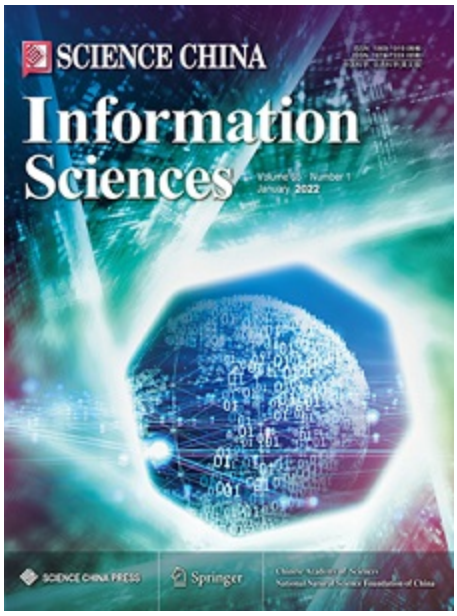

Multi-Beam Design: From Passive RIS to Active RIS

Journal:	*SCIENCE CHINA Information Sciences*
Manuscript ID	SCIS-2023-0063
Manuscript Type:	Research Paper
Date Submitted by the Author:	21-Jan-2023
Complete List of Authors:	Su, Ruochen; Tsinghua University, Department of Electronic Engineering Zhang, Zijian; Tsinghua University, Department of Electronic Engineering Dai, Linglong; Tsinghua University, Department of Electronic Engineering
Keywords:	Reconfigurable intelligent surface (RIS), precoding, multi-beam design, active RIS, communication systems
Speciality:	signal processing for communications < Information & Communication Engineering

SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

Multi-Beam Design: From Passive RIS to Active RIS

Ruochen Su, Zijian Zhang & Linglong Dai*

Department of Electronic Engineering, Tsinghua University,
Beijing National Research Center for Information Science and Technology (BNRist), Beijing 100084, *China*

Abstract Thanks to the capability of intelligently manipulating the propagation environment, reconfigurable intelligent surface (RIS) has emerged as a promising technology for the sixth generation (6G) wireless communications. Precoding is essential to enable RIS with high array gains. In most existing works, the precoding design of RIS relies on the complete channel state information (CSI) knowledge of RIS reflected links. For more practical cases when the CSI is not completely known, an efficient precoding design scheme for RIS to serve multiple users is the angle-based multi-beam design, which is however challenging due to the constant modulus constraint of RIS. To address this challenge, in this paper, we introduce the recently proposed active RIS and propose an adaptive multi-beam design framework. Specifically, we first formulate a multi-beam design problem for active RIS with power constraints, which accounts for the power allocation among multiple users. To solve the problem, we propose a discrete Fourier transform (DFT)-based codebook design scheme to generate multiple beams with low complexity. To further improve the beam gain, we develop a majorization-minimization (MM) algorithm-based scheme to solve the problem in an iterative way. Simulation results demonstrate that proposed schemes for active RIS can generate high-gain multiple beams, which realize high sum-rate in active RIS-assisted wireless communication systems.

Keywords Reconfigurable intelligent surface (RIS), precoding, multi-beam design, active RIS

Citation

1 Introduction

Reconfigurable intelligent surface (RIS) has attracted extensive attentions owing to the potential to enhance the future sixth generation (6G) wireless communications [1,2]. Thanks to the recent development of metamaterials, RIS composed of massive tuned elements enables the controllable manipulation of reflected signals [3,4]. By properly adjusting reflection coefficients of RIS to make reflected signals coherently or deconstructively add up at the user, RIS is capable of improving the signal-to-noise ratio (SNR) and suppressing interferences by providing additional signal propagation paths, which thus leads to the increased spectrum efficiency [5], enlarged wireless coverage [6], reduced power consumption, etc [7]. As a result, RIS has become a candidate technology for future 6G networks [8].

1.1 Prior Works

To unleash the benefits of RIS, precoding design is essential by jointly optimizing the transmit precoding vector at the base station (BS) and the precoding matrix at the RIS in RIS-assisted communications. Existing studies on the precoding design of RIS typically assume the complete knowledge of channel state information (CSI) [9–11]. For example, in [9], the precoding design at the RIS and the BS were alternately optimized to maximize the sum-rate through a block coordinate descent method. Besides, a time-division based dynamic beamforming framework of RIS was proposed to achieve the multi-beam gain for a multi-user wireless energy transfer system [10]. Nevertheless, the acquisition of the complete CSI for multiple users will result in high complexity and high or even unaffordable pilot overhead [12]. In practical scenarios, the information of spatial angles from the RIS to users, i.e., the incomplete CSI, is easy to be acquired through beam training. Thus, compared with most existing precoding methods

* Corresponding author (email: daill@tsinghua.edu.cn)

based on the complete CSI, the multi-beam design with incomplete CSI is more practical to serve multiple users [13].

The multi-beam design with incomplete CSI has been extensively studied and practically realized in multiple-input multiple-output (MIMO) communications [14–16]. For instance, a hierarchical codebook design was investigated to search the best beam direction in millimeter-wave communications based on the techniques of deactivation antenna processing and sub-array [14]. A multi-resolution codebook design was proposed to search the dominant directions for the single-path channel based on a generalized discrete Fourier transform (DFT) matrix [15]. In addition, the authors in [16] proposed a two-step codeword design in millimeter-wave massive MIMO by considering the hardware constraints, which aims at generating desired beam patterns. However, existing codebook design schemes for MIMO systems cannot be directly applied to RIS due to the constant modulus constraint of RIS. This is because the precoding designs for classical passive RIS consisting of massive passive elements mainly optimize the phase shifting of each RIS element, which cannot adjust the amplitude of the RIS element [5, 17]. If we extract the phase information of codewords in the designed codebook to satisfy the constant modulus constraint, a severe loss of beam gains will be introduced. Therefore, the multi-beam design for RIS-assisted communication systems with incomplete CSI is still a critical problem. To the best of our knowledge, there is no practical solution in the literature.

1.2 Our Contributions

To fill in the gap, we introduce active RIS and propose an adaptive multi-beam design framework with incomplete CSI in this paper[1)]. The main contributions are summarized as follows.

