

Hybrid Active-Passive RIS Empowered Secure Communications: Joint Architecture Design and Beamforming Optimization

Shiqi Gong, Yue Ju, Heng Liu, Liang Liu, *Senior Member, IEEE*,
and Chengwen Xing, *Member, IEEE*

Abstract—Reconfigurable intelligent surface (RIS) has recently emerged as a promising solution to significantly enhance the security of wireless communication systems. By combining the advantages of the conventional fully-active and fully-passive RISs, a novel hybrid active-passive RIS has been anticipated to achieve the excellent communication performance at a low cost. In this paper, we aim to maximize the secrecy rate in a hybrid active-passive RIS assisted multi-input single-output multi-antenna Eve (MISOME) system, where both fixed and dynamic hybrid RIS architectures are considered. To tackle this intractable problem, we jointly optimize the transmit covariance matrix at the base station (BS), the reflection matrices of active and passive sub-RISs, as well as the element allocation matrix for the dynamic hybrid RIS. Specifically, we firstly explore the rank-1 structure of the optimal BS transmit covariance matrix. Then, for the fixed hybrid RIS, we develop an efficient two-loop successive convex approximation (SCA) based iterative algorithm, where the optimal semi-closed-form solution to each subproblem can be obtained. For the dynamic RIS, this proposed algorithm is still applicable by relaxing the binary active/passive elements allocation variables into exponent-based continuous ones. Simulation results validate the superior secrecy performance of the proposed designs over the existing fully-active and fully-passive RIS designs. Moreover, it is demonstrated that the dynamic hybrid RIS is able to strike a good balance between the passive beamforming gain and the power amplification gain to adapt to the varying propagation environment.

Index Terms—Fixed and dynamic hybrid active-passive RISs, secrecy rate, MISOME system, two-loop SCA based iterative algorithm, power amplification gain.

I. INTRODUCTION

Driven by the rapidly growing demands for high-speed data transmission and low-latency communication, security is of paramount importance in the next-generation wireless communication systems [1], [2]. However, due to the broadcast nature

of wireless propagation channels, the transmitted confidential signals are susceptible to the information leakage [3]. In recent years, physical layer security (PLS) has become a promising candidate to enable secure communications, which originates from Shannon's information theory. By exploiting the inherent properties of wireless channels, various PLS techniques have been proposed to reduce the information leakage, and thus enhance system secrecy rate. There have been two widely adopted PLS techniques in current systems, i.e., the so-called artificial noise (AN) and cooperative beamforming, which are respectively applied to disrupt eavesdroppers and strengthen the signal reception at the legitimate user, thereby both enhancing the system security [4], [5]. Nevertheless, the former technique generally requires additional power allocation between the transmit signal and AN at the transmitter. Moreover, the implementation of cooperative beamforming relies on the collaboration among multiple nodes, which inevitably increases burden to the system deployment [6]. Thus, a cost-effective and energy-efficient approach to enhance system security is in urgent need.

Reconfigurable intelligent surface (RIS), which is known as a cost-effective technique, has emerged as a promising solution to conquer these dilemmas. The signals reflected by the RIS can be either combined with the signals from the direct path constructively at the legitimate user to increase the received signal power or combined destructively at the Eve to suppress the information leakage, thus enhancing the secrecy rate. Different from the traditional relay equipped with energy-intensive radio frequency (RF) chains, RIS is in essence a metamaterial-based planar array consisting of abundant low-cost tunable elements and is able to adjust the phases and amplitudes of the incident signals via a smart controller, thereby reshaping the electromagnetic propagation environment [7]. More importantly, the deployment of RIS is able to provide additional optimization degrees of freedom (DoFs) for improving system performance. Inspired by these remarkable advantages of RIS, there have been a steady stream of works focusing on RIS assisted secure sensing and communications. For example, the authors of [8] leveraged RIS to prevent passive sensing by malicious devices. The authors of [9] further developed a RIS aided electromagnetic stealth system to evade radar sensing.

From the perspective of secure communications, the integration of RIS and PLS techniques has attracted a widespread research attention. It is worth noting that most existing works

Manuscript received June 30, 2024; revised October 26, 2024 and December 1, 2024; accepted December 11, 2024. This work was supported in part by the National Natural Science Foundation of China under Grants 62471032 and 62231004, and in part by the Research Grants Council, Hong Kong, China, under Grant 15203222. The associate editor coordinating the review of this article and approving it for publication was Prof. Yu Han. (*Corresponding authors: Heng Liu; Yue Ju.*)

S. Gong is with the School of Cyberspace Science and Technology, Beijing Institute of Technology, Beijing 100081, China (e-mail: gsqyx@163.com).

Y. Ju, H. Liu and C. Xing are with the School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China (e-mails: yueju.ee@gmail.com; heng_liu_bit_ee@163.com; xingchengwen@gmail.com).

L. Liu is with the Department of Electrical and Electronic Engineering, The Hong Kong Polytechnic University, Hong Kong, China (e-mail: liang-eie.liu@polyu.edu.hk).

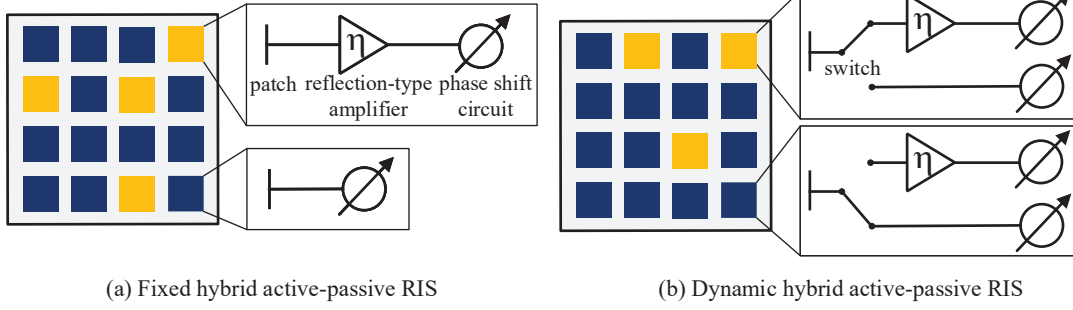


Fig. 1. Two types of hybrid active-passive RIS architectures.

focused on fully-passive RIS assisted communication systems, such as multi-input single-output single-antenna Eve (MISOSE) [10], [11], multi-input single-output multi-antenna Eve (MISOME) [12] and multi-input multiple-output multi-antenna Eve (MIMOME) systems [13]. Although the fully-passive RIS has a low-cost advantage, it usually suffers from severe multiplicative path loss when not deployed in close proximity to the transmitter/receiver, resulting in non-negligible performance degradation. To address this challenge, extensive research works are further extended to the fully-active RIS assisted communication systems [14]. Compared with the fully-passive RIS, the fully-active RIS additionally integrates a low-power reflection-type amplifier into each reflecting element, which can amplify the incident signals and effectively combat the multiplicative path loss [15]. In [16], the authors investigated the secrecy rate maximization (SRM) problem by jointly optimizing the transmit beamforming matrix at the base station (BS), along with phase shifts and amplitudes of RIS elements in the fully-active RIS assisted secure system. However, the hardware overheads and power consumption of the fully-active RIS is much higher than those of the fully-passive RIS. Therefore, the fully-active and fully-passive RISs have complementary advantages in terms of performance improvement and energy consumption [17]. By combining their advantages, a hybrid active-passive RIS has been proposed, whose architecture is shown in Fig. 1(a). The hybrid RIS leverages its active and passive elements to achieve both amplification and reflection capabilities. As compared to the traditional relay technique, the hybrid RIS eliminates the need for costly and energy-intensive RF chains, reduces the signal processing complexity and also mitigates severe self-interference. Therefore, the hybrid RIS is anticipated to strike a good balance between system performance and energy consumption.

