Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Beamforming design for RIS-aided amplify-and-forward relay networks*

Xuehui WANG[†1], Feng SHU[†‡1,2], Riqing CHEN[3], Peng ZHANG[1], Qi ZHANG[1], Guiyang XIA[4], Weiping SHI[2], Jiangzhou WANG[5]

[1]*School of Information and Communication Engineering, Hainan University, Haikou 570228, China*

[2]*School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China*

[3]*Digital Fujian Institute of Big Data for Agriculture, Fujian Agriculture and Forestry University, Fuzhou 350002, China*

[4]*Institute of Intelligent Agriculture, Anhui Agricultural University, Hefei 230036, China*

[5]*School of Engineering, University of Kent, Canterbury CT2 7NT, UK*

[†]E-mail: wangxuehui0503@163.com; shufeng0101@163.com

Received Feb. 27, 2023; Revision accepted Aug. 20, 2023; Crosschecked Dec. 1, 2023

Abstract: The use of a reconfigurable intelligent surface (RIS) in the enhancement of the rate performance is considered to involve the limitation of the RIS being a passive reflector. To address this issue, we propose a RIS-aided amplify-and-forward (AF) relay network in this paper. By jointly optimizing the beamforming matrix at AF relay and the phase-shift matrices at RIS, two schemes are put forward to address a maximizing signal-to-noise ratio (SNR) problem. First, aiming at achieving a high rate, a high-performance alternating optimization (AO) method based on Charnes–Cooper transformation and semidefinite programming (CCT-SDP) is proposed, where the optimization problem is decomposed into three subproblems solved using CCT-SDP, and rank-one solutions can be recovered using Gaussian randomization. However, the optimization variables in the CCT-SDP method are matrices, leading to extremely high complexity. To reduce the complexity, a low-complexity AO scheme based on Dinkelbachs transformation and successive convex approximation (DT-SCA) is proposed, where the variables are represented in vector form, and the three decoupling subproblems are solved using DT-SCA. Simulation results verify that compared to three benchmarks (i.e., a RIS-assisted AF relay network with random phase, an AF relay network without RIS, and a RIS-aided network without AF relay), the proposed CCT-SDP and DT-SCA schemes can harvest better rate performance. Furthermore, it is revealed that the rate of the low-complexity DT-SCA method is close to that of the CCT-SDP method.

Key words: Reconfigurable intelligent surface (RIS); Amplify-and-forward (AF) relay; Beamforming; Phase shift; Semidefinite programming; Successive convex approximation

https://doi.org/10.1631/FITEE.2300118 **CLC number:** TN92

1 Introduction

With the exponential growth of smart devices and data traffic, a large and complex communication network has been formed. As a corollary, the demands that are placed in terms of data rate, connectivity, and reliability continue to become more stringent (An et al., 2023b), and thus an urgent need has emerged for the development of an innovative, efficient, highly reliable, and resource-saving

‡ Corresponding author

* Project supported by the National Natural Science Foundation of China (Nos. U22A2002 and 62071234), the Hainan Province Science and Technology Special Fund, China (No. ZDKJ2021022), and the Scientific Research Fund Project of Hainan University, China (No. KYQD(ZR)-21008)

ORCID: Xuehui WANG, https://orcid.org/0000-0002-1090-5650; Feng SHU, https://orcid.org/0000-0003-0073-1965; Jiangzhou WANG, https://orcid.org/0000-0003-0881-3594

wireless network that can be used in the future, when these demands can be expected to greatly intensify. Meanwhile, two major practical limitations are immediately identifiable when discussing the challenges that are faced in the design and implementation of signal propagation in wireless networks. The first is that the communication process involves the consumption of a vast amount of energy, and the second is the nature of uncontrollability characterizing the transmission channel, which causes serious signal attenuation. Due to its ability to create friendly multipaths by adjusting the reflection coefficient of each passive element (Shu et al., 2021b), reconfigurable intelligent surface (RIS) technology is considered to have the potential to improve the communication quality (Chen et al., 2022; Tian et al., 2022). Passive RIS is characterized by the advantages of low cost and low power consumption, as well as relative ease of deployment (Wu and Zhang, 2019). The proper deployment of a RIS in the communication environment, especially when the direct channel link is blocked, enables the receipt of the transmitted signal by the desired receiver over an area of extended coverage.

Recently, RIS has become a research hotspot and has drawn much attention. RIS has been widely researched in different novel communication applications, and as some immediately familiar examples we may mention secure communication (Hong et al., 2021; Jiang WH et al., 2021; Shi et al., 2021a), covert communication (Zhou XB et al., 2022), satellite communication (Lee et al., 2021; Zheng et al., 2022), simultaneous wireless information and power transfer (SWIPT) (Shi et al., 2021b), vehicle networks (Jiang W and Schotten, 2022), and spatial modulation (Shu et al., 2021a, 2022). A RIS-aided multiple-input multiple-output secure network was proposed in Jiang WH et al. (2021), where the transmit beamforming and RIS phase-shift matrix were jointly optimized using an alternating optimization (AO) algorithm for maximum sum secrecy rate. In Zheng et al. (2022), a RIS-assisted low-Earth orbit satellite system was considered. A flexible beamforming could be achieved by controlling phase shifts of RIS in a programmable way, so that the time-varying channel could be cost-effectively handled. In Jiang W and Schotten (2022), a RIS was applied to a traditional vehicular network, where a method of maximizing instantaneous signal-to-noise ratio (SNR) was proposed to jointly optimize the transmit beamforming and RIS reflection-coefficient matrix. Evaluation of the sum spectral efficiency revealed that the capacity of a moving vehicular network could be enhanced by introducing a RIS. Guo et al. (2020) presented a RIS-aided multiuser multiple-input single-output downlink communication system. Aiming at maximizing the weighted sum-rate (WSR) of all users, a low-complexity method based on fractional programming was used to jointly design transmit beamforming and RIS phase in the case of perfect channel state information (CSI). Then the proposed algorithm was extended to the case of imperfect CSI, and stochastic successive convex approximation (SCA) was adopted to achieve maximum WSR.

Although the RIS and relay technologies are similar in forwarding signal, there are essential differences between them. RIS is generally considered as a low-cost and low-power-consumption green reflector whose exclusive task is to reflect the incident signal; however, so far as signal-processing ability is concerned, it has none, whereas in the case of relay, this ability is strong, and thus the latter can actively process and forward the received signal (Shu et al., 2014). It has been proved that RIS with massive reflecting elements can match up to a decode-and-forward (DF) relay in terms of rate performance (Björnson et al., 2020), while the rate achieved by RIS is not always improved commensurate with an increase in the number of elements (Wang MX et al., 2022). Additionally, as a result of high wind resistance, the deployment of RIS with large metasurface is more challenging in conventional communication scenarios. Further, the maintenance cost may be increased, as well.

At present, there are some studies in the literature researching a hybrid network encompassing both RIS and relay. These studies described the outcomes obtained from combining the RIS and relay approaches, as well as the benefits derivable from such an amalgamation in terms of rate enhancement (Bie et al., 2022; Khalid et al., 2022; Wang XH et al., 2023), coverage extension (Abdullah et al., 2022), spectral efficiency and energy efficiency improvement (Li et al., 2022; Obeed and Chaaban, 2022), and channel estimation (Sun et al., 2023) of the communication network. A RIS-aided two-way DF relay wireless network was proposed in Wang XH et al. (2023), where power allocation parameters of two

users and DF relay were optimized using the SCA method and the maximizing determinant method to obtain the maximum rate. In the hybrid network, the position of RIS plays a key role in the rate maximization problem. Bie et al. (2022) proposed a method of optimizing RIS deployment to enhance the rate, where the closed-form expressions of optimal RIS deployment were derived. Then it was verified that when RIS was close to relay, the maximum rate could be obtained. With the aim of extended coverage, three different double-RIS networks (i.e., a double-RIS network without relay, a relay-aided double-RIS network, and a double-relay-aided double-RIS network) were proposed in Abdullah et al. (2022), where the AO and majorization minimization (MM) schemes were designed to optimize the double-RIS phase. Furthermore, it was demonstrated that the double-relay-aided double-RIS network can achieve a higher rate than the two other networks. Obeed and Chaaban (2022) proposed a multiuser downlink network aided by relay and RIS. Their research, seeking to solve the optimization problem of maximizing the energy efficiency, consisted of the individual applications of singular value decomposition (SVD), semidefinite programming (SDP), and function approximations towards the optimization of the beamforming matrices of transmitter and relay, and RIS phase shifts. In terms of channel estimation, a pilot pattern was proposed in Sun et al. (2023) to separate direct and cascaded channels for the channel estimation of a discrete-phase-shifter RIS-assisted two-way relay network, and the performance loss of finite quantization was derived.

Given the aforementioned studies, it is evident that the combination networks consisting of relay and RIS, where these can work together in coordination, have drawn researchers' attention. The hybrid network can not only make full use of the advantages of relay and RIS, but also maintain a balance between cost and rate performance (Wang XH et al., 2022). Compared to a RIS-aided network, a hybrid RIS and DF relay network can achieve the same rate improvement with a smaller number of RIS elements, and it thereby becomes possible to also achieve savings in terms of the number of RIS units that need to be deployed (Abdullah et al., 2020). Moreover, when the channels are of poor quality, and especially when the three direct links from source to destination, from source to relay, and from relay to destination are weak, or are completely unavailable because of an obstacle, employing the hybrid RIS and relay network would still be viable; that is to say, the communication between the source and destination can still be proceeded with based on the joint assistance of RIS and relay in such a case, which is a testament to this method's robustness and aptness for use in such realistic abominable communication scenarios. Meanwhile, Yildirim et al. (2021) indicated that adding a relay in a RIS-assisted network is helpful for the enhancement of the rate performance.