- First, active RIS is introduced to compensate for the severe performance loss of multi-beam design caused by the constant modulus constraint of passive RIS. Different form passive RIS, active RIS is capable of amplifying reflected signals, which can break through the constant modulus constraint and increase the degree of freedom (DoF) for design. Based on the active RIS, we formulate the adaptive multi-beam design problem with incomplete CSI subject to the power constraint of the BS and the radiation power constraint of the active RIS.
- Then, to generate multiple beams towards target directions of multiple users, we propose a heuristic DFT-based codebook design scheme to solve the formulated problem with low complexity. To further improve the beam gains, we propose a majorization-minimization (MM) algorithm-based scheme to solve the multi-beam design problem, which compensates for the performance loss of the DFT-based scheme. In particular, the MM-based algorithm decomposes the non-convex optimization problem into tractable sub-problems. Then, the sub-problems are iteratively solved to approach the optimal solution, based on which the precoding scheme at the BS and RIS is developed. The convergence and the complexity of proposed schemes are analyzed accordingly.
- Finally, we optimize the power allocation among multiple users to further improve the system performance. The optimal power allocation to target multiple beams aiming at maximizing the minimum user rate or the users' sum-rate is provided, respectively. Simulation results show that our proposed DFT-based scheme and MM-based scheme are capable of generating high-gain multiple beams, which provide practical solutions to the multi-beam design for active RIS with incomplete CSI.

Organization: The rest of this paper is organized as follows. We introduce the system model and the formulated problem in Section 2. We propose the DFT-based and the MM-based multi-beam design schemes for active RIS in Section 3. Simulations results are provided in Section 4, followed by the conclusions in Section 5.

Notation: Lower-case boldface letter $\mathbf{a}$ and upper-case boldface letter $\mathbf{A}$ represent vectors and matrices, respectively. The conjugate transpose of $\mathbf{a}$ and $\mathbf{A}$ are represented by $\mathbf{a}^H$ and $\mathbf{A}^H$, respectively. The i-th column of $\mathbf{A}$ is given by $\mathbf{A}_{[:,i]}$. The l_2 norm of $\mathbf{a}$ is represented by $\|\mathbf{a}\|_2$. $\mathrm{diag}\{\mathbf{a}\}$ represents the diagonalization operation of $\mathbf{a}$. $\mathbf{A} \otimes \mathbf{B}$ represents the Kronecker product of $\mathbf{A}$ and $\mathbf{B}$. Besides, $\mathcal{U}(-a, a)$ is the probability density function of uniform distribution within $(-a, a)$, and the probability density function of complex Gaussian distribution can be denoted as $\mathcal{CN}\left(\mu, \sigma^2\right)$ with mean μ and variance σ^2. The expectation of random variable X is expressed as $E[X]$. Finally, the real part of complex value a is denoted by $\mathrm{Re}\{a\}$.

1) Simulation codes will be provided to reproduce the results after the presentation of this paper in the following link: http://oa.ee.tsinghua.edu.cn/dailinglong/publications/publications.html.

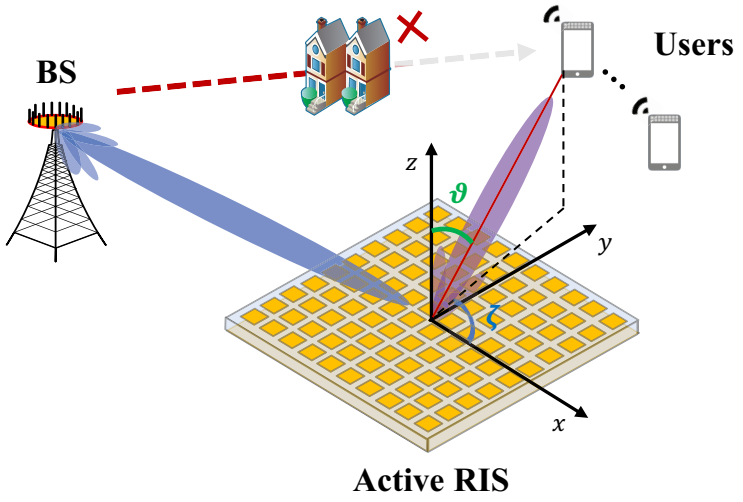

Figure 1 The active RIS-assisted multi-user system.

2 System Model

The model of active RIS-assisted communication systems is first introduced in Subsection 2.1. Then, we formulate the multi-beam design problem in Subsection 2.2.

2.1 Channel Model

The passive RIS with phase shifting circuits in each element can only reflect incident signals passively by tuning the phase shifting of RIS elements. In passive RIS-assisted communication systems, the multiplicative fading effect is a major challenge [18]. In specific, the path loss of the equivalent BS-RIS-user channel could be given by the product of the loss of BS-RIS channel paths and that of RIS-user channel paths. The recently proposed active RIS with reflection-type power amplifiers enables more flexible adjustment of reflection coefficients of RIS elements including the amplitude and the phase [19]. Thus, the active RIS with affordable power consumption is capable of reflecting incident signals with power amplification, which can overcome the system capacity limitation caused by the multiplicative fading effect [20–23]. Notice that different from passive RIS, the signal amplification of active RIS will generate non-negligible thermal noise [17].

As shown in Fig. 1, an active RIS-assisted multi-user system is considered. A BS with N_t antennas deployed on a uniform linear array (ULA) serves K single-antenna users, aided by a RIS employing $N = N_1 \times N_2$ elements arranged on a uniform planar array (UPA). The direct path between users and the BS is blocked. Note that the frequency division multiplexing can be utilized to serve multiple users. Therefore, the received signal y_k at user k is denoted as

$$y_k = \mathbf{h}_k^T \mathbf{\Theta} \mathbf{G} \mathbf{x} + \mathbf{h}_k^T \mathbf{\Theta} \mathbf{n}_R + n_k, \tag{1}$$

where $\mathbf{G} \in \mathbb{C}^{N \times N_t}$ is the BS-RIS channel, and $\mathbf{h}_k^T = [h_{k,1}, h_{k,2}, \cdots, h_{k,N}] \in \mathbb{C}^{1 \times N}$ represents the channel from the RIS to user k. $\mathbf{x} = \mathbf{w}s \in \mathbb{C}^{N_t \times 1}$ denotes the transmitted signal at the BS, where $\mathbf{w} \in \mathbb{C}^{N_t \times 1}$ represents the transmit precoding vector, and $s \in \mathbb{C}$ represents the transmitted information symbol satisfying $E[|s|^2] = 1$. $\mathbf{n}_R$ represents the noise induced by active RIS satisfying $\mathbf{n}_R \sim \mathcal{CN}\left(\mathbf{0}_N, \sigma^2 \mathbf{I}_N\right)$, and n_k denotes the additive white Gaussian noise with zero mean and variance σ_k^2. Besides, the reflection coefficients $\mathbf{\Theta}$ at the RIS could be expressed as

$$\mathbf{\Theta} = \mathrm{diag}(\boldsymbol{\theta}) = \mathrm{diag}([\beta_1 e^{j\theta_1}, \beta_2 e^{j\theta_2}, \cdots, \beta_N e^{j\theta_N}]^T). \tag{2}$$

