MIMO communications with unitary approximate message passing, arXiv preprint arXiv: 2112.15281, 2021.

[67] D. Shen and L. Dai, Dimension reduced channel feedback for reconfigurable intelligent surface aided wireless communications, *IEEE Trans. Commun.*, vol. 69, no. 11, pp. 7748–7760, 2021.

[68] Z. He and X. Yuan, Cascaded channel estimation for large intelligent metasurface assisted massive MIMO, *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 210–214, 2019.

[69] J. Mirza and B. Ali, Channel estimation method and phase shift design for reconfigurable intelligent surface assisted MIMO networks, *IEEE Trans. Cog. Commun. Netw.*, vol. 7, no. 2, pp. 441–451, 2021.

[70] J. He, M. Leinonen, H. Wymeersch, and M. Juntti, Channel estimation for RIS-aided mmWave MIMO channels, arXiv preprint arXiv: 2002.06453, 2020.

[71] G. T. D. Araújo and A. F. D. Almeida, PARAFAC-based channel estimation for intelligent reflective surface assisted MIMO system, in *Proc. 2020 IEEE 11th Sensor Arr. and Multichannel Signal Process. WkSp. (SAM)*, Hangzhou, China, 2020, pp. 1–5.

[72] B. Ning, Z. Chen, W. Chen, Y. Du, and J. Fang, Channel estimation and hybrid beamforming for reconfigurable intelligent surfaces assisted THz communications, arXiv preprint arXiv: 1912.11662, 2019.

[73] H. Liu, X. Yuan, and Y. Zhang, Matrix-calibration-based cascaded channel estimation for reconfigurable intelligent surface assisted multiuser MIMO, *IEEE J. Sel.*

Areas Commun., vol. 38, no. 11, pp. 2621–2636, 2020.

[74] J. Huang, C. Xing, and C. Wang, Simultaneous wireless information and power transfer: Technologies, applications, and research challenges, *IEEE Commun. Mag.*, vol. 55, no. 11, pp. 26–32, 2017.

[75] T. G. Kolda and B. W. Bader, Tensor decompositions and applications, *SIAM Review*, vol. 51, no. 3, pp. 455–500, 2009.

[76] Q. Wu, G. Y. Li, W. Chen, D. W. K. Ng, and R. Schober, An overview of sustainable green 5G networks, *IEEE Wireless Commun. Lett.*, vol. 24, no. 4, pp. 72–80, 2017.

[77] I. Krikidis, S. Timotheou, S. Nikolaou, G. Zheng, D. W. K. Ng, and R. Schober, Simultaneous wireless information and power transfer in modern communication systems, *IEEE Commun. Mag.*, vol. 52, no. 11, pp. 104–110, 2014.

[78] C. Pan, H. Ren, M. Elkashlan, A. Nallanathan, J. Wang, and L. Hanzo, Intelligent reflecting surface aided MIMO broadcasting for simultaneous wireless information and power transfer, *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1719–1734, 2020.

[79] Q. Wu and R. Zhang, Weighted sum power maximization for intelligent reflecting surface aided SWIPT, *IEEE Wireless Commun. Lett.*, vol. 9, no. 5, pp. 586–590, 2019.

[80] D. Mishra and H. Johansson, Channel estimation and low-complexity beamforming design for passive intelligent surface assisted MISO wireless energy transfer, in *Proc. 2019 IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Brighton, UK, 2019, pp. 4659–4663.

[81] L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-lin, and Z. Wang, Non-orthogonal multiple access for 5G: Solutions, challenges, opportunities, and future research trends, *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 74–81, 2015.

[82] Y. Liu, Z. Qin, M. Elkashlan, Z. Ding, A. Nallanathan, and L. Hanzo, Nonorthogonal multiple access for 5G and beyond, *Proc. IEEE*, vol. 105, no. 12, pp. 2347–2381, 2017.

[83] Z. Ding and H. V. Poor, A simple design of IRS-NOMA transmission, *IEEE Commun. Lett.*, vol. 24, no. 5, pp. 1119–1123, 2020.

[84] G. Yang, X. Xu, and Y. Liang, Intelligent reflecting surface assisted non-orthogonal multiple access, in *Proc. 2020 IEEE Wireless Commun. Net. Conf. (WCNC)*, Seoul, Republic of Korea, 2020, pp. 1–6.

[85] X. Liu, Y. Liu, Y. Chen, and H. V. Poor, RIS enhanced massive non-orthogonal multiple access networks: Deployment and passive beamforming design, *IEEE J. Sel. Areas Commun.*, vol. 39, no. 4, pp. 1057–1071, 2020.