To our knowledge, studies in the literature pertaining to hybrid networks focus on DF relay and RIS, and there are few studies dealing with a hybrid network consisting of amplify-and-forward (AF) relay and RIS. Considering the benefits of the hybrid relay and RIS network, the present study discusses a RIS-aided AF relay network. To obtain the needed rate performance improvement, two AO methods are proposed to jointly design beamforming for maximum SNR. Our contributions are summarized as follows:

1. A RIS-aided AF relay network is proposed, in which a non-convex optimization problem of maximizing SNR is formulated. Since the optimization variables (i.e., the beamforming matrix at AF relay and the reflection-coefficient matrices at RIS) are coupled, a high-performance AO method based on Charnes–Cooper transformation and semidefinite programming (CCT-SDP) is proposed. To solve such a non-convex problem, vectorization and trace function are applied to convert the problem into a fractional or linear optimization problem. After dropping rank-one constraints, the optimization problem can be addressed using the CCT-SDP algorithm. Finally, rank-one solutions are recovered using the Gaussian randomization method. Simulation results demonstrate that the proposed CCT-SDP scheme exhibits a significant rate improvement over three benchmarks (i.e., a RIS-assisted AF relay network with random phase, an AF relay network without RIS, and a RIS-aided network without AF relay).

2. The optimization variables in the proposed CCT-SDP method are matrices, involving an extremely high computational complexity, i.e., $\mathcal{O}(M^{13}+N^{6.5})$, where M and N are the numbers of AF relay antennas and RIS units, respectively. To reduce the complexity, a low-complexity maximizing SNR scheme based on Dinkelbachs transformation

and SCA (DT-SCA) is presented. By performing DT and first-order Taylor approximation, the non-convex optimization problem is transformed into a convex problem. Since the optimization variables are vectors, the complexity of the DT-SCA method (i.e., $\mathcal{O}(M^6+N^{3.5})$) is much lower than that of the CCT-SDP method. Moreover, from the simulation results, it is found that the low-complexity DT-SCA scheme outperforms the three benchmarks. Additionally, its rate is approximate to that of the CCT-SDP method.

Notations: Throughout this paper, we represent scalars, vectors, and matrices using letters of lower case, bold lower case, and bold upper case, respectively. The conjugate, transpose, and conjugate transpose of matrix $\boldsymbol{A}$ are expressed as $\boldsymbol{A}^*$, $\boldsymbol{A}^{\mathrm{T}}$, and $\boldsymbol{A}^{\mathrm{H}}$, respectively. The signs $\mathbb{E}\{\cdot\}$, $|\cdot|$, $\|\cdot\|$, and $\angle(\cdot)$ represent the expectation operation, the modulus of a scalar, 2-norm, and the phase of a complex number, respectively. $\mathrm{tr}(\boldsymbol{A})$ and $\lambda_{\max}(\boldsymbol{A})$ represent the trace and the largest eigenvalue of matrix $\boldsymbol{A}$, respectively. $\boldsymbol{I}_n$ stands for the n-dimensional identity matrix, and $\otimes$ denotes the Kronecker product.

2 System model

2.1 Signal model

Fig. 1 depicts a RIS-aided AF relay network, where there is a half-duplex AF relay with M transmit antennas, a RIS with N reflecting elements, a single-antenna source (S), and a single-antenna destination (D). Here, the network is operated in half-duplex mode. It is assumed that there is no direct path from S to D, and S transmits the signal toward D with the help of RIS and AF relay. Any signal that is reflected by the RIS twice or more can be ignored because of the serious path loss, and we consider that the RIS reflects the signal once in each time slot.

In the first time slot, the received signal at AF relay is written as

$$\boldsymbol{y}_{\mathrm{r}}=\sqrt{P_{\mathrm{s}}}(\boldsymbol{h}_{\mathrm{sr}}+\boldsymbol{H}_{\mathrm{ir}}\boldsymbol{\Theta}_1\boldsymbol{h}_{\mathrm{si}})x+\boldsymbol{n}_{\mathrm{r}}, \tag{1}$$

where x is the signal transmitted from S with $\mathbb{E}[|x|^2]=1$, and P_{s} is the total transmission power. We assume that all channels follow Rayleigh fading, and $\boldsymbol{h}_{\mathrm{sr}}\in\mathbb{C}^{M\times 1}$, $\boldsymbol{h}_{\mathrm{si}}\in\mathbb{C}^{N\times 1}$, and $\boldsymbol{H}_{\mathrm{ir}}\in\mathbb{C}^{M\times N}$ are the channels from S to AF relay, from S to RIS, and from RIS to AF relay, respectively. $\boldsymbol{\Theta}_1=\mathrm{diag}(\mathrm{e}^{\mathrm{j}\theta_{11}},\mathrm{e}^{\mathrm{j}\theta_{12}},...,\mathrm{e}^{\mathrm{j}\theta_{1N}})$ is a diagonal

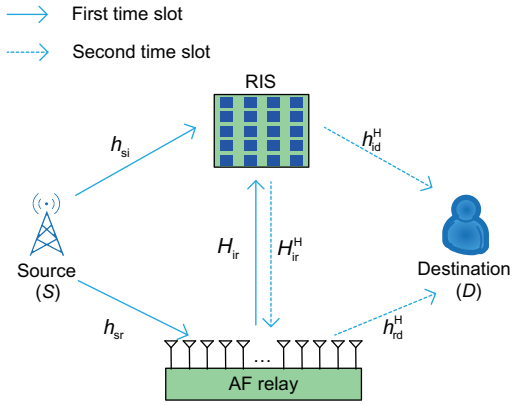

Fig. 1 A RIS-assisted AF relay wireless system

reflection-coefficient matrix of RIS in the first time slot, where $\theta_{1i}\in[0,2\pi)$ is the phase shift for the i^{th} RIS reflecting element. $\boldsymbol{n}_{\mathrm{r}}\in\mathbb{C}^{M\times 1}$ denotes the complex additive white Gaussian noise (AWGN) vector at AF relay with distribution $\boldsymbol{n}_{\mathrm{r}}\sim\mathcal{CN}(\boldsymbol{0},\sigma_{\mathrm{r}}^2\boldsymbol{I}_M)$. After receive and transmit beamforming operations are performed to AF relay, the signal transmitted at AF relay can be expressed as

$$\begin{aligned}\boldsymbol{y}_{\mathrm{t}}&=\boldsymbol{A}\boldsymbol{y}_{\mathrm{r}}\\&=\sqrt{P_{\mathrm{s}}}\boldsymbol{A}(\boldsymbol{h}_{\mathrm{sr}}+\boldsymbol{H}_{\mathrm{ir}}\boldsymbol{\Theta}_1\boldsymbol{h}_{\mathrm{si}})x+\boldsymbol{A}\boldsymbol{n}_{\mathrm{r}},\end{aligned} \tag{2}$$

where $\boldsymbol{A}\in\mathbb{C}^{M\times M}$ represents the beamforming matrix. The transmission power at AF relay can be written as

$$P_{\mathrm{r}}^{\mathrm{t}}=P_{\mathrm{s}}\|\boldsymbol{A}(\boldsymbol{h}_{\mathrm{sr}}+\boldsymbol{H}_{\mathrm{ir}}\boldsymbol{\Theta}_1\boldsymbol{h}_{\mathrm{si}})\|^2+\|\boldsymbol{A}\|_{\mathrm{F}}^2\sigma_{\mathrm{r}}^2\le P_{\mathrm{r}}, \tag{3}$$

where P_{r} is the transmission power budget at AF relay. In the second time slot, the received signal at D is given by

$$\begin{aligned}y_{\mathrm{d}}=&(\boldsymbol{h}_{\mathrm{rd}}^{\mathrm{H}}+\boldsymbol{h}_{\mathrm{id}}^{\mathrm{H}}\boldsymbol{\Theta}_2\boldsymbol{H}_{\mathrm{ir}}^{\mathrm{H}})\boldsymbol{y}_{\mathrm{t}}+n_{\mathrm{d}}\\=&\sqrt{P_s}(\boldsymbol{h}_{\mathrm{rd}}^{\mathrm{H}}+\boldsymbol{h}_{\mathrm{id}}^{\mathrm{H}}\boldsymbol{\Theta}_2\boldsymbol{H}_{\mathrm{ir}}^{\mathrm{H}})\boldsymbol{A}(\boldsymbol{h}_{\mathrm{sr}}+\boldsymbol{H}_{\mathrm{ir}}\boldsymbol{\Theta}_1\boldsymbol{h}_{\mathrm{si}})x\\&+(\boldsymbol{h}_{\mathrm{rd}}^{\mathrm{H}}+\boldsymbol{h}_{\mathrm{id}}^{\mathrm{H}}\boldsymbol{\Theta}_2\boldsymbol{H}_{\mathrm{ir}}^{\mathrm{H}})\boldsymbol{A}\boldsymbol{n}_{\mathrm{r}}+n_{\mathrm{d}},\end{aligned} \tag{4}$$