For the small-scale fading, we adopt the geometric channel model [24]. The BS-RIS channel is expressed as

$$\mathbf{G} = \sum_{i=1}^{L_1} g_{1,i} e^{-j2\pi\tau_{1,i}f} \mathbf{a}\left(\vartheta_{1,i}^r, \zeta_{1,i}^r\right) \mathbf{b}(\phi_i)^H, \tag{3}$$

where $\tau_{1,i} \in \mathbb{R}^+$ and $g_{1,i} \in \mathbb{C}$ represent path delay and path gain of the i-th path, respectively, and L_1 denotes the number of paths. The elevation and azimuth angle of the angle-of-arrival (AoA) at the RIS for the i-th path are represented by $\vartheta_{1,i}^t$ and $\zeta_{1,i}^t$, respectively. ϕ_i represents the transmit angle for the i-th path at the BS. $\mathbf{a}$ denotes the steering vector of the UPA, which is denoted as

$$\begin{aligned}\mathbf{a}\left(\vartheta,\zeta\right) = \tfrac{1}{\sqrt{N_1 N_2}}\left[1,\cdots,e^{j2\pi\frac{f}{c}(N_1-1)d\sin\vartheta\sin\zeta}\right]^T \\ \otimes\left[1,\cdots,e^{j2\pi\frac{f}{c}(N_2-1)d\cos\vartheta}\right]^T,\end{aligned} \tag{4}$$

and $\mathbf{b}$ denotes the steering vector of the ULA, which is given by

$$\mathbf{b}\left(\phi\right) = \tfrac{1}{\sqrt{N_{\mathrm{t}}}}\left[1,\cdots,e^{j2\pi\frac{f}{c}(N_{\mathrm{t}}-1)d\sin\phi}\right]^T. \tag{5}$$

Notice that the antenna spacing d is set to half of the wavelength. Similarly, the RIS-user channel can be given by

$$\mathbf{h}_k = \sum_{i=1}^{L_2} g_{2,k,i} e^{-j2\pi\tau_{2,k,i}f}\mathbf{a}\big(\vartheta_{2,k,i}^t,\zeta_{2,k,i}^t\big), \tag{6}$$

where $\tau_{2,k,i} \in \mathbb{R}^+$ and $g_{2,k,i} \in \mathbb{C}$ denote path delay and path gain of the i-th path for user k, respectively, and L_2 represents the number of channel paths. $\vartheta_{2,k,i}^t$ and $\zeta_{2,k,i}^t$ denote the elevation and azimuth angle of the angle-of-departure (AoD) from the RIS to user k for the i-th path, respectively. The angles satisfy $\vartheta \in \mathcal{U}(0,\pi)$, $\zeta \in \mathcal{U}(0,\pi)$, and $\phi \in \mathcal{U}(-\pi/2,\pi/2)$.

The large-scale fading of the end-to-end BS-RIS-user channel is expressed as [11]

$$f(d_1,d_2) = C d_1^{-\alpha_1} d_2^{-\alpha_2}, \tag{7}$$

where C represents the channel fading, α_1 and α_2 represent the path loss. Besides, d_1 and d_2 represent the distance between the RIS and BS, and that between the RIS and users, respectively.

2.2 Problem Formulation

Let $\mathbf{C} \in \mathbb{C}^{N\times N}$ denote a DFT matrix corresponding to the DFT codebook, where the n-th column of $\mathbf{C}$, i.e., $\mathbf{c}_n \in \mathbb{C}^{N\times 1}$, represents the n-th codeword. The incomplete CSI of the RIS-user channel is available by sequential beam training of K users based on the DFT codebook. In a more practical case, we can obtain the equivalent spatial angle $\vartheta_{2,k}^t$ and $\zeta_{2,k}^t$ for user k as shown in Fig. 1, as well as the corresponding beam gain in the spatial angle grid. Therefore, the target multiple beams $\mathbf{v} \in \mathbb{C}^{N\times 1}$ to serve all users can be adaptively designed by superposing the beam gains in K angle grids. Besides, the deployment of the RIS can be well-designed in a way that allows for the acquisition of BS-RIS channel $\mathbf{G}$ through statistical information in practice.

To fully utilize the benefits of active RIS, precoding design is required to balance the amplification of incident signals and thermal noise. We focus on minimizing the error of multi-beam design l_0 by jointly optimizing the precoding vector $\mathbf{w}$ at the BS and the precoding matrix $\mathbf{\Theta}$ at the RIS. The adaptive multi-beam design problem subject to the power constraints is formulated as

$$\mathcal{P}_1: \min_{\mathbf{w},\mathbf{\Theta}} \; l_0 = \left\|\mathbf{v} - \mathbf{C}^H\mathbf{\Theta}\mathbf{G}\mathbf{w}\right\|_2^2 \tag{8a}$$

$$\text{s.t.} \;\; C_1: \|\mathbf{w}\|_2^2 \leqslant P_1, \tag{8b}$$

$$C_2: \|\mathbf{\Theta}\mathbf{G}\mathbf{w}\|_2^2 + \|\mathbf{\Theta}\|_F^2\,\sigma^2 \leqslant P_2, \tag{8c}$$

where C_1 denotes the power constraint of the BS, C_2 represents the radiation power constraint of the RIS [19].