Recently, the hybrid active-passive RIS has become a research hotspot and been integrated into many scenarios, such as the multi-input multi-output (MIMO) system [18], the integrated sensing and communication (ISAC) system [19], and the Terahertz (THz) communication system [20]. However, all the aforementioned works aimed at the fixed hybrid RIS, i.e., the numbers and positions of active and passive elements are fixed. Obviously, the fixed hybrid RIS has limited flexibility in optimizing the reflection matrix, and potentially leads to performance degradation. Therefore, a dynamic hybrid active-passive RIS has been proposed. As shown in Fig. 1(b), each element of the dynamic hybrid RIS can be switched adaptively to active or passive mode according to the realistic

propagation environments. Compared with the fixed hybrid RIS in Fig. 1(a), this dynamic architecture provides more spatial DoFs to enhance system performance by allowing the dynamic allocation between active and passive elements, which means that both the numbers and positions of active and passive elements can be optimized. However, the dynamic active/passive elements allocation is quite challenging to solve due to its discrete and nonconvex nature [21]. Currently, most existing works on the dynamic hybrid RIS devote efforts to improving system spectral/energy efficiency. For example, the authors of [22] and [23] respectively investigated the ergodic capacity maximization and energy efficiency maximization problems for the dynamic hybrid RIS assisted communication systems, where the optimal numbers of active and passive elements were studied. In contrast, the authors of [24] focused on determining the optimal positions of active/passive elements given their corresponding numbers. To explore the performance potential of the dynamic hybrid RIS, Xie *et al.* jointly optimized the BS transmit beamforming vector, the RIS reflection coefficients and the element allocation matrix by utilizing the continuous relaxation and successive convex approximation (SCA) methods [25].

To our best knowledge, there have been a few research works that study the integration of hybrid RIS and PLS techniques. Ma *et al.* employed the semidefinite relaxation (SDR) and particle swarm optimization (PSO) algorithms to tackle a SRM problem in the downlink MISOSE communication system in the presence of a malicious multi-antenna jammer [26]. Similarly, the authors of [27] and [28] aimed at improving the secrecy capacity by jointly optimizing the BS transmit beamforming vector and the hybrid RIS reflection matrix in MIMOME systems. Nevertheless, the works in [26] and [27] did not explore the inherent structure of the corresponding optimization problem to reduce its computational complexity, and [28] neglected the eavesdropping channel from the BS to the Eve via the hybrid RIS. Moreover, all proposed algorithms in [26], [27] and [28] aimed at the fixed hybrid RIS, which cannot be directly extended to the dynamic hybrid RIS case. Motivated by the above facts, in this paper, we aim to investigate a hybrid active-passive RIS assisted downlink secure MISOME system, where both fixed and dynamic hybrid RIS architectures are considered. Specifically, we jointly optimize the transmit covariance matrix at the BS, the hybrid RIS reflection matrix and the element allocation matrix for the dynamic hybrid RIS to maximize the secrecy rate, subject to both the transmit power budget at the BS and the amplification power budget at the hybrid RIS. The main

contributions of this work are summarized as follows.

- Firstly, in order to cope with the NP-hard SRM problems for fixed and dynamic hybrid RIS empowered secure communication systems, we validate the rank-1 property of the optimal BS transmit covariance matrix, based on which the original optimization problem can be equivalently simplified as a tractable fractional programming (FP) problem. Then, we propose a two-loop SCA-based iterative algorithm to jointly optimize the BS transmit beamforming vector, the reflection matrices of active and passive sub-RISs, as well as the element allocation matrix for the dynamic hybrid RIS.
- Secondly, for the fixed hybrid RIS, the optimal BS transmit beamforming vector can be derived depending on the tightness of the double power constraints at the optimum. Furthermore, we jointly apply the Dinkelbach's method and the SCA technique to convert the nonconvex FP problem into a tractable one, based on which the optimal semi-closed-form solutions of reflection matrices of active and passive sub-RISs are available through the Lagrangian dual theory and the first-order Taylor's series approximation, respectively.
- Finally, in order to solve the mixed-integer programming problem for the dynamic hybrid RIS, we firstly utilize an exponent-based continuous relaxation method to relax it into a continuous one, where the relaxation becomes tighter with the increasing exponent. As such, the proposed two-loop SCA-based iterative algorithm is still applicable. Numerical results demonstrate the superior secrecy performance of our proposed algorithms for both fixed and dynamic hybrid RISs. Particularly, the dynamic hybrid RIS is able to strike a flexible trade-off between the beamforming gain and the power amplification gain by adjusting the active/passive elements allocation.

Notations: Vectors and matrices are denoted by boldface lowercase letters and boldface uppercase letters, respectively. $\mathbb{R}_+$ and $\mathbb{H}_+^N$ denote the sets containing all nonnegative real numbers and all Hermitian positive semidefinite matrices of size $N \times N$, respectively. $\mathbb{C}^{N_1 \times N_2}$ denotes the space of $N_1 \times N_2$ complex-valued matrices. $\mathcal{CN}(\mathbf{0}, \mathbf{I})$ denotes the distribution of circularly symmetric complex Gaussian (CSCG) with zero mean and unit covariance matrix $\mathbf{I}$. $\text{Tr}(\mathbf{A})$, $\text{rank}(\mathbf{A})$ and $\det(\mathbf{A})$ denote the trace, rank and determinant of the matrix $\mathbf{A}$, respectively. $\mathbf{I}_N$ represents an identity matrix of size $N \times N$. The operator $\text{vecd}(\mathbf{A})$ denotes stacking the diagonal entries of $\mathbf{A}$ into a vector. $\text{diag}(\mathbf{a})$ and $\text{diag}(\mathbf{A})$ represent two diagonal matrices with diagonal elements being the elements of $\mathbf{a}$ and being those of $\mathbf{A}$, respectively. $[\mathbf{a}]_n$ denotes the n -th element of the vector $\mathbf{a}$, while $[\mathbf{\Omega}]_{m,n}$ denotes the (m, n) -th entry of the matrix $\mathbf{\Omega}$. $\mathbf{u}(\mathbf{A}, \mathbf{B})$ denotes the normalized dominant eigenvector of the matrix pencil $(\mathbf{A}, \mathbf{B})$. $|\mathbf{a}|$ denotes the amplitude vector and $\angle \mathbf{a}$ represents the phase vector. $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^H$ represent the conjugate, transpose and conjugate transpose of a vector or a matrix, respectively. $\|\mathbf{a}\|_2$ denotes the 2-norm for the vector $\mathbf{a}$ and $\|\mathbf{A}\|_F$ denotes the Frobenius norm of the matrix $\mathbf{A}$. $\otimes$ and $\odot$ denote the Kronecker and Hadamard product of two matrices,

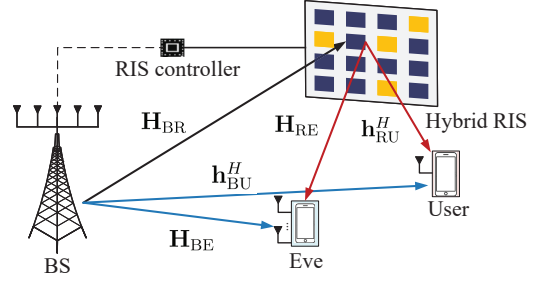


Fig. 2. A hybrid RIS assisted downlink secure communication system.

respectively. $\Re(\cdot)$ operates on a complex value and returns its real part. The phrase “with respect to” is abbreviated as “w.r.t.”.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. System Model

As illustrated in Fig. 2, we consider a hybrid RIS-empowered secure downlink communication system, where the BS equipped with N_t antennas sends confidential information to a single-antenna legitimate user with the aid of a hybrid RIS composed of M reflecting elements denoted by the set $\mathcal{M} = \{1, \dots, M\}$. Meanwhile, an Eve equipped with N_e antennas attempts to intercept the transmitted downlink signal from the BS. The hybrid RIS is assumed to be composed of M_a active and M_p passive elements, satisfying $M = M_a + M_p$, and operate in a full-duplex mode. To be specific, its involved active elements denoted by the set $\mathcal{A} \subset \mathcal{M}$ are able to simultaneously amplify and reflect the incident signals through reflection amplifiers and phase shifters, and thus suffer from non-negligible amplification noise. In contrast, its involved passive elements implemented by PIN diodes without RF components are usually free of thermal noise and intend to tune the phase shifts of incident signals. We define the reflection matrix of the hybrid RIS as a diagonal matrix, i.e., $\Psi = \text{diag}(\psi_1, \psi_2, \dots, \psi_M)$, where $\psi_m, \forall m \in \mathcal{M}$ denotes the reflecting coefficient of the m th hybrid RIS element and is modeled as

$$\psi_m = \begin{cases} a_m e^{j\theta_m}, & \text{if } m \in \mathcal{A} \\ e^{j\theta_m}, & \text{otherwise} \end{cases}, \quad (1)$$

where a_m denotes the amplitude of the m th hybrid RIS reflecting element if it is active, and θ_m denotes the phase shift of the m th hybrid RIS reflecting element no matter it is active or passive. Inspired by (1), we consider decomposing Ψ into the following two separate diagonal matrices Ψ_a and Ψ_p , which represent the reflection matrices of the active sub-RIS and the passive sub-RIS, respectively, i.e.,