where channels $\boldsymbol{h}_{\mathrm{rd}}^{\mathrm{H}}\in\mathbb{C}^{1\times M}$, $\boldsymbol{h}_{\mathrm{id}}^{\mathrm{H}}\in\mathbb{C}^{1\times N}$, and $\boldsymbol{H}_{\mathrm{ir}}^{\mathrm{H}}\in\mathbb{C}^{N\times M}$ denote the links AF–D, RIS–D, and AF relay–RIS, respectively. The diagonal reflection-coefficient matrix of RIS in the second time slot is expressed as $\boldsymbol{\Theta}_2=\mathrm{diag}(\mathrm{e}^{\mathrm{j}\theta_{21}},\mathrm{e}^{\mathrm{j}\theta_{22}},...,\mathrm{e}^{\mathrm{j}\theta_{2N}})$, where $\theta_{2i}\in[0,2\pi)$ is the phase shift for the i^{th} RIS reflecting element. n_{d} is the complex AWGN at D with distribution $n_{\mathrm{d}}\sim\mathcal{CN}(0,\sigma_{\mathrm{d}}^2)$. It is assumed that $\sigma_{\mathrm{r}}^2=\sigma_{\mathrm{d}}^2=\sigma^2$ and $\gamma_{\mathrm{s}}=\frac{P_{\mathrm{s}}}{\sigma^2}$. The achievable rate can then be defined as

$$R=\frac{1}{2}\log_2(1+\mathrm{SNR}), \tag{5}$$

where

$$\mathrm{SNR}=\frac{\gamma_{\mathrm{s}}|(\boldsymbol{h}_{\mathrm{rd}}^{\mathrm{H}}+\boldsymbol{h}_{\mathrm{id}}^{\mathrm{H}}\boldsymbol{\Theta}_2\boldsymbol{H}_{\mathrm{ir}}^{\mathrm{H}})\boldsymbol{A}(\boldsymbol{h}_{\mathrm{sr}}+\boldsymbol{H}_{\mathrm{ir}}\boldsymbol{\Theta}_1\boldsymbol{h}_{\mathrm{si}})|^2}{\|(\boldsymbol{h}_{\mathrm{rd}}^{\mathrm{H}}+\boldsymbol{h}_{\mathrm{id}}^{\mathrm{H}}\boldsymbol{\Theta}_2\boldsymbol{H}_{\mathrm{ir}}^{\mathrm{H}})\boldsymbol{A}\|^2+1}. \tag{6}$$

2.2 Discussion on CSI acquisition

Indeed, CSI is vital for the joint optimization of the beamforming matrix $\boldsymbol{A}$ and the reflection-coefficient matrices $\boldsymbol{\Theta}_1$ and $\boldsymbol{\Theta}_2$. Since channel reciprocity holds for each channel, the acquisition of the uplink or downlink channel coefficient is mainly considered (An et al., 2023a). The CSI for the RIS-aided AF relay network can be estimated via, e.g., the parallel factor based method (Wei et al., 2021), the anchor-aided method (Guan et al., 2022), the Bayesian learning based method (Yang et al., 2023), and the novel least-squares based method (Sun et al., 2023).

2.3 Problem formulation

To enhance the system rate performance, it is necessary to maximize SNR. The corresponding optimization problem can be formulated as

$$\max_{\boldsymbol{\Theta}_1,\boldsymbol{\Theta}_2,\boldsymbol{A}}\ \frac{\gamma_{\mathrm{s}}|(\boldsymbol{h}_{\mathrm{rd}}^{\mathrm{H}}+\boldsymbol{h}_{\mathrm{id}}^{\mathrm{H}}\boldsymbol{\Theta}_2\boldsymbol{H}_{\mathrm{ir}}^{\mathrm{H}})\boldsymbol{A}(\boldsymbol{h}_{\mathrm{sr}}+\boldsymbol{H}_{\mathrm{ir}}\boldsymbol{\Theta}_1\boldsymbol{h}_{\mathrm{si}})|^2}{\|(\boldsymbol{h}_{\mathrm{rd}}^{\mathrm{H}}+\boldsymbol{h}_{\mathrm{id}}^{\mathrm{H}}\boldsymbol{\Theta}_2\boldsymbol{H}_{\mathrm{ir}}^{\mathrm{H}})\boldsymbol{A}\|^2+1} \tag{7a}$$

$$\text{s.t. } |\Theta_1(i,i)|=1,\ |\Theta_2(i,i)|=1, \tag{7b}$$

$$\gamma_{\mathrm{s}}\|\boldsymbol{A}(\boldsymbol{h}_{\mathrm{sr}}+\boldsymbol{H}_{\mathrm{ir}}\boldsymbol{\Theta}_1\boldsymbol{h}_{\mathrm{si}})\|^2+\|\boldsymbol{A}\|_{\mathrm{F}}^2\le\gamma_{\mathrm{r}}, \tag{7c}$$

where $\gamma_{\mathrm{r}}=P_{\mathrm{r}}/\sigma^2$. By defining $\boldsymbol{u}_1=[\mathrm{e}^{\mathrm{j}\theta_{11}},\mathrm{e}^{\mathrm{j}\theta_{12}},...,\mathrm{e}^{\mathrm{j}\theta_{1N}}]^{\mathrm{T}}$, $\boldsymbol{u}_2=[\mathrm{e}^{\mathrm{j}\theta_{21}},\mathrm{e}^{\mathrm{j}\theta_{22}},...,\mathrm{e}^{\mathrm{j}\theta_{2N}}]^{\mathrm{T}}$, $\boldsymbol{v}_1=[\boldsymbol{u}_1;1]$, $\boldsymbol{v}_2=[\boldsymbol{u}_2;1]$, $\boldsymbol{H}_{\mathrm{sir}}=[\boldsymbol{H}_{\mathrm{ir}}\mathrm{diag}(\boldsymbol{h}_{\mathrm{si}}),\boldsymbol{h}_{\mathrm{sr}}]$, and $\boldsymbol{H}_{\mathrm{rid}}=[\mathrm{diag}(\boldsymbol{h}_{\mathrm{id}}^{\mathrm{H}})\boldsymbol{H}_{\mathrm{ir}}^{\mathrm{H}};\boldsymbol{h}_{\mathrm{rd}}^{\mathrm{H}}]$, we have $\boldsymbol{h}_{\mathrm{sr}}+\boldsymbol{H}_{\mathrm{ir}}\boldsymbol{\Theta}_1\boldsymbol{h}_{\mathrm{si}}=\boldsymbol{H}_{\mathrm{sir}}\boldsymbol{v}_1$ and $\boldsymbol{h}_{\mathrm{rd}}^{\mathrm{H}}+\boldsymbol{h}_{\mathrm{id}}^{\mathrm{H}}\boldsymbol{\Theta}_2\boldsymbol{H}_{\mathrm{ir}}^{\mathrm{H}}=\boldsymbol{v}_2^{\mathrm{H}}\boldsymbol{H}_{\mathrm{rid}}$. Therefore, the above optimization problem can be converted to the following form:

$$\max_{\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{A}}\ \frac{\gamma_{\mathrm{s}}|\boldsymbol{v}_2^{\mathrm{H}}\boldsymbol{H}_{\mathrm{rid}}\boldsymbol{A}\boldsymbol{H}_{\mathrm{sir}}\boldsymbol{v}_1|^2}{\|\boldsymbol{v}_2^{\mathrm{H}}\boldsymbol{H}_{\mathrm{rid}}\boldsymbol{A}\|^2+1} \tag{8a}$$

$$\text{s.t. }\ |v_1(i)|=1,\ |v_2(i)|=1,\ \forall i=1,2,\cdots,N, \tag{8b}$$

$$v_1(N+1)=1,\ v_2(N+1)=1, \tag{8c}$$

$$\gamma_{\mathrm{s}}\|\boldsymbol{A}\boldsymbol{H}_{\mathrm{sir}}\boldsymbol{v}_1\|^2+\|\boldsymbol{A}\|_{\mathrm{F}}^2\le\gamma_{\mathrm{r}}, \tag{8d}$$

where the optimization variables $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, and matrix $\boldsymbol{A}$ are coupled. To solve problem (8), two schemes, namely the CCT-SDP-based AO method and the DT-SCA-based AO method, are proposed to optimize $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, and $\boldsymbol{A}$ for maximum SNR.

3 Proposed high-performance CCT-SDP-based AO method

Let us define $\boldsymbol{V}_1=\boldsymbol{v}_1\boldsymbol{v}_1^{\mathrm{H}}$ and $\boldsymbol{V}_2=\boldsymbol{v}_2\boldsymbol{v}_2^{\mathrm{H}}$. Then problem (8) can be rewritten as follows:

$$\max_{\boldsymbol{V}_1,\boldsymbol{V}_2,\boldsymbol{A}}\ \frac{\gamma_{\mathrm{s}}\mathrm{tr}(\boldsymbol{V}_1\boldsymbol{H}_{\mathrm{sir}}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{rid}}^{\mathrm{H}}\boldsymbol{V}_2\boldsymbol{H}_{\mathrm{rid}}\boldsymbol{A}\boldsymbol{H}_{\mathrm{sir}})}{\mathrm{tr}(\boldsymbol{V}_2\boldsymbol{H}_{\mathrm{rid}}\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{rid}}^{\mathrm{H}})+1} \tag{9a}$$

$$\text{s.t. } V_1(n,n)=1,V_2(n,n)=1,\forall n=1,2,\cdots,N+1, \tag{9b}$$

$$\gamma_{\mathrm{s}}\mathrm{tr}(\boldsymbol{V}_1\boldsymbol{H}_{\mathrm{sir}}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{H}_{\mathrm{sir}})+\|\boldsymbol{A}\|_{\mathrm{F}}^2\le\gamma_{\mathrm{r}}, \tag{9c}$$

$$\mathrm{rank}(\boldsymbol{V}_1)=1,\ \boldsymbol{V}_1\succeq\boldsymbol{0}, \tag{9d}$$

$$\mathrm{rank}(\boldsymbol{V}_2)=1,\ \boldsymbol{V}_2\succeq\boldsymbol{0}, \tag{9e}$$

where the problem is non-convex with rank-one constrains. Using the AO algorithm, problem (9) can be decoupled into three subproblems, as described in the forthcoming subsections.