Notice that the power constraints C_1, C_2, and the objective function l_0 are non-convex, which makes it difficult to directly solve $\mathcal{P}_1$. In the next section, two multi-beam design schemes will be proposed.

3 Proposed Multi-Beam Design Scheme

We first propose a DFT-based codebook design scheme to generate multiple beams with low complexity in this section. Then, a MM algorithm-based scheme is proposed to solve $\mathcal{P}_1$, which provides improved beam gain at the cost of higher complexity compared with the DFT-based scheme. Besides, we consider the power allocation among multiple users to further enhance the system performance.

3.1 Proposed DFT-Based Scheme

In this subsection, we propose a DFT-based codebook design scheme with low complexity to generate the target multiple beams $\mathbf{v}$.

Firstly, the transmit precoding vector $\mathbf{w}$ can be calculated through the singular value decomposition (SVD) precoding. We can express the ordered SVD of $\mathbf{G}$ as $\mathbf{G} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ according to statistical information of BS-RIS channel $\mathbf{G}$. Note that $\mathbf{V}_{[:,1]}$ represents the eigenvector corresponding to the largest eigenvalue of $\mathbf{G}$. Thus, the transmit precoding vector $\mathbf{w}$ can be given by

$$\mathbf{w} = \frac{\mathbf{V}_{[:,1]}}{\|\mathbf{V}_{[:,1]}\|_2}\sqrt{P_1}. \tag{9}$$

which satisfies the power constraint in (8b).

Then, in a beam training procedure between the active RIS and user k, the precoding design of active RIS traverses all codewords from $\mathbf{C}$ [14]. We select the codeword $\mathbf{c}_n$ corresponding to maximum received power as the precoding vector $\mathbf{f}_k$, which can be expressed as

$$\mathbf{f}_k = \arg\max_{\mathbf{c}_n} \left|\mathbf{h}_k^T \mathbf{c}_n\right|. \tag{10}$$

After acquiring all precoding vectors for K users, we can superimpose them with corresponding weights to obtain the precoding design $\mathbf{\Theta}$ at the active RIS, which can be given by

$$\mathbf{\Theta} = \sum_{k=1}^{K} \omega_k \mathbf{f}_k. \tag{11}$$

Note that $\mathbf{\Theta}$ should be multiplied by a coefficient to satisfy the power constraint in (8c). The weight ω_k is determined by the power allocation, which will be discussed in Subsection 3.3.

It is worth noting that, the DFT-based scheme only focuses on RIS precoding, which does not consider the incident signal from the BS and the amplification of thermal noise at the active RIS. Consequently, significant performance degradation will occur when utilizing the weighted superposition of DFT codewords in active RIS-assisted wireless communication systems.

3.2 Proposed MM-Based Scheme

To compensate for the performance loss of the DFT-based scheme, an alternative scheme based on the MM algorithm is proposed in this subsection, which can realize higher beam gain at the cost of increased complexity.

Due to the non-convex property, we decouple two optimization variables in $\mathcal{P}_1$. The transmit precoding vector $\mathbf{w}$ at the BS can be acquired according to (9). To optimize the precoding matrix $\mathbf{\Theta}$, we derive the expansion of (8a) and (8c) and carry out variable substitution. We have $\mathbf{\Theta}\mathbf{G}\mathbf{w} = \mathrm{diag}(\boldsymbol{\theta})\mathbf{G}\mathbf{w} = \mathrm{diag}(\mathbf{G}\mathbf{w})\boldsymbol{\theta}$. By defining

$$\mathbf{A} = \mathbf{C}^H \mathrm{diag}(\mathbf{G}\mathbf{w}),\ \mathbf{B} = \mathrm{diag}(\mathbf{G}\mathbf{w})^H \mathrm{diag}(\mathbf{G}\mathbf{w}) + \sigma^2\mathbf{I},$$

$$\boldsymbol{\psi} = \mathbf{B}^{\frac{1}{2}}\boldsymbol{\theta},\ \tilde{\mathbf{\Lambda}} = \mathbf{B}^{-\frac{1}{2}}\mathbf{A}^H\mathbf{A}\mathbf{B}^{-\frac{1}{2}}, \text{and } \tilde{\beta} = \mathbf{B}^{-\frac{1}{2}}\mathbf{A}^H\mathbf{v},$$

$\mathcal{P}_1$ can be converted to

$$\mathcal{P}_2: \min_{\boldsymbol{\psi}}\ f(\boldsymbol{\psi}) = \boldsymbol{\psi}^H\tilde{\mathbf{\Lambda}}\boldsymbol{\psi} - 2\mathrm{Re}\left\{\tilde{\beta}^H\boldsymbol{\psi}\right\} \tag{12a}$$

$$\text{s.t. } C_1: \|\boldsymbol{\psi}\|_2^2 \leqslant P_2. \tag{12b}$$

The objective function $f(\boldsymbol{\psi})$ in (12a) is still non-convex. To solve $\mathcal{P}_2$, the MM algorithm is adopted [25]. The crux is to construct a series of approximate sub-problems and iteratively solve them. In specific, we introduce a series of proxy functions $g(\boldsymbol{\psi}|\boldsymbol{\psi}^t)$ as the objective function of corresponding sub-problems. Let $\boldsymbol{\psi}_t$ represent the solution to sub-problems in the t-th iteration. $g(\boldsymbol{\psi}|\boldsymbol{\psi}^t)$ serves as an upper bound of the original function in the $(t+1)$-th iteration, whose solution can be denoted as $\boldsymbol{\psi}_{t+1} = \arg\min_{\boldsymbol{\psi}} g(\boldsymbol{\psi}|\boldsymbol{\psi}^t)$. The proxy function $g(\boldsymbol{\psi}|\boldsymbol{\psi}^t)$ should satisfy the following three features:

$$f(\boldsymbol{\psi}^t) = g(\boldsymbol{\psi}^t|\boldsymbol{\psi}^t), \tag{13a}$$

$$f(\boldsymbol{\psi}) \leqslant g(\boldsymbol{\psi}|\boldsymbol{\psi}^t), \tag{13b}$$

$$\nabla_{\boldsymbol{\psi}} f(\boldsymbol{\psi})|_{\boldsymbol{\psi}=\boldsymbol{\psi}^t} = \nabla_{\boldsymbol{\psi}} g(\boldsymbol{\psi}|\boldsymbol{\psi}^t)|_{\boldsymbol{\psi}=\boldsymbol{\psi}^t}. \tag{13c}$$

(13a) means that the value of the proxy function and the original one at point $\boldsymbol{\psi}^t$ should be equal. (13b) represents that the proxy function serves as an upper bound of the original one. (13c) represents that the first-order gradient of the original function is the same as that of the proxy one.