$$\Psi = \underbrace{\text{diag}(\mathbf{1}_M^{\mathcal{A}})}_{\Psi_a} \Psi + \underbrace{(\mathbf{I}_M - \text{diag}(\mathbf{1}_M^{\mathcal{A}}))}_{\Psi_p} \Psi, \quad (2)$$

where $\mathbf{1}_M^{\mathcal{A}}$ denotes an M -dimensional vector whose non-zero elements are all one and their indexes are determined by the set $\mathcal{A}$. It follows from (2) that the hybrid RIS reduces to the fully-passive (fully-active) RIS when $\mathcal{A} = \emptyset$ ($\mathcal{A} = \mathcal{M}$). Denote by $\mathbf{H}_{BR} \in \mathbb{C}^{M \times N_t}$, $\mathbf{h}_{BU}^H \in \mathbb{C}^{1 \times N_t}$ ($\mathbf{H}_{BE} \in \mathbb{C}^{N_e \times N_t}$), and $\mathbf{h}_{RU}^H \in \mathbb{C}^{1 \times M}$ ($\mathbf{H}_{RE} \in \mathbb{C}^{N_e \times M}$) the baseband equivalent channels from the BS to the RIS, from the BS to the legitimate user (Eve), and from the RIS to the legitimate user (Eve),

respectively, which are all assumed to be quasi-static flat-fading and remain constant over the whole transmission block. We also suppose that the global CSI of the whole system is perfectly available at the BS for the sake of characterizing the system performance limit [29].

Let $\mathbf{x} \in \mathbb{C}^{N_t \times 1}$ denote the transmitted signal vector at the BS, and thus the transmit covariance matrix is defined as $\mathbf{Q} = \mathbb{E}[\mathbf{x}\mathbf{x}^H] \in \mathbb{C}^{N_t \times N_t}$. Moreover, we consider a total transmit power constraint at the BS, i.e., $\text{Tr}(\mathbf{Q}) \leq P_t$ with P_t being the maximum transmit power budget. Then, the signal reflected by the hybrid RIS can be expressed as

$$\mathbf{y}_{\text{RIS}} = \mathbf{\Psi} \mathbf{H}_{\text{BR}} \mathbf{x} + \mathbf{\Psi}_a \mathbf{n}_r, \quad (3)$$

where $\mathbf{n}_r \sim \mathcal{CN}(0, \sigma_r^2 \mathbf{I}_M)$ models both thermal noise and self-interference induced by the amplification operation at M potential active elements, with σ_r^2 denoting the average noise power. Let us define the maximum amplification power at the hybrid RIS as P_r , and then we have

$$\text{Tr}(\mathbf{\Psi}_a (\mathbf{H}_{\text{BR}} \mathbf{Q} \mathbf{H}_{\text{BR}}^H + \sigma_r^2 \mathbf{I}_M) \mathbf{\Psi}_a^H) + M_a P_{dc} \leq P_r, \quad (4)$$

where P_{dc} is the DC power required to drive each active element [30]. The signals reflected by the hybrid RIS more than once are negligible due to the large path loss. Consequently, the received signals at the legitimate user and the Eve are respectively written as

$$\begin{aligned} y_u &= (\mathbf{h}_{\text{BU}}^H + \mathbf{h}_{\text{RU}}^H \mathbf{\Psi} \mathbf{H}_{\text{BR}}) \mathbf{x} + \mathbf{h}_{\text{RU}}^H \mathbf{\Psi}_a \mathbf{n}_r + n_u, \\ y_e &= (\mathbf{H}_{\text{BE}} + \mathbf{H}_{\text{RE}} \mathbf{\Psi} \mathbf{H}_{\text{BR}}) \mathbf{x} + \mathbf{H}_{\text{RE}} \mathbf{\Psi}_a \mathbf{n}_r + \mathbf{n}_e, \end{aligned} \quad (5)$$

where $n_u \sim \mathcal{CN}(0, \sigma_u^2)$ and $\mathbf{n}_e \sim \mathcal{CN}(0, \sigma_e^2 \mathbf{I}_{N_e})$ denote the additive white Gaussian noise (AWGN) at the legitimate user and the Eve, respectively. Based on (5), the secrecy rate of the considered hybrid RIS empowered system is derived as [31]

$$R_s(\mathbf{Q}, \mathbf{\Psi}_a, \mathbf{\Psi}_p) = [R_u - R_e]^+, \quad (6)$$

where $[a]^+ = \max\{a, 0\}$, and R_u and R_e are respectively given as

$$\begin{aligned} R_u &= \log_2(1 + (\mathbf{h}_{\text{BU}}^H + \mathbf{h}_{\text{RU}}^H \mathbf{\Psi} \mathbf{H}_{\text{BR}}) \mathbf{Q} (\mathbf{h}_{\text{BU}} + \mathbf{H}_{\text{BR}}^H \mathbf{\Psi}^H \mathbf{h}_{\text{RU}}) K_{\mathbf{\Psi}_a}^{-1}), \\ R_e &= \log_2 \det(\mathbf{I}_{N_e} + (\mathbf{H}_{\text{BE}} + \mathbf{H}_{\text{RE}} \mathbf{\Psi} \mathbf{H}_{\text{BR}}) \mathbf{Q} \\ &\quad \times (\mathbf{H}_{\text{BE}} + \mathbf{H}_{\text{RE}} \mathbf{\Psi} \mathbf{H}_{\text{BR}})^H \mathbf{R}_{\mathbf{\Psi}_a}^{-1}), \end{aligned} \quad (7)$$

where $K_{\mathbf{\Psi}_a} = \sigma_r^2 \mathbf{h}_{\text{RU}}^H \mathbf{\Psi}_a \mathbf{\Psi}_a^H \mathbf{h}_{\text{RU}} + \sigma_u^2$ and $\mathbf{R}_{\mathbf{\Psi}_a} = \sigma_r^2 \mathbf{H}_{\text{RE}} \mathbf{\Psi}_a \mathbf{\Psi}_a^H \mathbf{H}_{\text{RE}}^H + \sigma_e^2 \mathbf{I}_{N_e}$. From an information theoretical perspective, the legitimate user can obtain confidential information from the BS at the rate of $R_s(\mathbf{Q}, \mathbf{\Psi}_a, \mathbf{\Psi}_p)$, whereas the Eve almost cannot retrieve any useful information from its received signal.

B. Problem Formulation

In this paper, we aim to jointly optimize the BS transmit covariance matrix $\mathbf{Q}$, the active sub-RIS reflection matrix $\mathbf{\Psi}_a$ and the passive sub-RIS reflection matrix $\mathbf{\Psi}_p$ at the hybrid RIS to maximize the secrecy rate $R_s(\mathbf{Q}, \mathbf{\Psi}_a, \mathbf{\Psi}_p)$ of the considered system, which is mathematically formulated as

$$\max_{\mathbf{Q} \succeq \mathbf{0}, \mathbf{\Psi}_a, \mathbf{\Psi}_p} R_s(\mathbf{Q}, \mathbf{\Psi}_a, \mathbf{\Psi}_p) \quad (8a)$$

$$\text{s.t.} \quad \text{Tr}(\mathbf{Q}) \leq P_t, \quad (8b)$$

$$\text{Tr}(\mathbf{\Psi}_a (\mathbf{H}_{\text{BR}} \mathbf{Q} \mathbf{H}_{\text{BR}}^H + \sigma_r^2 \mathbf{I}_M) \mathbf{\Psi}_a^H) + M_a P_{dc} \leq P_r, \quad (8c)$$

$$|[\mathbf{\Psi}_p]_{m,m}| = 1, \quad \forall m \in \mathcal{M} \setminus \mathcal{A}, \quad (8d)$$

where (8b) and (8c) respectively represent the transmit power constraints at the BS and the hybrid RIS, and (8d) denotes the unit-modulus constraints imposed on the passive sub-RIS. Obviously, problem (8) is nonconvex and challenging to solve due to the following reasons. Firstly, the optimization variables $\mathbf{Q}$, $\mathbf{\Psi}_a$ and $\mathbf{\Psi}_p$ are nonlinearly coupled in both the objective function (8a) and the constraint (8c). Secondly, the nonconvex unit-modulus constraint (8d) renders problem (8) more complicated. In the following section, we firstly explore the inherent characteristics of the optimal solutions with the aid of an upper-bound approximation, based on which problem (8) is greatly simplified and becomes much easier to handle. Then, an efficient two-loop AO algorithm is proposed to obtain a high-quality suboptimal solution to problem (8).