3.1 Optimizing $\boldsymbol{A}$ given $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$

When $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$ are fixed, let us define $\boldsymbol{a}=\mathrm{vec}(\boldsymbol{A})\in\mathbb{C}^{M^2\times1}$, and problem (9) can be simplified as

$$\max_{\boldsymbol{a}}\ \frac{\boldsymbol{a}^{\mathrm{H}}\boldsymbol{B}\boldsymbol{a}}{\boldsymbol{a}^{\mathrm{H}}\boldsymbol{C}\boldsymbol{a}+1} \tag{10a}$$

$$\text{s.t.}\ \ \boldsymbol{a}^{\mathrm{H}}(\boldsymbol{D}+\boldsymbol{I}_{M^2})\boldsymbol{a}\le\gamma_{\mathrm{r}}, \tag{10b}$$

where matrix $\boldsymbol{B}=\gamma_{\mathrm{s}}(\boldsymbol{H}_{\mathrm{sir}}^{*}\boldsymbol{V}_1^{*}\boldsymbol{H}_{\mathrm{sir}}^{\mathrm{T}})\otimes(\boldsymbol{H}_{\mathrm{rid}}^{\mathrm{H}}\boldsymbol{V}_2\boldsymbol{H}_{\mathrm{rid}})$, $\boldsymbol{C}=\boldsymbol{I}_M\otimes(\boldsymbol{H}_{\mathrm{rid}}^{\mathrm{H}}\boldsymbol{V}_2\boldsymbol{H}_{\mathrm{rid}})$, and $\boldsymbol{D}=\gamma_{\mathrm{s}}(\boldsymbol{H}_{\mathrm{sir}}^{*}\boldsymbol{V}_1^{*}\boldsymbol{H}_{\mathrm{sir}}^{\mathrm{T}})\otimes\mathbf{I}_M$. Upon defining $\overline{\boldsymbol{A}}=\boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}\in\mathbb{C}^{M^2\times M^2}$ and removing the $\mathrm{rank}(\overline{\boldsymbol{A}})=1$ constraint, an SDR problem of Eq. (10) is written as

$$\max_{\overline{\boldsymbol{A}}}\ \frac{\mathrm{tr}(\boldsymbol{B}\overline{\boldsymbol{A}})}{\mathrm{tr}(\boldsymbol{C}\overline{\boldsymbol{A}})+1} \tag{11a}$$

$$\text{s.t.}\ \ \mathrm{tr}\{(\boldsymbol{D}+\boldsymbol{I}_{M^2})\overline{\boldsymbol{A}}\}\le\gamma_{\mathrm{r}},\ \ \overline{\boldsymbol{A}}\succeq\boldsymbol{0}. \tag{11b}$$

It is a linear fractional problem, which can be solved using CCT (Charnes and Cooper, 1962). We introduce the following transformation of variables:

$$\widetilde{\boldsymbol{A}}=m\overline{\boldsymbol{A}}, \tag{12}$$

where $m>0$ needs to meet

$$\mathrm{tr}(\boldsymbol{C}\widetilde{\boldsymbol{A}})+m=1. \tag{13}$$

We multiply the numerator and denominator of the objective function in (11a) and the constraint in (11b) by m, together with regarding Eq. (13)

as a constraint. Thereafter, problem (11) can be transformed into the following linear programming problem:

$$\max_{\widetilde{\boldsymbol{A}},m} \quad \mathrm{tr}(\boldsymbol{B}\tilde{\boldsymbol{A}}) \tag{14a}$$

$$\text{s.t.} \quad \mathrm{tr}(\boldsymbol{C}\widetilde{\boldsymbol{A}}) + m = 1, \ \ m > 0, \tag{14b}$$

$$\mathrm{tr}\{(\boldsymbol{D} + \boldsymbol{I}_{M^2})\widetilde{\boldsymbol{A}}\} \le m\gamma_{\mathrm{r}}, \ \ \widetilde{\boldsymbol{A}} \succeq \boldsymbol{0}. \tag{14c}$$

It is an SDP problem (Zhou X et al., 2019), which can be directly solved via CVX. While the constraint $\mathrm{rank}(\overline{\boldsymbol{A}}) = 1$ is ignored in the SDR problem, we apply the Gaussian randomization method to achieve a solution $\overline{\boldsymbol{A}}$ with $\mathrm{rank}(\overline{\boldsymbol{A}}) = 1$. Further, $\boldsymbol{a}$ can be obtained through the eigenvalue decomposition of $\overline{\boldsymbol{A}}$, and the AF relay beamforming matrix $\boldsymbol{A}$ is thereby achieved.

3.2 Optimizing $\boldsymbol{V_1}$ given $\boldsymbol{V_2}$ and $\boldsymbol{A}$

Upon fixing $\boldsymbol{V}_2$ and $\boldsymbol{A}$, problem (9) without considering $\mathrm{rank}(\boldsymbol{V}_1) = 1$ can be reduced to

$$\max_{\boldsymbol{V}_1} \quad \gamma_{\mathrm{s}}\mathrm{tr}(\boldsymbol{V}_1\boldsymbol{H}_{\mathrm{sir}}^{\mathrm{H}}\boldsymbol{a}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{rid}}^{\mathrm{H}}\boldsymbol{V}_2\boldsymbol{H}_{\mathrm{rid}}\boldsymbol{A}\boldsymbol{H}_{\mathrm{sir}}) \tag{15a}$$

$$\text{s.t.} \quad V_1(n,n) = 1, \ \forall n = 1,2,\cdots,N+1, \ \boldsymbol{V}_1 \succeq \boldsymbol{0}, \tag{15b}$$

$$\gamma_{\mathrm{s}}\mathrm{tr}(\boldsymbol{V}_1\boldsymbol{H}_{\mathrm{sir}}^{\mathrm{H}}\boldsymbol{a}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{H}_{\mathrm{sir}}) + \|\boldsymbol{A}\|_{\mathrm{F}}^2 \le \gamma_{\mathrm{r}}, \tag{15c}$$

which is a standard SDP problem and can be efficiently solved through CVX. Similarly, rank-one solution $\boldsymbol{V}_1$ is recovered via a Gaussian randomization method.

3.3 Optimizing $\boldsymbol{V_2}$ given $\boldsymbol{V_1}$ and $\boldsymbol{A}$

In the case of fixing $\boldsymbol{V}_1$ and $\boldsymbol{A}$, together with not considering $\mathrm{rank}(\boldsymbol{V}_2) = 1$, optimization problem (9) can be simplified as

$$\max_{\boldsymbol{V}_2} \quad \frac{\gamma_{\mathrm{s}}\mathrm{tr}(\boldsymbol{V}_2\boldsymbol{H}_{\mathrm{rid}}\boldsymbol{A}\boldsymbol{H}_{\mathrm{sir}}\boldsymbol{V}_1\boldsymbol{H}_{\mathrm{sir}}^{\mathrm{H}}\boldsymbol{a}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{rid}}^{\mathrm{H}})}{\mathrm{tr}(\boldsymbol{V}_2\boldsymbol{H}_{\mathrm{rid}}\boldsymbol{A}\boldsymbol{a}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{rid}}^{\mathrm{H}}) + 1} \tag{16a}$$

$$\text{s.t.} \quad V_2(n,n) = 1, \ \forall n = 1,2,\cdots,N+1, \ \boldsymbol{V}_2 \succeq \boldsymbol{0}. \tag{16b}$$

Clearly, problem (16) is similar to problem (11), so that the mean derived from solving $\overline{\boldsymbol{A}}$ can be applied to seek $\boldsymbol{V}_2$. In the same manner, the SDR problem (16) can be transformed into the following form:

$$\max_{\widetilde{\boldsymbol{V}}_2,p} \quad \gamma_{\mathrm{s}}\mathrm{tr}(\widetilde{\boldsymbol{V}}_2\boldsymbol{H}_{\mathrm{rid}}\boldsymbol{A}\boldsymbol{H}_{\mathrm{sir}}\boldsymbol{V}_1\boldsymbol{H}_{\mathrm{sir}}^{\mathrm{H}}\boldsymbol{a}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{rid}}^{\mathrm{H}}) \tag{17a}$$

$$\text{s.t.} \quad \widetilde{V}_2(n,n) = p, \ \forall n = 1,2,\cdots,N+1, \ \widetilde{\boldsymbol{V}}_2 \succeq \boldsymbol{0}, \tag{17b}$$

$$\mathrm{tr}(\widetilde{\boldsymbol{V}}_2\boldsymbol{H}_{\mathrm{rid}}\boldsymbol{A}\boldsymbol{a}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{rid}}^{\mathrm{H}}) + p = 1, \ \ p > 0, \tag{17c}$$

where $p = (\mathrm{tr}(\boldsymbol{V}_2\boldsymbol{H}_{\mathrm{rid}}\boldsymbol{A}\boldsymbol{a}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{rid}}^{\mathrm{H}}) + 1)^{-1}$. Solution $\widetilde{\boldsymbol{V}}_2$ can be found via CVX, and then the Gaussian randomization method can be applied to seek $\boldsymbol{V}_2$ with $\mathrm{rank}(\boldsymbol{V}_2) = 1$.