Based on the principles, we construct the proxy function expressed as

$$g(\boldsymbol{\psi}|\boldsymbol{\psi}^t) = y\left(\boldsymbol{\psi}|\boldsymbol{\psi}^t\right) - 2\mathrm{Re}\left\{\tilde{\beta}^H\boldsymbol{\psi}\right\}, \tag{14}$$

where $y(\boldsymbol{\psi})$ is defined according to [9], which can be given by

$$\begin{aligned} y\left(\boldsymbol{\psi}|\boldsymbol{\psi}^t\right) \triangleq\ & \boldsymbol{\psi}^{\mathrm{H}}\mathbf{X}\boldsymbol{\psi} - 2\,\mathrm{Re}\left\{\boldsymbol{\psi}^{\mathrm{H}}(\mathbf{X}-\tilde{\boldsymbol{\Lambda}})\boldsymbol{\psi}^t\right\} \\ & + \left(\boldsymbol{\psi}^t\right)^{\mathrm{H}}(\mathbf{X}-\tilde{\boldsymbol{\Lambda}})\boldsymbol{\psi}^t \geqslant \boldsymbol{\psi}^{\mathrm{H}}\tilde{\boldsymbol{\Lambda}}\boldsymbol{\psi}, \end{aligned} \tag{15}$$

where $\mathbf{X} = \lambda_{\max}\mathbf{I}_N$, and $\lambda_{\max}$ is the maximum eigenvalue of $\tilde{\boldsymbol{\Lambda}}$. According to (14)-(15), $\mathcal{P}_2$ can be converted to a series of sub-problems expressed as

$$\mathcal{P}_3: \ \max_{\boldsymbol{\psi}}\ l_1 = \mathrm{Re}\left\{\boldsymbol{\psi}^H\mathbf{q}^t\right\} \tag{16a}$$

$$\text{s.t.}\ \ C_1: \|\boldsymbol{\psi}\|_2^2 = P_2, \tag{16b}$$

where $\mathbf{q}^t = \left(\lambda_{\max}\mathbf{I} - \tilde{\boldsymbol{\Lambda}}\right)\boldsymbol{\psi}^t + \tilde{\boldsymbol{\beta}}$. The equal sign in (16b) ensures the maximum utilization of total power to improve the system performance. Thus, the solution to $\mathcal{P}_3$ can be given by

$$\boldsymbol{\psi}^{t+1} = \frac{\mathbf{q}^t}{\|\mathbf{q}^t\|_2}\sqrt{P_2}. \tag{17}$$

After the iterations, we can approach the optimal solution of $\boldsymbol{\psi}$. Then, the optimal precoding design $\mathbf{w}^{\mathrm{opt}}$ and $\boldsymbol{\Theta}^{\mathrm{opt}}$, i.e., the solution to $\mathcal{P}_1$, is obtained based on $\boldsymbol{\psi}$. Details of our proposed MM-based scheme is shown in **Algorithm 1**.

Algorithm 1 MM-based multi-beam design scheme

Require: Target multiple beams $\mathbf{v}$; DFT codebook $\mathbf{C}$; BS-RIS channel $\mathbf{G}$; BS power P_1; RIS power P_2; number of iterations T;
Ensure: Transmit precoding $\mathbf{w}^{\mathrm{opt}}$; RIS precoding $\boldsymbol{\Theta}^{\mathrm{opt}}$;
1: Randomly initialize the precoding vector $\boldsymbol{\theta} = \boldsymbol{\theta}^1$;
2: **for** $t \in \{1, 2, \cdots, T\}$ **do**
3: Calculate $\boldsymbol{\psi}^t = \mathbf{B}^{\frac{1}{2}}\boldsymbol{\theta}^t$;
4: $\mathbf{q}^t = \left(\lambda_{\max}\mathbf{I} - \tilde{\boldsymbol{\Lambda}}\right)\boldsymbol{\psi}^t + \tilde{\boldsymbol{\beta}}$;
5: $\boldsymbol{\psi}^{t+1} = \frac{\mathbf{q}^t}{\|\mathbf{q}^t\|_2}\sqrt{P_2}$;
6: **end for**
7: $\boldsymbol{\psi}^{\mathrm{opt}} = \boldsymbol{\psi}^t$, $\boldsymbol{\theta}^{\mathrm{opt}} = \mathbf{B}^{-\frac{1}{2}}\boldsymbol{\psi}^{\mathrm{opt}}$;
8: $\mathbf{G} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^H$, $\mathbf{V}_{[:,1]} = \mathbf{V}_{[:,1]}$;
9: $\mathbf{w}^{\mathrm{opt}} = \frac{\mathbf{V}_{[:,1]}}{\|\mathbf{V}_{[:,1]}\|_2}\sqrt{P_1}$, $\boldsymbol{\Theta}^{\mathrm{opt}} = \mathrm{diag}(\boldsymbol{\theta}^{\mathrm{opt}})$.

3.3 Power Allocation

We discuss the power allocation among multiple users to design the amplitude of multiple beams in this subsection.

Based on the sequential beam training of K users through the DFT codebook, K spatial angles of users and the corresponding beam gain λ_k in each spatial angle grid can be obtained. Here we consider two scenarios to adaptively design the target multiple beams $\mathbf{v}$ by weighted superposition. Specifically, one scenario is aimed at maximizing the minimum user rate, the other is to maximize the sum-rate.

The optimization problem of maximizing the minimum user rate is given by

$$\max \min_i \ R_i = \log(1 + \frac{p_i \lambda_i}{\sigma_i^2}) \tag{18a}$$

$$\text{s.t.} \ \ C_1 : \sum_i p_i = P, \tag{18b}$$

where p_i and λ_i represent the allocated power and the equivalent channel gain for user i. The following **Lemma 1** provides the solution to the optimization problem.