III. JOINT TRANSMIT BEAMFORMING AND FIXED HYBRID RIS OPTIMIZATION

In this section, the SRM problem (8) can be equivalently converted into a more tractable problem by exploring the inherent characteristics of the optimal solutions. Then, we develop an AO algorithm to jointly optimize the BS transmit covariance matrix $\mathbf{Q}$, the active sub-RIS reflection matrix $\mathbf{\Psi}_a$ and the passive sub-RIS reflection matrix $\mathbf{\Psi}_p$ under the fixed allocation between active and passive elements at the hybrid RIS.

A. Equivalent Transformation

To tackle the nonconvex objective function of problem (8), we firstly apply the determinant inequality $\det(\mathbf{I} + \mathbf{A}) \geq 1 + \text{Tr}(\mathbf{A})$ for any $\mathbf{A} \succeq \mathbf{0}$, where the equality holds if and only if $\text{rank}(\mathbf{A}) = 1$, to provide an upper-bound approximation of problem (8) as follows.

$$\max_{\mathbf{Q} \succeq \mathbf{0}, \mathbf{\Psi}_a, \mathbf{\Psi}_p} \frac{1 + \text{Tr}(\mathbf{Q} \mathbf{A}_{\Psi})}{1 + \text{Tr}(\mathbf{Q} \mathbf{B}_{\Psi})} \quad \text{s.t.} \quad (8b) \sim (8d), \quad (9)$$

where $\mathbf{A}_{\Psi} = K_{\mathbf{\Psi}_a}^{-1} (\mathbf{h}_{\text{BU}} + \mathbf{H}_{\text{BR}}^H \mathbf{\Psi}^H \mathbf{h}_{\text{RU}}) (\mathbf{h}_{\text{BU}}^H + \mathbf{h}_{\text{RU}}^H \mathbf{\Psi} \mathbf{H}_{\text{BR}})$ with $\text{rank}(\mathbf{A}_{\Psi}) = 1$ and $\mathbf{B}_{\Psi} = (\mathbf{H}_{\text{BE}} + \mathbf{H}_{\text{RE}} \mathbf{\Psi} \mathbf{H}_{\text{BR}})^H \mathbf{R}_{\mathbf{\Psi}_a}^{-1} (\mathbf{H}_{\text{BE}} + \mathbf{H}_{\text{RE}} \mathbf{\Psi} \mathbf{H}_{\text{BR}})$. Although the approximated fractional problem (9) is still nonconvex on $\{\mathbf{Q}, \mathbf{\Psi}_a, \mathbf{\Psi}_p\}$, we readily find that problem (9) with any given $\{\mathbf{\Psi}_a, \mathbf{\Psi}_p\}$ is quasi-linear on $\mathbf{Q}$, whose globally optimal solution is usually available by using the general quasi-convex optimization techniques such as the golden section search and the bisection method [32], [33]. However, such methods are usually very complex and cannot provide any useful insights into the optimal solutions. In order to address these issues, we further explore the inherent characteristics of the optimal solution to problem (9).

Proposition 1. *The optimal $\mathbf{Q}^*$ to problem (9) satisfies $\text{rank}(\mathbf{Q}^*) = 1$, which implies that problem (9) is actually a tight relaxation of the original SRM problem (8).*

Proof. Please see Appendix A. $\square$

According to **Proposition 1**, we can equivalently reexpress the optimal rank-one transmit covariance matrix as $\mathbf{Q} = \mathbf{q} \mathbf{q}^H$,

where $\mathbf{q}$ denotes the transmit beamforming vector for guaranteeing $\text{rank}(\mathbf{Q}) = 1$. Then, by substituting the rank-one solution into the relaxed SRM problem (9), we can obtain

$$\max_{\mathbf{q}, \Psi_a, \Psi_p} \frac{1 + \mathbf{q}^H \mathbf{A} \Psi \mathbf{q}}{1 + \mathbf{q}^H \mathbf{B} \Psi \mathbf{q}} \quad (10a)$$

$$\text{s.t. (8d), } \mathbf{q}^H \mathbf{q} \leq P_t, \quad (10b)$$

$$\text{Tr}(\Psi_a(\mathbf{H}_{\text{BR}} \mathbf{q} \mathbf{q}^H \mathbf{H}_{\text{BR}}^H + \sigma_r^2 \mathbf{I}_M) \Psi_a^H) + M_a P_{dc} \leq P_r. \quad (10c)$$

Thus, we have equivalently simplified the original SRM problem (8) into a constrained fractional quadratic problem (CFQP). However, problem (10) is still nonconvex due to the strongly coupled variables $\{\mathbf{q}, \Psi_a, \Psi_p\}$ and the unit-modulus constraint (8d). Hence, we develop an AO algorithm to find a high-quality suboptimal solution to problem (10). Furthermore, by exploring the inherent characteristics of problem (10), we have the following lemma.

Lemma 1. *There is at least one power constraint (i.e., (10b) or (10c)) being active at the optimum of problem (10).*

Proof. Please see Appendix B. $\square$

It is worth noting that **Lemma 1** can be leveraged to further simplify the optimization of $\mathbf{q}$, as elaborated in the following section.

B. Optimization of Transmit Beamforming Vector

With any given Ψ_a and Ψ_p , the subproblem of optimizing $\mathbf{q}$ is expressed as

$$\max_{\mathbf{q}} \frac{1 + \mathbf{q}^H \mathbf{A} \Psi \mathbf{q}}{1 + \mathbf{q}^H \mathbf{B} \Psi \mathbf{q}} \quad \text{s.t. (10b), (10c).} \quad (11)$$

It is well-known that problem (11) can be well addressed by jointly leveraging Charnes-Cooper transformation and SDR method. However, such methods are usually very complex and cannot provide any useful insights. By recalling **Lemma 1**, at least one power constraint is active at the optimum of problem (11). Hence, we can gain more insights into the optimal solution of $\mathbf{q}$, depending on the tightness of the constraints (10b) and (10c), which is detailed in the following proposition.

Proposition 2. *The optimal $\mathbf{q}^*$ to problem (11) can be summarized as the following three cases: tight (10b) and loose (10c), tight (10c) and loose (10b), as well as tight (10b) and (10c). Specifically, considering only one tight constraint, the optimal closed-form $\mathbf{q}^*$ is obtained as*

$$\mathbf{q}^* = \begin{cases} \sqrt{P_t} \mathbf{u}(\mathbf{I}_{N_t} + P_t \mathbf{A} \Psi, \mathbf{I}_{N_t} + P_t \mathbf{B} \Psi), & \text{tight (10b), loose (10c)} \\ \sqrt{P_c} \mathbf{u}(\mathbf{C} \Psi + P_q \mathbf{A} \Psi, \mathbf{C} \Psi + P_q \mathbf{B} \Psi), & \text{tight (10c), loose (10b)} \end{cases} \quad (12)$$

where $\mathbf{C} \Psi = \mathbf{H}_{\text{BR}}^H \Psi_a^H \Psi_a \mathbf{H}_{\text{BR}}$, $P_q = P_r - M_a P_{dc} - \sigma_r^2 \text{Tr}(\Psi_a \Psi_a^H)$ and $P_c = P_q / (\mathbf{u}^H(\mathbf{C} \Psi + P_q \mathbf{A} \Psi, \mathbf{C} \Psi + P_q \mathbf{B} \Psi) \mathbf{C} \Psi \mathbf{u}(\mathbf{C} \Psi + P_q \mathbf{A} \Psi, \mathbf{C} \Psi + P_q \mathbf{B} \Psi))$. Whereas for the case of both tight constraints (10b) and (10c), the optimal $\mathbf{q}^*$ can be obtained by recalling $\mathbf{Q} = \mathbf{q} \mathbf{q}^H$ and then solving the following semidefinite programming (SDP) problem [34].

$$\max_{\tilde{\mathbf{Q}} \succeq 0} \text{Tr}(\tilde{\mathbf{Q}}^H (\mathbf{I}_{N_t} + P_t \mathbf{A} \Psi)) \quad (13a)$$

$$\text{s.t. } \text{Tr}(\tilde{\mathbf{Q}}^H (\mathbf{I}_{N_t} + P_t \mathbf{B} \Psi)) = 1, \quad (13b)$$

$$\text{Tr}(\tilde{\mathbf{Q}}^H (P_t \mathbf{C} \Psi - P_q \mathbf{I}_{N_t})) = 0. \quad (13c)$$

Accordingly, the optimal $\mathbf{q}^*$ is given by $\mathbf{q}^* = \sqrt{\frac{P_t \bar{\lambda}_{\mathbf{q}}}{\text{Tr}(\tilde{\mathbf{Q}}^*)}} \mathbf{v}_{\mathbf{q}}$, where $\mathbf{v}_{\mathbf{q}}$ is the normalized eigenvector associated with the maximum eigenvalue $\bar{\lambda}_{\mathbf{q}}$ of the matrix $\tilde{\mathbf{Q}}^*$.