3.4 Overall algorithm and complexity analysis

The related procedure of the proposed CCT-SDP method is summarized in Algorithm 1. The corresponding total computational complexity is written as

$$\begin{aligned}&\mathcal{O}\{L_1[n_{\boldsymbol{A}}\sqrt{M^2+3}(M^6+3+n_{\boldsymbol{A}}(M^4+3)+n_{\boldsymbol{A}}^2)\\&+n_{\boldsymbol{V}_1}\sqrt{2N+3}((N+1)^3+N+2+n_{\boldsymbol{V}_1}((N+1)^2\\&+N+2)+n_{\boldsymbol{V}_1}^2)+n_{\boldsymbol{V}_2}\sqrt{2N+4}((N+1)^3+N+3\\&+n_{\boldsymbol{V}_2}((N+1)^2+N+3)+n_{\boldsymbol{V}_2}^2)]\ln(1/\varepsilon)\}\ \text{FLOPs},\end{aligned} \tag{18}$$

where ε denotes the computational accuracy, and $n_{\boldsymbol{A}} = M^4+1$, $n_{\boldsymbol{V}_1} = (N+1)^2$, and $n_{\boldsymbol{V}_2} = (N+1)^2+1$ are the numbers of variables in Eqs. (14), (15), and (17), respectively. L_1 is the iterative number required to be deployed to reach the convergence condition of the CCT-SDP-based AO method. Clearly, the highest order of computational complexity, as specified in expression (18), is M^{13} and $N^{6.5}$ FLOPs.

4 Proposed low-complexity DT-SCA-based AO method

In the previous section, we have proposed a high-performance CCT-SDP method to optimize the beamforming matrix $\boldsymbol{A}$, as well as the reflection-coefficient matrices $\boldsymbol{\Theta}_1$ and $\boldsymbol{\Theta}_2$, for a high achievable rate. However, the CCT-SDP method is related to matrix optimization, which results in high complexity. Aiming at reducing the high complexity, a low-complexity DT-SCA method is proposed. By optimizing one variable and fixing the two other variables, the optimization problem (8) can be decomposed into the three subproblems described in the forthcoming subsections.

Algorithm 1 Proposed CCT-SDP method

Initialize $\boldsymbol{A}^0$, $\boldsymbol{V}_1^0$, and $\boldsymbol{V}_2^0$. R^0 can be computed based on expression (9a)

Set $t=0$ and convergence accuracy ε

Repeat

Solve problem (14) to obtain $\widetilde{\boldsymbol{A}}^{t+1}$ with given $\boldsymbol{V}_1^t$ and $\boldsymbol{V}_2^t$, and apply Gaussian randomization to recover rank-one solution $\overline{\boldsymbol{A}}^{t+1}$ and obtain $\boldsymbol{A}^{t+1}$

Fix $\boldsymbol{A}^{t+1}$ and $\boldsymbol{V}_2^t$, compute $\boldsymbol{V}_1^{t+1}$ according to problem (15), and recover rank-one solution $\boldsymbol{V}_1^{t+1}$ via Gaussian randomization

Fix $\boldsymbol{A}^{t+1}$ and $\boldsymbol{V}_1^{t+1}$, obtain $\widetilde{\boldsymbol{V}}_2^{t+1}$ by solving problem (17), and use Gaussian randomization to recover rank-one solution $\boldsymbol{V}_2^{t+1}$

Compute R^{t+1}

Until $\left|R^{t+1}-R^t\right|\le\delta$

4.1 Optimization of A with fixed v_1 and v_2

We reformulate the fractional optimization problem (10) as follows:

$$\max_{\boldsymbol{a}} \quad \frac{\boldsymbol{a}^{\mathrm{H}}\overline{\boldsymbol{B}}\boldsymbol{a}}{\boldsymbol{a}^{\mathrm{H}}\overline{\boldsymbol{C}}\boldsymbol{a}+1} \tag{19a}$$

$$\text{s.t.} \quad \boldsymbol{a}^{\mathrm{H}}(\overline{\boldsymbol{D}}+\boldsymbol{I}_{M^2})\boldsymbol{a}\le\gamma_{\mathrm{r}}, \tag{19b}$$

where matrix $\overline{\boldsymbol{B}}=\gamma_{\mathrm{s}}(\boldsymbol{H}_{\mathrm{sir}}^{*}\boldsymbol{v}_1^{*}(\boldsymbol{H}_{\mathrm{sir}}\boldsymbol{v}_1)^{\mathrm{T}})\otimes(\boldsymbol{H}_{\mathrm{rid}}^{\mathrm{H}}\boldsymbol{v}_2\boldsymbol{v}_2^{\mathrm{H}}\boldsymbol{H}_{\mathrm{rid}})$, $\overline{\boldsymbol{C}}=\boldsymbol{I}_M\otimes(\boldsymbol{H}_{\mathrm{rid}}^{\mathrm{H}}\boldsymbol{v}_2\boldsymbol{v}_2^{\mathrm{H}}\boldsymbol{H}_{\mathrm{rid}})$, and $\overline{\boldsymbol{D}}=\gamma_{\mathrm{s}}(\boldsymbol{H}_{\mathrm{sir}}^{*}\boldsymbol{v}_1^{*}(\boldsymbol{H}_{\mathrm{sir}}\boldsymbol{v}_1)^{\mathrm{T}})\otimes\boldsymbol{I}_M$. Note that the numerator of the objective function in Eq. (19a) is convex, and that the above problem cannot be directly solved using the Dinkelbachs transformation (Shen and Yu, 2018). Here, we apply the SCA method to approximate $\boldsymbol{a}^{\mathrm{H}}\overline{\boldsymbol{B}}\boldsymbol{a}$ to a linear function given by

$$\boldsymbol{a}^{\mathrm{H}}\overline{\boldsymbol{B}}\boldsymbol{a}\ge2\Re\{\boldsymbol{a}^{\mathrm{H}}\overline{\boldsymbol{B}}\widetilde{\boldsymbol{a}}\}-\widetilde{\boldsymbol{a}}^{H}\overline{\boldsymbol{B}}\widetilde{\boldsymbol{a}}. \tag{20}$$

The above inequality is the corresponding first-order Taylor expansion of $\boldsymbol{a}^{\mathrm{H}}\overline{\boldsymbol{B}}\boldsymbol{a}$ at feasible point $\widetilde{\boldsymbol{a}}$. Therefore, problem (19) can be transformed into the following form:

$$\max_{\boldsymbol{a}} \quad 2\Re\{\boldsymbol{a}^{\mathrm{H}}\overline{\boldsymbol{B}}\widetilde{\boldsymbol{a}}\}-\widetilde{\boldsymbol{a}}^{H}\overline{\boldsymbol{B}}\widetilde{\boldsymbol{a}}-\mu(\boldsymbol{a}^{\mathrm{H}}\overline{\boldsymbol{C}}\boldsymbol{a}+1) \tag{21a}$$

$$\text{s.t.} \quad \boldsymbol{a}^{\mathrm{H}}(\overline{\boldsymbol{D}}+\mathbf{I}_{M^2})\boldsymbol{a}\le\gamma_{\mathrm{r}}, \tag{21b}$$

where μ is iteratively updated in accordance with

$$\mu(t+1)=\frac{2\Re\{\boldsymbol{a}^{\mathrm{H}}(t)\overline{\boldsymbol{B}}\widetilde{\boldsymbol{a}}\}-\widetilde{\boldsymbol{a}}^{H}\overline{\boldsymbol{B}}\widetilde{\boldsymbol{a}}}{\boldsymbol{a}^{\mathrm{H}}(t)\overline{\boldsymbol{C}}\boldsymbol{a}(t)+1}, \tag{22}$$

where t is the iteration index. Since μ is nondecreasing, the convergence of the objective function in (21a) can be guaranteed. Obviously, problem (21), consisting of a concave objective function and a convex constraint, is convex, and its solution $\boldsymbol{a}$ can be computed via CVX. Consequently, $\boldsymbol{a}$ can be obtained.