[86] N. Yang, L. Wang, G. Geraci, M. Elkashlan, J. Yuan, and M. D. Renzo, Safeguarding 5G wireless communication networks using physical layer security, *IEEE Commun. Mag.*, vol. 53, no. 4, pp. 20–27, 2015.

[87] S. Hong, C. Pan, H. Ren, K. Wang, and A. Nallanathan, Artificial-noise-aided secure MIMO wireless communications via intelligent reflecting surface, *IEEE Trans. Commun.*, vol. 68, no. 12, pp. 7851–7866, 2020.

[88] J. Chen, Y. Liang, Y. Pei, and H. Guo, Intelligent reflecting surface: A programmable wireless environment for physical layer security, *IEEE Access*, vol. 7, pp. 82599–82612, 2019.

[89] X. Yu, D. Xu, Y. Sun, D. W. K. Ng, and R. Schober, Robust and secure wireless communications via intelligent reflecting surfaces, *IEEE J. Sel. Areas in Commun.*, vol. 38, no. 11, pp. 2637–2652, 2020.

[90] X. Guan, Q. Wu, and R. Zhang, Intelligent reflecting surface assisted secrecy communication via joint beamforming and jamming, arXiv preprint arXiv: 1907.12839v1, 2019.

[91] X. Yu, D. Xu, and R. Schober, Enabling secure wireless communications via intelligent reflecting surfaces, in *Proc. 2019 IEEE Global Commun. Conf. (GLOBECOM)*, Waikoloa, HI, USA, 2019, pp. 1–6.

[92] M. Cui, G. Zhang, and R. Zhang, Secure wireless communication via intelligent reflecting surface, *IEEE Wireless Commun. Lett.*, vol. 8, no. 5, pp. 1410–1414, 2019.

[93] S. Li, B. Duo, X. Yuan, Y. Liang, and M. D. Renzo, Reconfigurable intelligent surface assisted UAV communication: Joint trajectory design and passive beamforming, *IEEE Wireless Commun. Lett.*, vol. 9, no. 5, pp. 716–720, 2020.

[94] L. Yang, F. Meng, J. Zhang, M. O. Hasna and M. D. Renzo, On the performance of RIS-assisted dual-hop UAV communication systems, *IEEE Trans. Veh. Tech.*, vol. 69, no. 9, pp. 10385–10390, 2020.

[95] W. Zhao, G. Wang, S. Atapattu, T. A. Tsiftsis, and X. Ma, Performance analysis of large intelligent surface aided backscatter communication systems, *IEEE Wireless Commun. Lett.*, vol. 9, no. 7, pp. 962–966, 2020.

[96] S. Basharat, S. A. Hassan, A. Mahmood, Z. Ding, and M. Gidlund, Reconfigurable intelligent surface-assisted backscatter communication: A new frontier for enabling 6G IoT Networks, arXiv preprint arXiv: 2107.07813, 2021.

[97] F. Liu, A. Pitilakis, M. S. Mirmoosa, O. Tsilipakos, X. Wang, A. C. Tasolamprou, S. Abadal, A. Cabellos-Aparicio, E. Alarcón, C. Liaskos, et al., Programmable metasurfaces: State of the art and prospects, in *Proc.*

2018 IEEE Int. Symp. Circuits Syst. (ISCAS), Florence, Italy, 2018, pp. 1–5.

[98] H. -T. Chen, A. J. Taylor, and N. Yu, A review of metasurfaces: Physics and applications, *Rep. Prog. Phys.*, vol. 79, no. 7, p. 076401, 2016.

[99] S. Foo, Liquid-crystal reconfigurable metasurface reflectors, in *Proc. 2017 IEEE Int. Symp. Antennas Propag. & USNC/URSI Nat. Radio Sci. Meeting*, San Diego, CA, USA, 2017, pp. 2069–2070.

[100] D. Shrekenhamer, W. -C. Chen, and W. J. Padilla, Liquid crystal tunable metamaterial absorber, *Phys. Rev. Lett.*, vol. 110, no. 17, p. 177403, 2013.

[101] C. Huang, C. Zhang, J. Yang, B. Sun, B. Zhao, and X. Luo, Reconfigurable metasurface for multifunctional control of electromagnetic waves, *Adv. Opt. Mater.*, vol. 5, no. 22, p. 1700485, 2017.

[102] S. Kim, M. S. Jang, V. W. Brar, K. W. Mauser, L. Kim, and H. A. Atwater, Electronically tunable perfect absorption in graphene, *Nano Lett.*, vol. 18, no. 2, pp. 971–979, 2018.