Proof. Please see Appendix C. $\square$

It is worth noting that the optimal $\mathbf{q}^*$ to problem (11) with any fixed Ψ_a, Ψ_p has been summarized in **Proposition 2**. Next, we turn our attention towards deriving the optimal active sub-RIS reflection matrix Ψ_a and the optimal passive sub-RIS reflection matrix Ψ_p to problem (10).

C. Optimization of Reflection Matrices of Active and Passive Sub-RISs

In this subsection, we consider the optimization of Ψ_a and Ψ_p with the obtained $\mathbf{q}$. To be specific, the subproblem of jointly optimizing Ψ_a and Ψ_p is expressed as

$$\max_{\Psi_a, \Psi_p} \frac{1 + K_{\Psi_a}^{-1} |c + \mathbf{h}_{\text{RU}}^H \Psi \mathbf{H}_{\text{BR}} \mathbf{q}|^2}{1 + (\mathbf{a} + \mathbf{H}_{\text{RE}} \Psi \mathbf{H}_{\text{BR}} \mathbf{q})^H \mathbf{R}_{\Psi_a}^{-1} (\mathbf{a} + \mathbf{H}_{\text{RE}} \Psi \mathbf{H}_{\text{BR}} \mathbf{q})} \quad (14)$$

s.t. (8d), (10c),

where $c = \mathbf{h}_{\text{BU}}^H \mathbf{q}$ and $\mathbf{a} = \mathbf{H}_{\text{BE}} \mathbf{q}$. Inspired by the diagonal structure of the reflection matrices Ψ_a and Ψ_p , we introduce two novel variables $\psi_a = \text{vecd}(\Psi_a)$ and $\psi_p = \text{vecd}(\Psi_p)$ to reformulate problem (14) as

$$\max_{\psi_a, \psi_p} \frac{1 + K_{\Psi_a}^{-1} |c^* + \psi^H \mathbf{b}|^2}{1 + (\mathbf{a} + \mathbf{G} \psi)^H \mathbf{R}_{\Psi_a}^{-1} (\mathbf{a} + \mathbf{G} \psi)} \quad (15a)$$

$$\text{s.t. } \psi_a^H \Omega_{\text{BR}} \psi_a \leq P_r - M_a P_{dc}, \quad (15b)$$

$$|\psi_p]_m| = 1, \quad \forall m = 1, \dots, M_p, \quad (15c)$$

where $\psi = \psi_a + \psi_p$, $\Omega_{\text{BR}} = \text{diag}(\mathbf{H}_{\text{BR}} \mathbf{Q} \mathbf{H}_{\text{BR}}^H + \sigma_r^2 \mathbf{I}_M)$, $\mathbf{b} = \text{diag}(\mathbf{q}^H \mathbf{H}_{\text{BR}}^H) \mathbf{h}_{\text{RU}}$ and $\mathbf{G} = \mathbf{H}_{\text{RE}} \text{diag}(\mathbf{H}_{\text{BR}} \mathbf{q})$. Owing to the nonconvex fractional objective function (15a) and the unit-modulus constraint (15c), problem (15) is still nonconvex and challenging to solve. To effectively tackle this problem, we propose to jointly optimize ψ_a and ψ_p by applying the Dinkelbach's method and the SCA technique, which is elaborated in the following section.

D. Proposed Method to Solve Problem (15)

In order to obtain the optimal solutions to problem (15), we firstly consider employing the Dinkelbach's method to transform its fractional objective function (15a) into the following equivalent function in a parametric subtractive form.

$$F(\gamma, \psi_a, \psi_p) = 1 + K_{\Psi_a}^{-1} |c^* + \psi^H \mathbf{b}|^2 - \gamma (1 + (\mathbf{a} + \mathbf{G} \psi)^H \mathbf{R}_{\Psi_a}^{-1} (\mathbf{a} + \mathbf{G} \psi)), \quad (16)$$

where γ is the introduced auxiliary variable. It's readily inferred that the function (16) is still jointly nonconcave on ψ_a and ψ_p . Therefore, we consider applying the SCA technique to find a locally tight lower bound of $F(\psi_a, \psi_p, \gamma)$, as shown at the bottom of the next page, where the superscript (t_{out}) denotes the index of the outer iterations and $\bar{\lambda}_{\mathbf{G}}$ denotes the maximum eigenvalue of the matrix $\mathbf{G}^H \mathbf{G}$. Other auxiliary parameters are defined as

$$\mathbf{A}_{\text{RE}} = \mathbf{G} - \sigma_e^{-2} \mathbf{H}_{\text{RE}} \text{diag}(\mathbf{m}^{(t_{\text{out}})}), \quad \Omega_{\text{RU}} = \text{diag}(\mathbf{h}_{\text{RU}} \mathbf{h}_{\text{RU}}^H),$$

$$\begin{aligned}
\mathbf{p}_1^{(t_{\text{out}})} &= \sigma_e^{-2}(\bar{\lambda}_G \mathbf{I}_M - \mathbf{G}^H \mathbf{G}) \boldsymbol{\psi}_p^{(t_{\text{out}})} - \sigma_e^{-2} \mathbf{G}^H \mathbf{a}, \\
w^{(t_{\text{out}})} &= (c + \mathbf{b}^H \boldsymbol{\psi}^{(t_{\text{out}})})(K_{\Psi_a}^{(t_{\text{out}})})^{-1}, \\
c_1^{(t_{\text{out}})} &= 2\Re\{w^{(t_{\text{out}})} c^*\} - |w^{(t_{\text{out}})}|^2 \sigma_u^2 + 1, \\
c_2^{(t_{\text{out}})} &= -\sigma_e^{-2} \bar{\lambda}_G M_p - \sigma_e^{-2} (\boldsymbol{\psi}_p^{(t_{\text{out}})})^H (\bar{\lambda}_G \mathbf{I}_M - \mathbf{G}^H \mathbf{G}) \boldsymbol{\psi}_p^{(t_{\text{out}})}, \\
\mathbf{m}^{(t_{\text{out}})} &= (\sigma_r^{-2} \mathbf{I}_M + \sigma_e^{-2} (\boldsymbol{\Psi}_a^{(t_{\text{out}})})^H \mathbf{H}_{\text{RE}}^H \mathbf{H}_{\text{RE}} \boldsymbol{\Psi}_a^{(t_{\text{out}})})^{-1} \\
&\quad \times (\boldsymbol{\Psi}_a^{(t_{\text{out}})})^H \mathbf{H}_{\text{RE}}^H (\mathbf{a} + \mathbf{G} \boldsymbol{\psi}^{(t_{\text{out}})}). \tag{18}
\end{aligned}$$

The equality (a) in (17) holds due to the Woodbury equality $(\mathbf{A} + \mathbf{UCV})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{U}(\mathbf{C}^{-1} + \mathbf{VA}^{-1} \mathbf{U})^{-1} \mathbf{VA}^{-1}$, while (b) holds due to $\text{Tr}(\mathbf{X}^H \mathbf{Y}^{-1} \mathbf{X}) \geq 2\Re\{\text{Tr}((\mathbf{X}^{(t_{\text{out}})})^H (\mathbf{Y}^{(t_{\text{out}})})^{-1} \mathbf{X})\} - \text{Tr}((\mathbf{Y}^{(t_{\text{out}})})^{-1} \mathbf{X}^{(t_{\text{out}})} (\mathbf{X}^{(t_{\text{out}})})^H (\mathbf{Y}^{(t_{\text{out}})})^{-1} \mathbf{Y})$ [35]. Moreover, the inequality (c) holds based on the first-order Taylor series expansion of the function $\boldsymbol{\psi}_p^H (\bar{\lambda}_G \mathbf{I}_M - \mathbf{G}^H \mathbf{G}) \boldsymbol{\psi}_p$ at the local point $\boldsymbol{\psi}_p^{(t_{\text{out}})}$.