Lemma 1. Assuming that the noise power at each user is the same. The optimization problem of maximizing the minimum user rate achieves the optimal performance, if and only if the received SNR of each user is the same, i.e., $p_i \propto \frac{1}{\lambda_i}$.

Proof: This proof can be completed through the proof by contradiction. ■

From **Lemma 1**, the minimum user rate can be achieved by weighted superposition-based multi-beam design. The beam gains in K angle grids can be superimposed with the weight $\omega_k = \frac{1}{\lambda_k}$ to design $\mathbf{v}$.

Besides, the sum-rate of K users can be maximized by applying the water-filling algorithm, wherein the target multiple beams $\mathbf{v}$ could be designed by superimposing beam gains with the weight $\omega_k = \lambda_k$.

3.4 Convergence and Complexity

Firstly, the convergence of our proposed MM-based scheme is analyzed. According to [9], the proxy function $g(\boldsymbol{\psi}|\boldsymbol{\psi}^t)$ in (14) satisfies the features in (13). Thus, we have

$$f(\boldsymbol{\psi}^{t+1}) \leqslant g(\boldsymbol{\psi}^{t+1}|\boldsymbol{\psi}^t) \leqslant g(\boldsymbol{\psi}^t|\boldsymbol{\psi}^t) = f(\boldsymbol{\psi}^t). \tag{19}$$

It can be revealed that, the sequence $f(\boldsymbol{\psi}^t)$ is monotonically decreasing during iterations. Thus, the MM algorithm-based scheme will converge to the optimal solution of $\mathcal{P}_1$.

In addition, we discuss the computational complexity of the MM-based scheme, which mainly includes two aspects. First, the complexity of calculating the maximum eigenvalue of $\tilde{\mathbf{\Lambda}}$ could be given by $\mathcal{O}(N^3)$. Second, $\boldsymbol{\psi}^t$ and $\mathbf{q}^t$ are updated in each iteration, the complexity of which is given by $\mathcal{O}(N^2)$. Notice that T represent the number of iterations. Thus, the computational complexity of our proposed MM-based scheme is expressed as $\mathcal{O}(N^3) + \mathcal{O}(TN^2)$.

As for the DFT-based scheme, since the weighted superposition of DFT codewords provides the closed-form solution, the computational complexity is $\mathcal{O}(1)$.

4 Simulation Results

Simulation results are provided in this section to validate the performance of proposed schemes.

If not specifically illustrated, the number of RIS elements can be given by $N = 256$ with $N_1 = 16$ and $N_2 = 16$. The number of paths is given by $L = 4$. The path delay τ_l, path gain g_l, and other parameters are generated as follows: $\tau_l \sim \mathcal{U}\left(0, 20 \times 10^{-9}\right)$, $g_l \sim \mathcal{CN}\left(0, 1\right)$, $\alpha_1 = 2$, $\alpha_2 = 2$, $d_1 = 10$ m, $d_2 = 250$ m, $C = -60$ dB, $N_t = 4$, $K = 4$, and $f = 5$ GHz. The noise power is given by $\sigma^2 = \sigma_k^2 = -110$ dBm. The number of iterations is selected as $T = 5$. Five hundred independent experiments are performed with randomly generated channels.

We present the performance of our proposed MM-based scheme and DFT-based scheme for active RIS to design multiple beams. The DFT-based scheme for passive RIS serves as the baseline, which only extracts the phase information after the weighted superposition of DFT codebooks. Note that the power consumption is set to be the same. Specifically, The total power of the RIS-assisted communication system $P = P_1 + P_2$ ranges from 5 W to 30 W. As for the active RIS-assisted communication system,

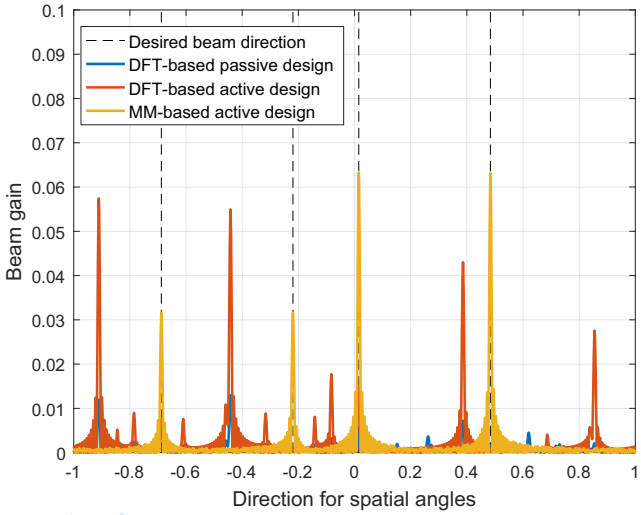

Figure 2 The designed multiple beams.

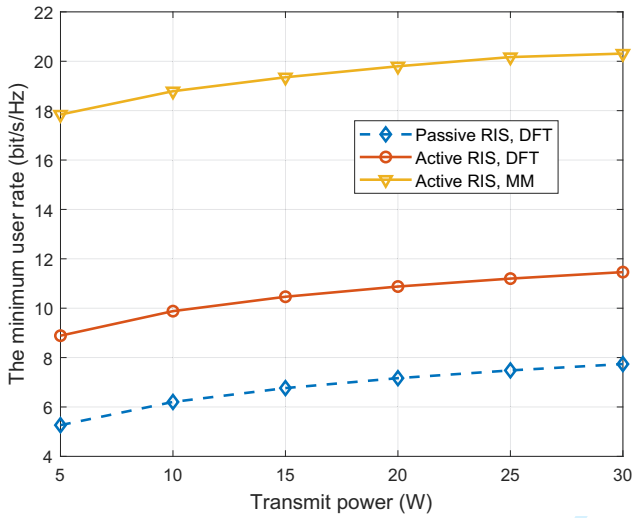

Figure 3 Minimum user rate against transmit power.

the radiation power of active RIS is set to one thousandth of the total power expressed as $P_2 = \frac{P}{1000}$. For comparison, the transmit power of BS in the passive RIS-assisted system is equal to the total power expressed as $P_1 = P$.