4.2 Optimization of v_1 with fixed v_2 and a

Given $\boldsymbol{v}_2$ and $\boldsymbol{a}$, problem (8) can be reduced to the following form:

$$\max_{\boldsymbol{v}_1} \quad \boldsymbol{v}_2^{\mathrm{H}}\boldsymbol{E}\boldsymbol{v}_1 \tag{23a}$$

$$\text{s.t.} \quad |v_1(i)|=1,\ \forall i=1,2,\cdots,N, \tag{23b}$$

$$v_1(N+1)=1, \tag{23c}$$

$$\boldsymbol{v}_2^{\mathrm{H}}\boldsymbol{F}\boldsymbol{v}_1+\|\boldsymbol{a}\|_{\mathrm{F}}^2\le\gamma_{\mathrm{r}}, \tag{23d}$$

where matrix $\boldsymbol{E}=\gamma_{\mathrm{s}}\boldsymbol{H}_{\mathrm{sir}}^{\mathrm{H}}\boldsymbol{a}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{rid}}^{\mathrm{H}}\boldsymbol{v}_2\boldsymbol{v}_2^{\mathrm{H}}\boldsymbol{H}_{\mathrm{rid}}\boldsymbol{a}\boldsymbol{H}_{\mathrm{sir}}$ and $\boldsymbol{F}=\gamma_{\mathrm{s}}\boldsymbol{H}_{\mathrm{sir}}^{\mathrm{H}}\boldsymbol{a}^{\mathrm{H}}\boldsymbol{a}\boldsymbol{H}_{\mathrm{sir}}$. The above problem is non-convex because of the non-concave objective function in (23a) and the unit-modulus constraint in (23b). (23a) can be an approximated linear function through its first-order Taylor expansion at feasible vector $\widetilde{\boldsymbol{v}}_1$, which is

$$\boldsymbol{v}_2^{\mathrm{H}}\boldsymbol{E}\boldsymbol{v}_1\ge2\Re\{\boldsymbol{v}_2^{\mathrm{H}}\boldsymbol{E}\widetilde{\boldsymbol{v}}_1\}-\widetilde{\boldsymbol{v}}_1^{\mathrm{H}}\boldsymbol{E}\widetilde{\boldsymbol{v}}_1. \tag{24}$$

The unit-modulus constraint in Eq. (23b) can be relaxed to

$$|v_1(i)|^2\le1,\ \forall i=1,2,\cdots,N. \tag{25}$$

After the above transformation, we have the following optimization problem:

$$\max_{\boldsymbol{v}_1} \quad 2\Re\{\boldsymbol{v}_2^{\mathrm{H}}\boldsymbol{E}\widetilde{\boldsymbol{v}}_1\}-\widetilde{\boldsymbol{v}}_1^{H}\boldsymbol{E}\widetilde{\boldsymbol{v}}_1 \tag{26a}$$

$$\text{s.t.} \quad \text{(23c), (23d), (25).} \tag{26b}$$

The above is a convex optimization problem, and its solution, denoted as $\widehat{\boldsymbol{v}}_1$, is achieved via CVX. Accordingly, the solution to problem (23) is given by

$$\boldsymbol{v}_1=\mathrm{e}^{\mathrm{j}\angle\frac{\widehat{\boldsymbol{v}}_1}{\widehat{v}_1(N+1)}}. \tag{27}$$

4.3 Optimization of v_2 with fixed v_1 and a

When $\boldsymbol{v}_1$ and $\boldsymbol{a}$ are given, problem (8) can be simplified as

$$\max_{\boldsymbol{v}_2} \quad \frac{\boldsymbol{v}_2^{\mathrm{H}}\boldsymbol{G}\boldsymbol{v}_2}{\boldsymbol{v}_2^{\mathrm{H}}\boldsymbol{J}\boldsymbol{v}_2} \tag{28a}$$

$$\text{s.t.} \quad |v_2(i)|=1,\ \forall i=1,2,\cdots,N, \tag{28b}$$

$$v_2(N+1)=1, \tag{28c}$$

where matrix $\boldsymbol{G} = \gamma_{\mathrm{s}}\boldsymbol{H}_{\mathrm{rid}}\boldsymbol{a}\boldsymbol{H}_{\mathrm{sir}}\boldsymbol{v}_1\boldsymbol{v}_2^{\mathrm{H}}\boldsymbol{H}_{\mathrm{sir}}^{\mathrm{H}}\boldsymbol{a}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{rid}}^{\mathrm{H}}$ and $\boldsymbol{J} = \boldsymbol{H}_{\mathrm{rid}}\boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}\boldsymbol{H}_{\mathrm{rid}}^{\mathrm{H}} + \frac{\boldsymbol{I}_{N+1}}{N+1}$. Similarly, the SCA method is applied to achieve the low bound of (28a). Its first-order Taylor expansion at feasible vector $\widetilde{\boldsymbol{v}}_2$ can be expressed as (Guan et al., 2020)

$$\frac{\boldsymbol{v}_2^{\mathrm{H}}\boldsymbol{G}\boldsymbol{v}_2}{\boldsymbol{v}_2^{\mathrm{H}}\boldsymbol{J}\boldsymbol{v}_2} \geq 2\Re\{\boldsymbol{g}^{\mathrm{H}}\boldsymbol{v}_2\} + d, \tag{29}$$

where

$$\boldsymbol{g}^{\mathrm{H}} = \frac{\widetilde{\boldsymbol{v}}_2^{\mathrm{H}}\boldsymbol{G}}{\widetilde{\boldsymbol{v}}_2^{\mathrm{H}}\boldsymbol{J}\widetilde{\boldsymbol{v}}_2} - \widetilde{\boldsymbol{v}}_2^{\mathrm{H}}(\boldsymbol{J} - \lambda_{\max}(\boldsymbol{J})\boldsymbol{I}_{N+1})\frac{\widetilde{\boldsymbol{v}}_2^{\mathrm{H}}\boldsymbol{G}\widetilde{\boldsymbol{v}}_2}{(\widetilde{\boldsymbol{v}}_2^{\mathrm{H}}\boldsymbol{J}\widetilde{\boldsymbol{v}}_2)^2}, \tag{30a}$$

$$d = -\left[2\lambda_{\max}(\boldsymbol{J})(N+1) - \widetilde{\boldsymbol{v}}_2^{\mathrm{H}}\boldsymbol{J}\widetilde{\boldsymbol{v}}_2\right]\frac{\widetilde{\boldsymbol{v}}_2^{\mathrm{H}}\boldsymbol{G}\widetilde{\boldsymbol{v}}_2}{(\widetilde{\boldsymbol{v}}_2^{\mathrm{H}}\boldsymbol{J}\widetilde{\boldsymbol{v}}_2)^2}. \tag{30b}$$

In the same manner, problem (28) can be transformed into the following form:

$$\max_{\boldsymbol{v}_2} \quad 2\Re\{\boldsymbol{g}^{\mathrm{H}}\boldsymbol{v}_2\} + d \tag{31a}$$

$$\text{s.t.} \quad |v_2(i)|^2 \leq 1,\ \forall i = 1, 2, \cdots, N, \tag{31b}$$

which is convex. Its solution $\widehat{\boldsymbol{v}}_2$ can be efficiently found using CVX. Therefore, the solution to problem (28) can be denoted as

$$\boldsymbol{v}_2 = \mathrm{e}^{\mathrm{j}\angle\frac{\widehat{\boldsymbol{v}}_2}{\widehat{v}_2(N+1)}}. \tag{32}$$

4.4 Overall algorithm and complexity analysis

The related procedure of DT-SCA is summarized in Algorithm 2. The corresponding total complexity of Algorithm 2 is

$$\begin{aligned}&\mathcal{O}\{L_2[n_{\boldsymbol{a}}\sqrt{2}(M^4 + n_{\boldsymbol{a}}^2) + n_{\boldsymbol{v}_1}\sqrt{2N+3}(n_{\boldsymbol{v}_1}\\&+ N + 1 + (N+2)^2 + n_{\boldsymbol{v}_1}^2) + n_{\boldsymbol{v}_2}\sqrt{2N+1}(n_{\boldsymbol{v}_2}\\&+ N + 1 + n_{\boldsymbol{v}_2}^2)]\ln(1/\varepsilon)\}\ \text{FLOPs}.\end{aligned} \tag{33}$$

The numbers of variables $n_{\boldsymbol{a}}$, $n_{\boldsymbol{v}_1}$, and $n_{\boldsymbol{v}_2}$ in Eqs. (21), (26), and (31) are equal to M^2, $N+1$, and $N+1$, respectively. L_2 is the iteration number satisfying the convergence condition of DT-SCA. The highest order of computational complexity, as specified in expression (33), is M^6 and $N^{3.5}$ FLOPs, which is much lower than that of the CCT-SDP method.

5 Simulation and numerical results

In this section, numerical simulations are applied to verify the rate performances of the two proposed schemes. For convenience, the positions of S, D, RIS, and AF relay are set as (0, 0, 0), (0, 100, 0), (−10, 50, 20), and (10, 50, 10) m, respectively, in the three-dimensional (3D) space. The path loss between the transmitter and receiver can be computed according to the expression $\mathrm{PL}(d) = \mathrm{PL}_0 - 10\alpha\lg(\frac{d}{d_0})$, where α represents the path loss exponent, d is the distance, and $\mathrm{PL}_0 = -30$ dB denotes the reference path loss at $d_0 = 1$ m. Here, the path loss exponents of channel link S–RIS, S–AF relay, RIS–AF relay, RIS–D, and AF relay–D are set as 2.0, 3.5, 2.0, 2.0, and 3.5, respectively. Additionally, let $\sigma^2 = -90$ dBm.