Then, with any given $\boldsymbol{\psi}_a$ and γ , we can readily observe from (17c) that the optimal closed-form $\boldsymbol{\psi}_p$ is given by

$$\boldsymbol{\psi}_p^* = e^{j\angle(\gamma \mathbf{p}_1^{(t_{\text{out}})} - \gamma \sigma_e^{-2} \mathbf{G}^H \mathbf{A}_{\text{RE}} \boldsymbol{\psi}_a + w^{(t_{\text{out}})} \mathbf{b})}. \tag{19}$$

Substituting the optimal $\boldsymbol{\psi}_p^*$ into (17c) further yields

$$\begin{aligned}
\tilde{F}(\gamma, \boldsymbol{\psi}_a, \boldsymbol{\psi}_p^*) &= -\boldsymbol{\psi}_a^H (\gamma \sigma_e^{-2} \mathbf{A}_{\text{RE}}^H \mathbf{A}_{\text{RE}} + |w^{(t_{\text{out}})}|^2 \sigma_r^2 \boldsymbol{\Omega}_{\text{RU}}) \boldsymbol{\psi}_a \\
&\quad + 2 \sum_{m \in \mathcal{M} \setminus \mathcal{A}} |\gamma \sigma_e^{-2} [\mathbf{G}^H \mathbf{A}_{\text{RE}}]_{m,:} \boldsymbol{\psi}_a - [\gamma \mathbf{p}_1^{(t_{\text{out}})} + w^{(t_{\text{out}})} \mathbf{b}]_m| \\
&\quad - 2\gamma \sigma_e^{-2} \Re\{\boldsymbol{\psi}_a^H \mathbf{A}_{\text{RE}}^H \mathbf{a}\} + 2w^{(t_{\text{out}})} \Re\{\boldsymbol{\psi}_a^H \mathbf{b}\} + c_1^{(t_{\text{out}})} \\
&\quad + \gamma(c_2^{(t_{\text{out}})} - 1 - \sigma_e^{-4} \sigma_r^{-2} \|\mathbf{m}^{(t_{\text{out}})}\|_2^2). \tag{20}
\end{aligned}$$

Thus, a tractable lower bound approximation of problem (15) can be reexpressed as

$$\max_{\gamma, \boldsymbol{\psi}_a} \tilde{F}(\gamma, \boldsymbol{\psi}_a, \boldsymbol{\psi}_p^*) \quad \text{s.t.} \quad (15b). \tag{21}$$

It is obvious that problem (21) is still nonconvex due to the strongly coupled variables γ and $\boldsymbol{\psi}_a$. As such, we develop an efficient inner-loop AO algorithm to tackle problem (21).

1) *Optimization of γ* : With any given $\boldsymbol{\psi}_a$ and the obtained closed-form $\boldsymbol{\psi}_p$, the optimal auxiliary variable γ can be obtained according to the Dinkelbach's update rule, which is given by [36]

$$\gamma^* = \frac{1 + (K_{\Psi_a}^{(t_{\text{in}})})^{-1} |c^* + (\boldsymbol{\psi}^{(t_{\text{in}})})^H \mathbf{b}_2|^2}{1 + (\mathbf{a} + \mathbf{G} \boldsymbol{\psi}^{(t_{\text{in}})})^H (\mathbf{R}_{\Psi_a}^{(t_{\text{in}})})^{-1} (\mathbf{a} + \mathbf{G} \boldsymbol{\psi}^{(t_{\text{in}})})}, \tag{22}$$

where the superscript (t_{in}) denotes the iteration index of this AO procedure.

2) *Optimization of $\boldsymbol{\psi}_a$* : With the obtained γ , we optimize $\boldsymbol{\psi}_a$ in problem (21). To be specific, since the second term in the right-hand side of (20) is convex, we consider taking the first-order Taylor series expansion of it at the local point $\boldsymbol{\psi}_a^{(t_{\text{in}})}$ to obtain a concave lower bound of $\tilde{F}(\boldsymbol{\psi}_a, \boldsymbol{\psi}_p^*; \gamma)$ as

$$\tilde{F}(\boldsymbol{\psi}_a, \boldsymbol{\psi}_p^*; \gamma) \geq -\boldsymbol{\psi}_a^H \mathbf{B}_a \boldsymbol{\psi}_a + 2\Re\{\boldsymbol{\psi}_a^H \mathbf{p}_2^{(t_{\text{in}})}\} + c_3^{(t_{\text{in}})}, \tag{23a}$$

where the auxiliary parameters are defined as

$$\begin{aligned}
\tilde{\mathbf{p}}_1^{(t_{\text{in}})} &= \gamma \sigma_e^{-2} [\mathbf{G}^H \mathbf{A}_{\text{RE}}]_{m,:} \boldsymbol{\psi}_a^{(t_{\text{in}})} - [\gamma \mathbf{p}_1^{(t_{\text{out}})} + w^{(t_{\text{out}})} \mathbf{b}]_m, \\
\mathbf{B}_a &= \gamma \sigma_e^{-2} \mathbf{A}_{\text{RE}}^H \mathbf{A}_{\text{RE}} + |w^{(t_{\text{out}})}|^2 \sigma_r^2 \boldsymbol{\Omega}_{\text{RU}}, \\
\mathbf{p}_2^{(t_{\text{in}})} &= \sum_{m \in \mathcal{M} \setminus \mathcal{A}} \gamma \sigma_e^{-2} e^{-j\angle \tilde{\mathbf{p}}_1^{(t_{\text{in}})}} [\mathbf{G}^H \mathbf{A}_{\text{RE}}]_{m,:} - \gamma \sigma_e^{-2} \mathbf{A}_{\text{RE}}^H \mathbf{a} + w^{(t_{\text{out}})} \mathbf{b}, \\
c_3^{(t_{\text{in}})} &= c_1^{(t_{\text{out}})} + \gamma(c_2^{(t_{\text{out}})} - 1 - \sigma_e^{-4} \sigma_r^{-2} \|\mathbf{m}^{(t_{\text{out}})}\|_2^2) + 2 \sum_{m \in \mathcal{M} \setminus \mathcal{A}} |\tilde{\mathbf{p}}_1^{(t_{\text{in}})}| \\
&\quad - 2\gamma \sigma_e^{-2} \sum_{m \in \mathcal{M} \setminus \mathcal{A}} \Re\{e^{-j\angle \tilde{\mathbf{p}}_1^{(t_{\text{in}})}} [\mathbf{G}^H \mathbf{A}_{\text{RE}}]_{m,:} \boldsymbol{\psi}_a^{(t_{\text{in}})}\}. \tag{24}
\end{aligned}$$

Following the lower-bound approximation in (23a), we further stack the non-zero elements of $\boldsymbol{\psi}_a$ into a novel vector $\tilde{\boldsymbol{\psi}}_a \in \mathbb{C}^{M_a \times 1}$, based on which the subproblem w.r.t. $\tilde{\boldsymbol{\psi}}_a$ can be ultimately recast as

$$\min_{\tilde{\boldsymbol{\psi}}_a} \tilde{\boldsymbol{\psi}}_a^H \tilde{\mathbf{B}}_a \tilde{\boldsymbol{\psi}}_a - 2\Re\{\tilde{\boldsymbol{\psi}}_a^H \tilde{\mathbf{p}}_2^{(t_{\text{in}})}\} \tag{25a}$$

$$\text{s.t.} \quad \tilde{\boldsymbol{\psi}}_a^H \tilde{\boldsymbol{\Omega}}_{\text{BR}} \tilde{\boldsymbol{\psi}}_a \leq P_r - M_a P_{dc}, \tag{25b}$$

where $\tilde{\mathbf{B}}_a \in \mathbb{C}^{M_a \times M_a}$ ($\tilde{\boldsymbol{\Omega}}_{\text{BR}} \in \mathbb{C}^{M_a \times M_a}$) contains M_a rows and M_a columns of $\mathbf{B}_a$ ($\boldsymbol{\Omega}_{\text{BR}}$), with their corresponding indexes determined by the set $\mathcal{A}$. The vector $\tilde{\mathbf{p}}_2^{(t_{\text{in}})} \in \mathbb{C}^{M_a \times 1}$ consists of M_a elements of $\mathbf{p}_2^{(t_{\text{in}})}$, whose indexes are also from the set $\mathcal{A}$. It can be readily inferred that problem (25) is a standard quadratically constrained quadratic programming (QCQP) problem, whose optimal solution can be derived in closed form by leveraging the Lagrangian dual theory. Specifically, by taking the first-order derivative of the Lagrangian function of problem (25) to zero, we obtain

$$\tilde{\boldsymbol{\psi}}_a^* = (\tilde{\mathbf{B}}_a + \lambda_a \tilde{\boldsymbol{\Omega}}_{\text{BR}})^{-1} \tilde{\mathbf{p}}_2^{(t_{\text{in}})}, \tag{26}$$

where $\lambda_a \geq 0$ is the dual variable associated with the constraint (25b), which can be determined by using the bisection search method [37].