[103] N. Dabidian, S. Dutta-Gupta, I. Kholmanov, K. Lai, F. Lu, J. Lee, M. Jin, S. Trendafilov, A. Knanikaev, B. Fallahazad, et al., Experimental demonstration of phase modulation and motion sensing using graphene-integrated metasurfaces, *Nano Lett.*, vol. 16, no. 6, pp. 3607–3615, 2016.

[104] Y. -J. Yang, Y. -J. Huang, G. -J. Wen, J. -P. Zhong, H. -B. Sun, and G. Oghenemuero, Tunable broadband metamaterial absorber consisting of ferrite slabs and a copper wire, *Chin. Phys. B*, vol. 21, no. 3, p. 038501, 2012.

[105] D. Wang, L. Zhang, Y. Gu, M. Q. Mehmood, Y. Gong, A. Srivastava, L. Jian, T. Venkatesan, C. Qiu, and M. Hong, Switchable ultrathin quarter-wave plate in terahertz using active phase-change metasurface, *Sci. Rep.*, vol. 5, p. 15020, 2015.

[106] B. Zhu, Y. Feng, J. Zhao, and T. Jiang, Switchable metamaterial reflector/absorber for different polarized electromagnetic waves, *Appl. Phys. Lett.*, vol. 97, no. 5, p. 051906, 2010.

[107] C. Mias and J. H. Yap, A varactor-tunable high impedance surface with a resistive-lumped-element biasing grid, *IEEE Trans. Antennas Propag.*, vol. 55, no. 7, pp. 1955–1962, 2007.

[108] S. V. Hum and J. Perruisseau-Carrier, Reconfigurable reflectarrays and array lenses for dynamic antenna beam control: A review, *IEEE Trans. Antennas Propag.*, vol. 62, no. 1, pp. 183–198, 2014.

[109] S. Hu, F. Rusek, and O. Edfors, Beyond massive MIMO: The potential of data transmission with large intelligent surfaces, *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2746–2758, 2018.

[110] H. R. Kunsch, E. Agrell, and F. A. Hamprecht, Optimal lattices for sampling, *IEEE Trans. Inf. Theory*, vol. 51, no. 2, pp. 634–647, 2005.

[111] H. Yang, F. Yang, S. Xu, Y. Mao, M. Li, X. Cao, and J. Gao, A 1-bit 10×10 reconfigurable reflectarray antenna: Design, optimization, and experiment, *IEEE Trans. Antennas Propag.*, vol. 64, no. 6, pp. 2246–2254, 2016.

[112] C. Liaskos, S. Nie, A. Tsioliaridou, A. Pitsillides, S. Ioannidis, and I. Akyildiz, Realizing wireless communication through software-defined hypersurface environments, in *Proc. 2018 IEEE 19th Int. Symp. World Wireless, Mobile Multimedia Netw. (WoWMoM)*, Chania, Greece, 2018, pp. 14–15.

[113] 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, *China Commun.*, vol. 16, no. 5, pp. 46–61, 2019.

[114] V. F. Fusco and Q. Chen, Direct-signal modulation using a silicon microstrip patch antenna, *IEEE Trans. Antennas Propag.*, vol. 47, no. 6, pp. 1025–1028, 1999.

[115] W. Yao and Y. Wang, Direct antenna modulation - A promise for ultra-wideband (UWB) transmitting, in *Proc. IEEE MTT-S Int. Microw. Symp. Dig.*, Fort Worth, TX, USA, 2004, pp. 1273–1276.

[116] M. P. Daly, E. L. Daly, and J. T. Bernhard, Demonstration of directional modulation using a phased array, *IEEE Trans. Antennas Propag.*, vol. 58, no. 5, pp. 1545–1550, 2010.

[117] W. Tang, M. Z. Chen, J. Y. Dai, Y. Zeng, X. Zhao, S. Jin, Q. Cheng, and T. J. Cui, Wireless communications with programmable metasurface: New paradigms, opportunities, and challenges on transceiver design, *IEEE Wireless Commun.*, vol. 27, no. 2, pp. 180–187, 2020.

[118] L. Dai, B. Wang, M. Wang, X. Yang, J. Tan, S. Bi, S. Xu, F. Yang, Z. Chen, M. D. Renzo, et al., Reconfigurable intelligent surface-based wireless communications: Antenna design, prototyping, and experimental results, *IEEE Access*, vol. 8, pp. 45913–45923, 2020.

[119] DOCOMO conducts worlds first successful trial of transparent dynamic metasurface, https://www.nttdocomo.co.jp/english/info/mediacenter/pr/2020/011700.html, 2020.

[120] DOCOMO, AGC and Ericsson achieve worlds first 5G communication using glass antenna for 28 GHz, https://www.nttdocomo.co.jp/english/info/mediacenter/pr/2019/052900.html, 2019.