Algorithm 2 Proposed DT-SCA method

Initialize $\boldsymbol{a}^0$, $\boldsymbol{v}_1^0$, and $\boldsymbol{v}_2^0$. Achieve R^0 in line with problem (8). Set $t = 0$ and convergence accuracy as ε

Repeat

Given $\boldsymbol{v}_1^t$ and $\boldsymbol{v}_2^t$, find $\boldsymbol{a}^{t+1}$ by solving problem (21), and obtain $\boldsymbol{a}^{t+1}$

Given $\boldsymbol{a}_1^{t+1}$ and $\boldsymbol{v}_2^t$, find $\boldsymbol{v}_1^{t+1}$ by solving problem (26)

Given $\boldsymbol{a}^{t+1}$ and $\boldsymbol{v}_1^{t+1}$, find $\boldsymbol{v}_2^{t+1}$ by solving problem (31)

Achieve R^{t+1}

Until $\left|R^{t+1} - R^t\right| \leq \delta$

Figs. 2 and 3 compare the computational complexities between the two proposed methods in terms of parameter N with $(M, L_1, L_2) = (2, 10, 10)$ and parameter M with $(N, L_1, L_2) = (256, 10, 10)$, respectively. Regardless of the inferences that we might draw from Figs. 2 and 3, it is obvious that the computational complexities of CCT-SDP and DT-SCA increase as N and M increase, while the complexity of DT-SCA is much lower than that of CCT-SDP, which is consistent with our analysis of the complexity in previous sections. Additionally, it is interesting to find that the complexity difference is becoming increasingly significant with the increases of N and M.

Fig. 4 verifies the convergence of the two proposed algorithms with $P_{\mathrm{s}} = 15$ dBm and $P_{\mathrm{s}} = 25$ dBm. When $P_{\mathrm{s}} = 15$ dBm, about four iterations are required for convergence to be achieved between CCT-SDP and DT-SCA. In the case of $P_{\mathrm{s}} = 25$ dBm, the two proposed methods can come up to the maximum rate within eight iterations. It is revealed that CCT-SDP and DT-SCA are convergent and feasible.

Fig. 5 depicts the achievable rate versus P_{s} with $(M, N, P_{\mathrm{r}}) = (2, 256, 30\ \mathrm{dBm})$. It is obvious that the

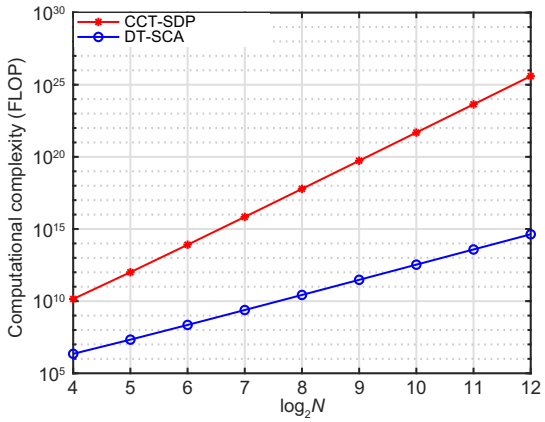

Fig. 2 Computational complexity versus N with $(M, L_1, L_2) = (2, 10, 10)$

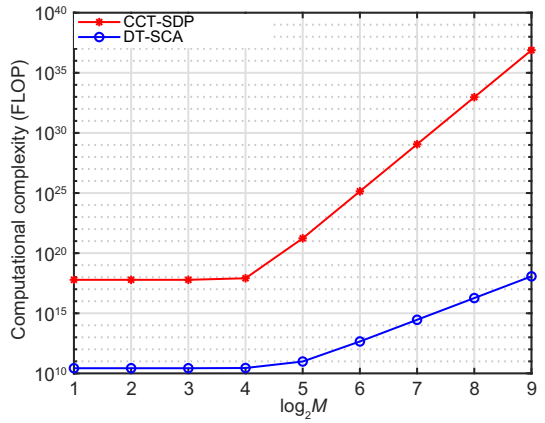

Fig. 3 Computational complexity versus M with $(N, L_1, L_2) = (256, 10, 10)$

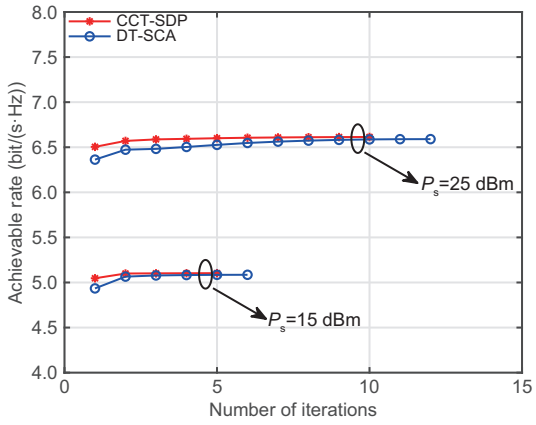

Fig. 4 Convergence of the two proposed methods with $(M, N, P_r) = (2, 256, 30\ \text{dBm})$

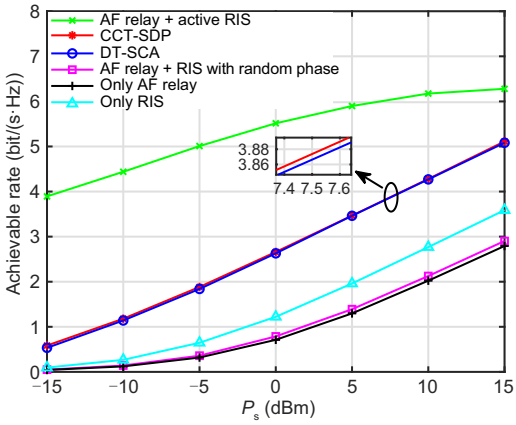

Fig. 5 Achievable rate versus P_s with $(M, N, P_r) = (2, 256, 30\ \text{dBm})$

rates of CCT-SDP and DT-SCA, as well as those of the four benchmark networks, AF relay + active RIS, AF relay + RIS with random phase, only AF relay, and only RIS, increase as the transmission power P_s at S increases. Since multiple paths have been created with the aid of RIS, and the beamforming matrix at AF relay and the phase shifts at RIS have been optimized, CCT-SDP and DT-SCA can perform much better than AF relay + RIS with random phase, only AF relay, and only RIS. Additionally, the rate of DT-SCA is slightly lower than that of CCT-SDP in the whole P_s region. Since the transmission power budget at active RIS is limited, as P_s increases, the rate performance gap between the two proposed methods and AF relay + active RIS decreases.

Fig. 6 shows the achievable rate versus the number N of RIS elements with $(M, P_s, P_r) = (2, 10\ \text{dBm}, 30\ \text{dBm})$. It can be seen that as N increases, the rates of the two proposed methods and two benchmark schemes (i.e., AF relay + active RIS and only RIS) increase, while those of the two other benchmark schemes (i.e., AF relay + RIS with random phase and only AF relay) are almost stable. For small-scale RIS, the rate difference among the two proposed methods, AF relay + RIS with random phase, and only AF relay is small. For medium- and large-scale RIS, the gap between CCT-SDP and DT-SCA grows smaller. When $N = 256$, the rate of DT-SCA is approximate to that of CCT-SDP. Meanwhile, compared to AF relay + RIS with random phase, only AF relay, and only RIS, the rate is improved by 104.0%, 113.0%, and 55.6% respectively using the two proposed schemes.

6 Conclusions

The present study investigated a RIS-aided AF relay wireless network. Two AO methods were proposed based on the criterion of maximum SNR, namely CCT-SDP and DT-SCA. Here, the

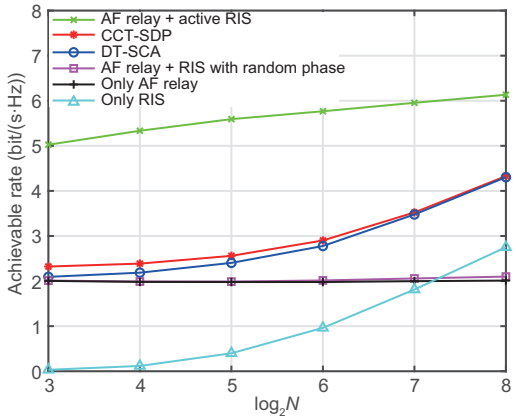

Fig. 6 Achievable rate versus N with $(M, P_s, P_r) = (2, 10\ \text{dBm}, 30\ \text{dBm})$

beamforming matrix at AF relay and phase-shift matrices at RIS were jointly optimized for rate performance enhancement. Simulations verified that the proposed CCT-SDP and DT-SCA schemes are convergent, and apparent rate improvement can be achieved compared to a RIS-assisted AF relay network with random phase, an AF relay network without RIS, and a RIS-aided network without AF relay. Additionally, it was proved that the rate performance achieved by the low-complexity DT-SCA method is slightly lower than that of the high-performance CCT-SDP method.

Contributors

Xuehui WANG and Feng SHU designed the research. Xuehui WANG, Riqing CHEN, and Peng ZHANG processed the data. Xuehui WANG and Qi ZHANG drafted the paper. Guiyang XIA and Weiping SHI helped organize the paper. Feng SHU and Jiangzhou WANG revised and finalized the paper.

Compliance with ethics guidelines

Xuehui WANG, Feng SHU, Riqing CHEN, Peng ZHANG, Qi ZHANG, Guiyang XIA, Weiping SHI, and Jiangzhou WANG declare that they have no conflict of interest.

Data availability

Due to the nature of this research, participants of this study did not agree for their data to be shared publicly, so supporting data are not available.