Firstly, the beam gain of acquired multiple beams by three schemes versus the direction for spatial angles are presented in Fig. 2. To simplify the illustration, we adopt the number of RIS elements as $N_1 = 256$ and $N_2 = 1$. The total power is set to 10 W. The ratio of allocated powers for multiple beams towards $K = 4$ users are set to $1 : 1 : 2 : 2$. It can be observed that the passive RIS can only generate beams with very low power due to the effect of multiplicative fading. Besides, the DFT-based schemes for both active RIS and passive RIS cannot generate beams towards desired directions with desired gains. The reason is that the DFT-based schemes do not take reflection of incident signals from the BS into account, which only aims at RIS precoding. Thus, the generated beams will point to other directions in a practical RIS-assisted system. By contrast, the generated multiple beams based on the MM algorithm point to target directions with predetermined beam gains, which validates the effectiveness.

Then, Fig. 3 shows the minimum user rate against the transmit power. It is noted that the rate performance of active RIS always outperforms that of passive RIS. It is because that active RIS can overcome the effect of multiplicative fading by signal amplification. Besides, for active RIS, the MM-based scheme improves more than 50% minimum user rate compared with the DFT-based scheme. The reason is that the DFT-based scheme cannot generate multiple beams towards desired directions. Besides,

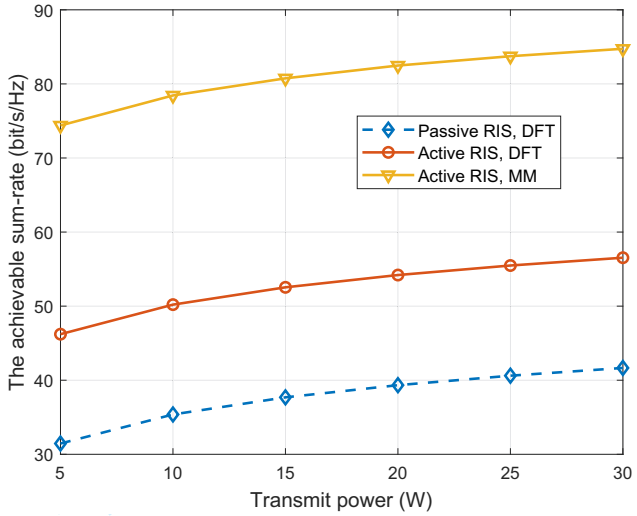

Figure 4 Achievable sum-rate against transmit power.

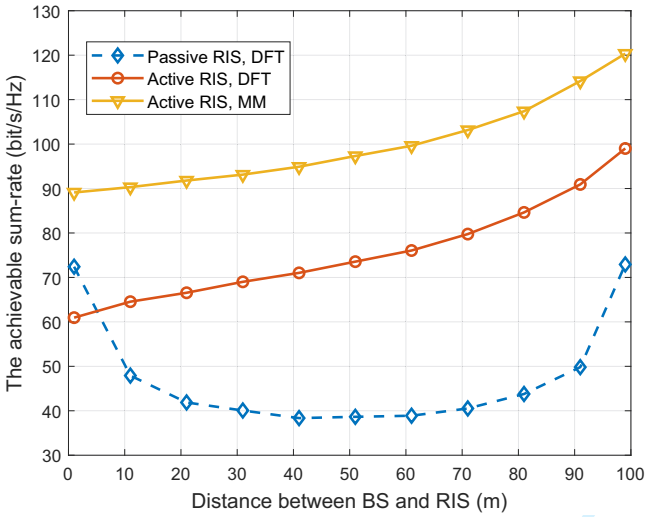

Figure 5 Achievable sum-rate against the distance.

the DFT-based scheme is unable to tackle with the amplification of thermal noise, which leads to serious performance loss. In contrast, due to the better balance of amplifying signals and thermal noise, the MM-based scheme is capable of achieving a much higher minimum rate, which effectively improves user fairness.

Moreover, the achievable sum-rate against the transmit power is presented in Fig. 4. It is revealed that with the increased transmit power, the system performance of all three schemes gradually increases, which is due to the improvement of SNR at users. One can also observe that the active RIS could achieve better performance than the passive RIS. Similarly, compared with the DFT-based scheme, the sum-rate achieved by the MM-based scheme increases by above 50% for active RIS. Therefore, we can conclude that the MM-based scheme is able to realize improved user rate performance compared with the DFT-based schemes, no matter for active RIS or passive RIS.

In addition, the influence of the location of active RIS is investigated. The total power is set to 10 W. The sum of the distance from the RIS to BS and that from the RIS to users is set to 100 m, where the location of RIS changes. We provide the sum-rate performance against the distance between the RIS and BS in Fig. 5. As for passive RIS-assisted systems, we can find that the sum-rate performance decreases at first and then increases with the increase of the distance between the RIS and BS. This is because the multiplicative fading effect in passive RIS-assisted systems can be alleviated with RIS located near either users or the BS. By contrast, the sum-rate performance of active RIS based on the DFT-based and MM-

based scheme increases with the increased distance between the RIS and BS. There are two main reasons. The first one is that when the active RIS is deployed closer to users, the amplification factors are bigger to compensate for the amplification noise, which significantly improves the signal power and overcomes the multiplicative fading. The second one is that as the distance between the RIS and users becomes smaller, the influence of incomplete CSI is gradually mitigated. Thus, one can observe that when the RIS is deployed very close to the BS where the advantages of active RIS are not fully utilized, the DFT-based scheme for passive RIS outperforms that for active RIS. Besides, the MM-based scheme always achieves better sum-rate compared with the DFT-based scheme due to beam alignment. Therefore, the MM-based scheme can provide a practical multi-beam design for active RIS with incomplete CSI.

5 Conclusions

In this paper, we proposed an adaptive multi-beam design framework based on active RIS with incomplete CSI. We first formulated the multi-beam design problem with power constraints and investigated the power allocation among multiple users. Then, a DFT-based codebook design scheme with low complexity was proposed to generate desired multiple beams. To compensate for the beam gain loss caused by the beam misalignment of the DFT-based scheme, we further proposed a MM algorithm-based scheme to improve the performance at the cost of increased complexity. Simulation results demonstrated that both the DFT-based and MM-based scheme for active RIS proposed in this paper can realize higher beam gain compared with passive RIS. Besides, the proposed MM-based scheme can generate multiple beams towards desired directions with expected beam gains, which thus achieved higher user rate performance compared with the DFT-based scheme. Therefore, the proposed adaptive multi-beam design framework based on active RIS in this paper provided an effective solution to multi-beam design for RIS. More advanced precoding algorithms to solve the optimization problem with lower complexity and improved performance are left for future works.