[121] Z. Zhang, H. Zhao, J. Wang, and Y. Shen, Signal-multiplexing ranging for network localization, *IEEE*

Trans. Wireless Commun., vol. 21, no. 3, pp. 1694–1709, 2021.

[122] T. L. Marzetta, Noncooperative cellular wireless with unlimited numbers of base station antennas, *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, 2010.

[123] X. Gao, L. Dai, S. Han, C. -L. I, and R. W. Heath, Energy-efficient hybrid analog and digital precoding for mmWave MIMO systems with large antenna arrays, *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 998–1009, 2016.

[124] J. Tan and L. Dai, Wideband beam tracking in THz massive MIMO systems, *IEEE J. Sel. Areas Commun.*, vol. 39, no. 6, pp. 1693–1710, 2021.

[125] E. Basar, Transmission through large intelligent surfaces: A new frontier in wireless communications, in *Proc. 2019 Eur. Conf. Netw. Commun. (EuCNC)*, Valencia, Spain, 2019, pp. 112–117.

[126] M. D. Renzo, M. Debbah, D. -T. Phan-Huy, A. Zappone, M. -S. Alouini, C. Yuen, V. Sciancalepore, G. C. Alexandropoulos, J. Hoydis, H. Gacanin, et al., Smart radio environments empowered by reconfigurable AI metasurfaces: An idea whose time has come, *EURASIP J. Wireless Commun. Netw.*, vol. 2019, no. 1, p. 129, 2019.

[127] Z. Zhang, H. Zhao, and Y. Shen, High-efficient ranging algorithms for wireless sensor network, in *Proc. 2019 11th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Xi'an, China, 2019, pp. 1–6.

[128] X. Gan, C. Zhong, C. Huang, and Z. Zhang, Ris-assisted multi-user MISO communications exploiting statistical CSI, *IEEE Trans. Commun.*, vol. 69, no. 10, pp. 6781–6792, 2021.

[129] D. Selimis, K. P. Peppas, G. C. Alexandropoulos, and F. I. Lazarakis, On the performance analysis of RIS-empowered communications over nakagami-m fading, *IEEE Commun. Lett.*, vol. 25, no. 7, pp. 2191–2195, 2021.

[130] J. Zhu, Z. Zhang, Z. Wan, and L. Dai, Finite-time capacity: Making exceed-shannon possible? arXiv preprint arXiv: 2111.00444, 2021.

[131] Z. Wan, J. Zhu, Z. Zhang, and L. Dai, Capacity for electromagnetic information theory, arXiv preprint arXiv: 2111.00496, 2021.

[132] K. Liu, Z. Zhang, L. Dai, S. Xu, and F. Yang, Active reconfigurable intelligent surface: Fully-connected or sub-connected? *IEEE Commun. Lett.*, vol. 26, no. 1, pp. 167–171, 2021.

[133] A. M. Salhab and L. Yang, Mixed RF/FSO relay networks: RIS-equipped RF source vs RIS-aided RF source, *IEEE Wireless Commun. Lett.*, vol. 10, no. 8, pp. 1712–1716, 2021.

[134] E. Basar and H. V. Poor, Present and future of reconfigurable intelligent surface-empowered communications [perspectives], *IEEE Signal Process. Mag.*, vol. 38, no. 6, pp. 146–152, 2021.

[135] I. Al-Nahhal, O. A. Dobre, E. Basar, T. M. N. Ngatched, and S. Ikki, Reconfigurable intelligent surface optimization for uplink sparse code multiple access, *IEEE Commun. Lett.*, vol. 26, no. 1, pp. 133–137, 2021.

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

[137] K. Liu, Z. Zhang, and L. Dai, User-side RIS: Realizing large-scale array at user side, in *Proc. 2021 IEEE Global Commun. Conf. (IEEE GLOBECOM)*, Madrid, Spain, pp. 1–6, 2021.

Talha Mir received the BS degree from Balochistan University of Information Technology, Engineering, and Management Sciences (BUITEMS) in 2007, the MS degree from Bradford University, UK, in 2011, and the PhD degree from Tsinghua University, Beijing, China, in 2019. He is currently serving as an assistant professor in the Department of Electronic Engineering, BUITEMS. His current research interests include 5G/6G wireless communication technologies (massive MIMO, millimeter-wave/Terahertz communications), and machine learning for future wireless communications.

Muhammad Zain Siddiqi is currently pursuing the PhD degree in electronics engineering at Tsinghua University, Beijing, China. His current research focuses on reconfigureable intelligent surface, massive MIMO, mmWave communications, resource, power allocation in wireless networks, and green communication.