References

Abdullah Z, Chen GJ, Lambotharan S, et al., 2020. A hybrid relay and intelligent reflecting surface network and its ergodic performance analysis. *IEEE Wirel Commun Lett*, 9(10):1653-1657. https://doi.org/10.1109/LWC.2020.2999918

Abdullah Z, Kisseleff S, Ntontin K, et al., 2022. Double-RIS communication with DF relaying for coverage extension: is one relay enough? IEEE Int Conf on Communications, p.2639-2644. https://doi.org/10.1109/ICC45855.2022.9839181

An JC, Yuen C, Huang CW, et al., 2023a. A tutorial on holographic MIMO communications—part I: channel modeling and channel estimation. *IEEE Commun Lett*, 27(7):1664-1668. https://doi.org/10.1109/LCOMM.2023.3278683

An JC, Yuen C, Huang CW, et al., 2023b. A tutorial on holographic MIMO communications—part II: performance analysis and holographic beamforming. *IEEE Commun Lett*, 27(7):1669-1673. https://doi.org/10.1109/LCOMM.2023.3278682

Bie QY, Liu Y, Wang YX, et al., 2022. Deployment optimization of reconfigurable intelligent surface for relay systems. *IEEE Trans Green Commun Netw*, 6(1):221-233. https://doi.org/10.1109/TGCN.2022.3145026

Björnson E, Özdogan Ö, Larsson EG, 2020. Intelligent reflecting surface versus decode-and-forward: how large surfaces are needed to beat relaying? *IEEE Wirel Commun Lett*, 9(2):244-248. https://doi.org/10.1109/LWC.2019.2950624

Charnes A, Cooper WW, 1962. Programming with linear fractional functionals. *Nav Res Log Q*, 9(3-4):181-186. https://doi.org/10.1002/nav.3800090303

Chen Z, Tang J, Zhang XY, et al., 2022. Hybrid evolutionary-based sparse channel estimation for IRS-assisted mmWave MIMO systems. *IEEE Trans Wirel Commun*, 21(3):1586-1601. https://doi.org/10.1109/TWC.2021.3105405

Guan XR, Wu QQ, Zhang R, 2020. Joint power control and passive beamforming in IRS-assisted spectrum sharing. *IEEE Commun Lett*, 24(7):1553-1557. https://doi.org/10.1109/LCOMM.2020.2979709

Guan XR, Wu QQ, Zhang R, 2022. Anchor-assisted channel estimation for intelligent reflecting surface aided multiuser communication. *IEEE Trans Wirel Commun*, 21(6):3764-3778. https://doi.org/10.1109/TWC.2021.3123674

Guo HY, Liang YC, Chen J, et al., 2020. Weighted sum-rate maximization for reconfigurable intelligent surface aided wireless netwroks. *IEEE Trans Wirel Commun*, 19(5):3064-3076. https://doi.org/10.1109/TWC.2020.2970061

Hong S, Pan CH, Ren H, et al., 2021. Robust transmission design for intelligent reflecting surface-aided secure communication systems with imperfect cascaded CSI. *IEEE Trans Wirel Commun*, 20(4):2487-2501. https://doi.org/10.1109/TWC.2020.3042828

Jiang W, Schotten HD, 2022. Intelligent reflecting vehicle surface: a novel IRS paradigm for moving vehicular networks. IEEE Military Communications Conf, p.793-798. https://doi.org/10.1109/MILCOM55135.2022.10017691

Jiang WH, Chen BL, Zhao J, et al., 2021. Joint active and passive beamforming design for the IRS-assisted MIMOME-OFDM secure communications. *IEEE Trans*

Veh Technol, 70(10):10369-10381. https://doi.org/10.1109/TVT.2021.3106351

Khalid W, Shahjalal M, Yu H, 2022. Outage performance analysis of hybrid relay-reconfigurable intelligent surface networks. Proc 27th Asia Pacific Conf on Communications, p.253-254. https://doi.org/10.1109/APCC55198.2022.9943604

Lee J, Shin W, Lee J, 2021. Performance analysis of IRS-assisted LEO satellite communication systems. Int Conf on Information and Communication Technology Convergence, p.323-325. https://doi.org/10.1109/ICTC52510.2021.9621010

Li GH, Yue DW, Jin SN, et al., 2022. Hybrid double-RIS and DF-relay for outdoor-to-indoor communication. *IEEE Access*, 10:126651-126663. https://doi.org/10.1109/ACCESS.2022.3225876

Obeed M, Chaaban A, 2022. Joint beamforming design for multi-user MISO downlink aided by a reconfigurable intelligent surface and a relay. *IEEE Trans Wirel Commun*, 21(10):8216-8229. https://doi.org/10.1109/TWC.2022.3164903

Shen KM, Yu W, 2018. Fractional programming for communication systems—part II: uplink scheduling via matching. *IEEE Trans Signal Process*, 66(10):2631-2644. https://doi.org/10.1109/TSP.2018.2812748

Shi WP, Zhou XB, Jia LQ, et al., 2021a. Enhanced secure wireless information and power transfer via intelligent reflecting surface. *IEEE Commun Lett*, 25(4):1084-1088. https://doi.org/10.1109/LCOMM.2020.3043475

Shi WP, Li JY, Xia GY, et al., 2021b. Secure multigroup multicast communication systems via intelligent reflecting surface. *China Commun*, 18(3):39-51. https://doi.org/10.23919/JCC.2021.03.004

Shu F, Lu Y, Chen YZ, et al., 2014. High-sum-rate beamformers for multi-pair two-way relay networks with amplify-and-forward relaying strategy. *Sci China Inform Sci*, 57(2):1-11. https://doi.org/10.1007/s11432-013-4980-9

Shu F, Teng Y, Li JY, et al., 2021a. Enhanced secrecy rate maximization for directional modulation networks via IRS. *IEEE Trans Commun*, 69(12):8388-8401. https://doi.org/10.1109/TCOMM.2021.3110598

Shu F, Jiang XY, Liu XY, et al., 2021b. Precoding and transmit antenna subarray selection for secure hybrid spatial modulation. *IEEE Trans Wirel Commun*, 20(3):1903-1917. https://doi.org/10.1109/TWC.2020.3037217

Shu F, Yang LL, Jiang XY, et al., 2022. Beamforming and transmit power design for intelligent reconfigurable surface-aided secure spatial modulation. *IEEE J Sel Top Signal Process*, 16(5):933-949. https://doi.org/10.1109/JSTSP.2022.3172682

Sun ZW, Wang XH, Feng SL, et al., 2023. Pilot optimization and channel estimation for two-way relaying network aided by IRS with finite discrete phase shifters. *IEEE Trans Veh Technol*, 72(4):5502-5507. https://doi.org/10.1109/TVT.2022.3230423

Tian Z, Chen ZC, Wang M, et al., 2022. Reconfigurable intelligent surface empowered optimization for spectrum sharing: scenarios and methods. *IEEE Veh Technol Mag*, 17(2):74-82. https://doi.org/10.1109/MVT.2022.3157070

Wang MX, Duan W, Zhang GA, et al., 2022. On the achievable capacity of cooperative NOMA networks: RIS or relay? *IEEE Wirel Commun Lett*, 11(8):1624-1628. https://doi.org/10.1109/LWC.2022.3169806

Wang XH, Shu F, Shi WP, et al., 2022. Beamforming design for IRS-aided decode-and-forward relay wireless network. *IEEE Trans Green Commun Netw*, 6(1):198-207. https://doi.org/10.1109/TGCN.2022.3145031

Wang XH, Zhang P, Shu F, et al., 2023. Power allocation for IRS-aided two-way decode-and-forward relay wireless network. *IEEE Trans Veh Technol*, 72(1):1337-1342. https://doi.org/10.1109/TVT.2022.3205725

Wei L, Huang CW, Alexandropoulos GC, et al., 2021. Channel estimation for RIS-empowered multi-user MISO wireless communications. *IEEE Trans Commun*, 69(6):4144-4157. https://doi.org/10.1109/TCOMM.2021.3063236

Wu QQ, Zhang R, 2019. Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming. *IEEE Trans Wirel Commun*, 18(11):5394-5409. https://doi.org/10.1109/TWC.2019.2936025

Yang SJ, Lyu W, Xiu Y, et al., 2023. Active 3D double-RIS-aided multi-user communications: two-timescale-based separate channel estimation via Bayesian learning. *IEEE Trans Commun*, 71(6):3605-3620. https://doi.org/10.1109/TCOMM.2023.3265115

Yildirim I, Kilinc F, Basar E, et al., 2021. Hybrid RIS-empowered reflection and decode-and-forward relaying for coverage extension. *IEEE Commun Lett*, 25(5):1692-1696. https://doi.org/10.1109/LCOMM.2021.3054819

Zheng BX, Lin SE, Zhang R, 2022. Intelligent reflecting surface-aided LEO satellite communication: cooperative passive beamforming and distributed channel estimation. *IEEE J Sel Areas Commun*, 40(10):3057-3070. https://doi.org/10.1109/JSAC.2022.3196119

Zhou X, Li J, Shu F, et al., 2019. Secure SWIPT for directional modulation-aided AF relaying networks. *IEEE J Sel Areas Commun*, 37(2):253-268. https://doi.org/10.1109/JSAC.2018.2872372

Zhou XB, Yan SH, Wu QQ, et al., 2022. Intelligent reflecting surface (IRS)-aided covert wireless communications with delay constraint. *IEEE Trans Wirel Commun*, 21(1):532-547. https://doi.org/10.1109/TWC.2021.3098099