In a nutshell, we propose a two-loop SCA-based iterative algorithm to solve the original problem (8). First of all, we initially set the amplitudes of all elements $\{a_m | m \in \mathcal{M}\}$ to 1 and randomly generate their phases $\{\theta_m | m \in \mathcal{M}\}$

$$F(\gamma, \boldsymbol{\psi}_a, \boldsymbol{\psi}_p) \stackrel{(a)}{=} \gamma \sigma_e^{-4} (\mathbf{a} + \mathbf{G} \boldsymbol{\psi})^H \mathbf{H}_{\text{RE}} \boldsymbol{\Psi}_a (\sigma_r^{-2} \mathbf{I}_M + \sigma_e^{-2} \boldsymbol{\Psi}_a^H \mathbf{H}_{\text{RE}}^H \mathbf{H}_{\text{RE}} \boldsymbol{\Psi}_a)^{-1} \boldsymbol{\Psi}_a^H \mathbf{H}_{\text{RE}}^H (\mathbf{a} + \mathbf{G} \boldsymbol{\psi}) + K_{\Psi_a}^{-1} |c^* + \boldsymbol{\psi}^H \mathbf{b}|^2 + 1 - \gamma - \gamma \sigma_e^{-2} (\mathbf{a} + \mathbf{G} \boldsymbol{\psi})^H (\mathbf{a} + \mathbf{G} \boldsymbol{\psi}) \tag{17a}$$

$$\stackrel{(b)}{\geq} -\gamma \sigma_e^{-2} \|\mathbf{a} + \mathbf{G} \boldsymbol{\psi}_p + \mathbf{A}_{\text{RE}} \boldsymbol{\psi}_a\|_2^2 - |w^{(t_{\text{out}})}|^2 \sigma_r^2 \boldsymbol{\psi}_a^H \boldsymbol{\Omega}_{\text{RU}} \boldsymbol{\psi}_a + 2\Re\{w^{(t_{\text{out}})} \boldsymbol{\psi}^H \mathbf{b}\} - \gamma \sigma_e^{-4} \sigma_r^{-2} \|\mathbf{m}^{(t_{\text{out}})}\|_2^2 - \gamma + c_1^{(t_{\text{out}})} \tag{17b}$$

$$\stackrel{(c)}{\geq} 2\Re\{\boldsymbol{\psi}_p^H (\gamma \mathbf{p}_1^{(t_{\text{out}})} - \gamma \sigma_e^{-2} \mathbf{G}^H \mathbf{A}_{\text{RE}} \boldsymbol{\psi}_a + w^{(t_{\text{out}})} \mathbf{b})\} - \gamma \sigma_e^{-2} \|\mathbf{A}_{\text{RE}} \boldsymbol{\psi}_a + \mathbf{a}\|^2 + 2\Re\{w^{(t_{\text{out}})} \boldsymbol{\psi}_a^H \mathbf{b}\} - |w^{(t_{\text{out}})}|^2 \sigma_r^2 \boldsymbol{\psi}_a^H \boldsymbol{\Omega}_{\text{RU}} \boldsymbol{\psi}_a + c_1^{(t_{\text{out}})} + \gamma(c_2^{(t_{\text{out}})} - 1 - \sigma_e^{-4} \sigma_r^{-2} \|\mathbf{m}^{(t_{\text{out}})}\|_2^2) = \tilde{F}(\gamma, \boldsymbol{\psi}_a, \boldsymbol{\psi}_p), \tag{17c}$$

Algorithm 1: Proposed two-loop SCA-based iterative algorithm for solving problem (9)

Input: Initial $\bar{\Psi}^{(0)}$, index of outer iterations $t_{\text{out}} = 0$, index of inner iterations $t_{\text{in}} = 0$, maximum iteration numbers $T_{\text{out}}, T_{\text{in}}$, convergence threshold $\epsilon = 10^{-4}$;

Output: The BS transmit covariance matrix $\mathbf{Q}$, the active sub-RIS reflection matrix $\bar{\Psi}_a$ and the passive sub-RIS reflection matrix $\bar{\Psi}_p$;

```

1 repeat
2    $t_{\text{out}} = t_{\text{out}} + 1$ ;
3   Update  $\mathbf{q}^{(t_{\text{out}})}$  based on Proposition 2;
4   repeat
5      $t_{\text{in}} = t_{\text{in}} + 1$ ;
6     Update  $\gamma^{(t_{\text{in}})}$  based on (22);
7     Update  $\psi_a^{(t_{\text{in}})}$  based on (26);
8   until  $t_{\text{in}} = T_{\text{in}}$  or  $|E_{\text{in}}^{(t_{\text{in}})} - E_{\text{in}}^{(t_{\text{in}}-1)}| < \epsilon E_{\text{in}}^{(t_{\text{in}}-1)}$ ;
9   Update  $\bar{\Psi}_a^{(t_{\text{out}})}$  based on  $\bar{\Psi}_a = \text{diag}(\psi_a^{(t_{\text{in}})})$ ;
10  Update  $\bar{\Psi}_p^{(t_{\text{out}})}$  based on (19) and
     $\bar{\Psi}_p = \text{diag}(\psi_p^{(t_{\text{in}})})$ ;
11 until  $t_{\text{out}} = T_{\text{out}}$  or  $|E_{\text{out}}^{(t_{\text{out}})} - E_{\text{out}}^{(t_{\text{out}}-1)}| < \epsilon E_{\text{out}}^{(t_{\text{out}}-1)}$ ;
12 Update  $\mathbf{Q} = \mathbf{q}\mathbf{q}^H$ .
```

from the set $[0, 2\pi)$. Then, in the inner-loop, we iteratively update the introduced auxiliary variable γ and the reflection coefficients of the active sub-RIS ψ_a with the given transmit beamforming vector $\mathbf{q}$ and the derived optimal closed-form reflection coefficients of the passive sub-RIS ψ_p . In the outer-loop, a high-quality suboptimal solution to problem (8) can be obtained by alternately optimizing $\mathbf{q}$ and $\{\bar{\Psi}_a, \bar{\Psi}_p\}$. Ultimately, the optimal $\mathbf{Q}$ can be obtained based on $\mathbf{Q} = \mathbf{q}\mathbf{q}^H$. The proposed algorithm is summarized in **Algorithm 1**, where E_{in} and E_{out} denote the objective values of problem (21) and problem (10), respectively. The precision ϵ is set as 10^{-4} . T_{out} and T_{in} denote the maximum numbers of iterations in the outer and inner loops, respectively.

IV. JOINT TRANSMIT BEAMFORMING AND DYNAMIC HYBRID RIS OPTIMIZATION

In this section, we mainly consider the dynamic hybrid RIS architecture and redefine the hybrid RIS reflection matrix as $\bar{\Psi}$, where the allocation of active/passive elements closely related to the propagation environment is additionally regarded as optimization variable. Specifically, we introduce an indicator matrix to represent the dynamic allocation of active/passive elements as

$$[\Lambda]_{i,j} = \begin{cases} 1, & \text{if } i = j \text{ and } i \in \mathcal{A} \\ 0, & \text{otherwise} \end{cases}, \quad (27)$$

where $[\Lambda]_{i,i} = 1$ or 0 indicates that the i th RIS element is active or passive. Then, by recalling (2), the reflection matrices of the active and passive sub-RISs are respectively redefined as

$$\bar{\Psi} = \underbrace{\Lambda \bar{\Psi}}_{\bar{\Psi}_a} + \underbrace{(\mathbf{I}_M - \Lambda) \bar{\Psi}}_{\bar{\Psi}_p}. \quad (28)$$