Acknowledgements This work was supported in part by the National Key Research and Development Program of China under Grant 2020YFB1807201, and in part by the European Commission through the H2020-MSCA-ITN META WIRELESS Research Project under Grant 956256.

References

1 C. Liu, F. Yang, S. Xu, and M. Li. Reconfigurable metasurface: A systematic categorization and recent advances. arXiv preprint arXiv:2301.00593, 2023.
2 C. Pan, H. Ren, K. Wang, J. F. Kolb, M. Elkashlan, M. Chen, M. Di Renzo, Y. Hao, J. Wang, A. L. Swindlehurst, X. You, and L. Hanzo. Reconfigurable intelligent surfaces for 6G systems: Principles, applications, and research directions. IEEE Commun. Mag., 2021, 59: 14–20.
3 L. Dai, B. Wang, M. Wang, X. Yang, J. Tan, S. Bi, S. Xu, F. Yang, Z. Chen, M. D. Renzo, C. Chae, and L. Hanzo. Reconfigurable intelligent surface-based wireless communications: Antenna design, prototyping, and experimental results. IEEE Access, 2020, 8: 45913– 45923.
4 R. Liu, Q. Wu, M. Di Renzo, and Y. Yuan. A path to smart radio environments: An industrial viewpoint on reconfigurable intelligent surfaces. IEEE Wireless Commun., 2022, 29: 202–208.
5 E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M.-S. Alouini, and R. Zhang. Wireless communications through reconfigurable intelligent surfaces. IEEE Access, 2019, 7: 116753–116773.
6 P. Wang, J. Fang, X. Yuan, Z. Chen, and H. Li. Intelligent reflecting surface-assisted millimeter wave communications: Joint active and passive precoding design. IEEE Trans. Veh. Technol., 2020, 69: 14960–14973.
7 Q. Wu and R. Zhang. Beamforming optimization for wireless network aided by intelligent reflecting surface with discrete phase shifts. IEEE Trans. Commun., 2020, 68: 1838–1851.
8 Y. Liu, X. Liu, X. Mu, T. Hou, J. Xu, M. Di Renzo, and N. Al-Dhahir. Reconfigurable intelligent surfaces: Principles and opportunities. IEEE Commun. Surv. Tutorials, 2021, 23: 1546–1577.
9 C. Pan, H. Ren, K. Wang, W. Xu, M. Elkashlan, A. Nallanathan, and L. Hanzo. Multicell MIMO communications relying on intelligent reflecting surfaces. IEEE Trans. Wireless Commun., 2020, 19: 5218–5233.
10 C. Qiu, Q. Wu, M. Hua, X. Guan, and Y. Wu. Achieving multi-beam gain in intelligent reflecting surface assisted wireless energy transfer. IEEE Trans. Veh. Technol., 2022.
11 R. Su, L. Dai, J. Tan, M. Hao, and R. MacKenzie. Capacity enhancement for reconfigurable intelligent surface-aided wireless network: From regular array to irregular array. IEEE Trans. Veh. Technol., 2023.
12 H. Alwazani, A. Kammoun, A. Chaaban, M. Debbah, and M.-S. Alouini. Intelligent reflecting surface-assisted multi-user MISO communication: Channel estimation and beamforming design. IEEE Open J. Commun. Soc., 2020, 1: 661–680.
13 C. Pfeffer, R. Feger, C. Wagner, and A. Stelzer. FMCW MIMO radar system for frequency-division multiple TX-beamforming. IEEE Trans. Microw. Theory Techn., 2013, 61: 4262-4274.
14 Z. Xiao, T. He, P. Xia, and X.-G. Xia. Hierarchical codebook design for beamforming training in millimeter-wave communication. IEEE Trans. Wireless Commun., 2016, 15: 3380–3392.
15 S. Noh, M. D. Zoltowski, and D. J. Love. Multi-resolution codebook and adaptive beamforming sequence design for millimeter wave beam alignment. IEEE Trans. Wireless Commun., 2017, 16: 5689– 5701.
16 K. Chen, C. Qi, G. Y. Li. Two-step codeword design for millimeter wave massive MIMO systems with quantized phase shifters. IEEE Trans. Signal Process., 2020, 68: 170-180.

17 Z. Kang, C. You, and R. Zhang. Active-IRS-aided wireless communication: Fundamentals, designs and open issues. arXiv preprint arXiv:2301.04311, 2023.
18 M. Najafi, V. Jamali, R. Schober, and H. V. Poor. Physics-based modeling and scalable optimization of large intelligent reflecting surfaces. IEEE Trans. Commun., 2021, 69: 2673–2691.
19 Z. Zhang, L. Dai, X. Chen, C. Liu, F. Yang, R. Schober, and H. V. Poor. Active RIS vs. passive RIS: Which will prevail in 6G?. IEEE Trans. Commun., 2022.
20 J. Loncar, Z. Sipus, and S. Hrabar. Ultrathin active polarization-selective metasurface at X-band frequencies. Physical Review B, 2019, 100: 075131.
21 J. Bousquet, S. Magierowski, and G. G. Messier. A 4-GHz active scatterer in 130-nm CMOS for phase sweep amplify-and-forward. IEEE Trans. Circuits Syst. I, 2012, 59: 529–540.
22 K. K. Kishor and S. V. Hum. An amplifying reconfigurable reflectarray antenna. IEEE Trans. Antennas Propag., 2012, 60: 197–205.
23 R. Long, Y.-C. Liang, Y. Pei, and E. G. Larsson. Active reconfigurable intelligent surface-aided wireless communications. IEEE Trans. Wireless Commun., 2021, 20: 4962–4975.
24 O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath. Spatially sparse precoding in millimeter wave MIMO systems. IEEE Trans. Wireless Commun., 2014, 13: 1499–1513.
25 Y. Sun, P. Babu, and D. P. Palomar. Majorization-minimization algorithms in signal processing, communications, and machine learning. IEEE Trans. Signal Process., 2017, 65: 794–816.