Moreover, it follows from **Proposition 1** that the optimal $\mathbf{Q}^*$ is also rank one, i.e., $\mathbf{Q} = \mathbf{q}\mathbf{q}^H$. Motivated by the above facts,

the original SRM problem (10) for the dynamic hybrid RIS can be formulated as

$$\max_{\mathbf{q}, \bar{\Psi}_a, \bar{\Psi}_p, \Lambda} g(\mathbf{q}, \bar{\Psi}_a, \bar{\Psi}_p) \quad (29a)$$

$$\text{s.t.} \quad (10b), \text{Tr}(\bar{\Psi}_a(\mathbf{H}_{\text{BR}}\mathbf{q}\mathbf{q}^H\mathbf{H}_{\text{BR}}^H + \sigma_r^2\mathbf{I}_M)\bar{\Psi}_a^H) + \sum_m [\Lambda]_{m,m} P_{dc} \leq P_r, \quad (29b)$$

$$|[\bar{\Psi}_p]_{m,m}| = 1 - [\Lambda]_{m,m}, \quad \forall m \in \mathcal{M}, \quad (29c)$$

$$[\Lambda]_{m,m} = \{0, 1\}, \quad \forall m \in \mathcal{M}. \quad (29d)$$

where $K_{\bar{\Psi}_a} = \sigma_r^2 \mathbf{h}_{\text{RU}}^H \bar{\Psi}_a \bar{\Psi}_a^H \mathbf{h}_{\text{RU}} + \sigma_u^2$ and $\mathbf{R}_{\bar{\Psi}_a} = \sigma_r^2 \mathbf{H}_{\text{RE}} \bar{\Psi}_a \bar{\Psi}_a^H \mathbf{H}_{\text{RE}}^H + \sigma_e^2 \mathbf{I}_{N_e}$ and

$$g(\mathbf{q}, \bar{\Psi}_a, \bar{\Psi}_p) = \frac{1 + K_{\bar{\Psi}_a}^{-1} |\mathbf{h}_{\text{BU}}^H \mathbf{q} + \mathbf{h}_{\text{RU}}^H \bar{\Psi} \mathbf{H}_{\text{BR}} \mathbf{q}|^2}{1 + (\mathbf{H}_{\text{BE}} \mathbf{q} + \mathbf{H}_{\text{RE}} \bar{\Psi} \mathbf{H}_{\text{BR}} \mathbf{q})^H \mathbf{R}_{\bar{\Psi}_a}^{-1} (\mathbf{H}_{\text{BE}} \mathbf{q} + \mathbf{H}_{\text{RE}} \bar{\Psi} \mathbf{H}_{\text{BR}} \mathbf{q})}. \quad (30)$$

In order to obtain a high-quality solution to the nonconvex mixed-integer fractional optimization problem (29), we firstly introduce a diagonal and positive semi-definite matrix $\mathbf{V} \in \mathbb{H}_+^M$ to represent the amplification power of all M potential active elements. Furthermore, by recalling the element allocation matrix Λ , we have

$$|[\bar{\Psi}_a]_{m,m}|^2 = [\Lambda]_{m,m} \mathbf{V}_{m,m}, \quad \forall m. \quad (31)$$

Then, we additionally define $K_{\mathbf{V}} = \sigma_r^2 \mathbf{h}_{\text{RU}}^H \mathbf{V} \mathbf{h}_{\text{RU}} + \sigma_u^2$ and combine it with (31) to convert problem (29) into

$$\max_{\mathbf{q}, \bar{\Psi}_a, \bar{\Psi}_p, \mathbf{V}} \tilde{g}(\mathbf{q}, \bar{\Psi}_a, \bar{\Psi}_p, \mathbf{V}, \Lambda) \quad (32a)$$

$$\text{s.t.} \quad (10b), (29b) \sim (29d),$$

$$|[\bar{\Psi}_a]_{m,m}|^2 \leq [\Lambda]_{m,m} [\mathbf{V}]_{m,m}, \quad \forall m \in \mathcal{M}, \quad (32b)$$

where

$$\tilde{g}(\mathbf{q}, \bar{\Psi}_a, \bar{\Psi}_p, \mathbf{V}) = \frac{1 + K_{\mathbf{V}}^{-1} |\mathbf{h}_{\text{BU}}^H \mathbf{q} + \mathbf{h}_{\text{RU}}^H \bar{\Psi} \mathbf{H}_{\text{BR}} \mathbf{q}|^2}{1 + (\mathbf{H}_{\text{BE}} \mathbf{q} + \mathbf{H}_{\text{RE}} \bar{\Psi} \mathbf{H}_{\text{BR}} \mathbf{q})^H \mathbf{R}_{\bar{\Psi}_a}^{-1} (\mathbf{H}_{\text{BE}} \mathbf{q} + \mathbf{H}_{\text{RE}} \bar{\Psi} \mathbf{H}_{\text{BR}} \mathbf{q})}. \quad (33)$$

It is worth noting that the equivalence between problem (32) and problem (29) is established in the following lemma.

Lemma 2. *Problem (32) is equivalent to problem (29). That is to say, they can achieve the same optimal solution.*

Proof. Please see Appendix D. $\square$

Unfortunately, problem (32) is still a mixed-integer optimization problem. In the following section, we intend to develop an exponent-based continuous relaxation method to efficiently solve problem (32).

A. A Tractable Continuous Relaxation of Problem (32)

In this subsection, we firstly relax the binary variables into bounded continuous variables to make the mixed-integer optimization problem (32) more tractable, i.e.,

$$0 \leq [\Lambda_{\text{rel}}]_{m,m} \leq 1, \quad \forall m \in \mathcal{M}. \quad (34)$$

Moreover, by referring to [38], we also consider imposing an exponent χ on $[\Lambda_{\text{rel}}]_{m,m}$, $\forall m$ for the sake of a tighter

continuous relaxation. Therefore, problem (32) can be relaxed into

$$\max_{\mathbf{q}, \bar{\Psi}_a, \bar{\Psi}_p, \mathbf{V}, \Lambda_{\text{rel}}} \tilde{g}(\mathbf{q}, \bar{\Psi}_a, \bar{\Psi}_p, \mathbf{V}) \quad (35a)$$

$$\text{s.t.} \quad (10b), (29b), (29c), (34),$$

$$|[\bar{\Psi}_a]_{m,m}|^2 \leq [\Lambda_{\text{rel}}]_{m,m}^\chi [\mathbf{V}]_{m,m}, \quad \forall m \in \mathcal{M}. \quad (35b)$$

Intuitively, the exponent χ in problem (35) serves as a penalty parameter, which penalizes the values of $[\Lambda_{\text{rel}}]_{m,m}, \forall m \in \mathcal{M}$, thereby encouraging them towards binary solutions despite their continuous nature. Furthermore, according to [38], the optimal objective value of problem (35) becomes closer to that of the original mixed-integer problem (29) with the increasing χ , thus leading to a tighter relaxation. In fact, problem (35) is still nonconvex due to the strongly coupled variables and the nonconvex constraints (29c) and (35b). To tackle it effectively, the proposed efficient two-loop SCA-based iterative algorithm is also applicable, as elaborated in the following.

B. Proposed Method to Solve Problem (35)

In this subsection, the two-loop SCA-based iterative algorithm is also applied to solve problem (35). Specifically, we firstly optimize the transmit beamforming vector $\mathbf{q}$ with any given $\{\bar{\Psi}_a, \bar{\Psi}_p, \mathbf{V}, \Lambda_{\text{rel}}\}$, which can be directly obtained from **Proposition 2**. Subsequently, with the obtained $\mathbf{q}$, we focus on optimizing $\{\bar{\Psi}_a, \bar{\Psi}_p, \mathbf{V}, \Lambda_{\text{rel}}\}$ for the dynamic hybrid RIS. Similar to Section III-C, we introduce two auxiliary vectors $\bar{\psi}_a = \text{vecd}(\bar{\Psi}_a)$ and $\bar{\psi}_p = \text{vecd}(\bar{\Psi}_p)$, based on which the subproblem of optimizing $\{\bar{\Psi}_a, \bar{\Psi}_p, \mathbf{V}, \Lambda_{\text{rel}}\}$ can be simplified as

$$\max_{\bar{\psi}_a, \bar{\psi}_p, \mathbf{V}, \Lambda_{\text{rel}}} \frac{1 + K_{\mathbf{V}}^{-1} |c^* + \bar{\psi}^H \mathbf{b}|^2}{1 + (\mathbf{a} + \mathbf{G}\bar{\psi})^H \mathbf{R}_{\bar{\Psi}_a}^{-1} (\mathbf{a} + \mathbf{G}\bar{\psi})} \quad (36a)$$

$$\text{s.t.} \quad (34),$$

$$\bar{\psi}_a^H \Omega_{\text{BR}} \bar{\psi}_a + \sum_m [\Lambda_{\text{rel}}]_{m,m} P_{dc} \leq P_r, \quad (36b)$$

$$|[\bar{\psi}_p]_m| = 1 - [\Lambda_{\text{rel}}]_{m,m}, \quad \forall m \in \mathcal{M}, \quad (36c)$$

$$|[\bar{\psi}_a]_m|^2 \leq [\Lambda_{\text{rel}}]_{m,m}^\chi [\mathbf{V}]_{m,m}, \quad \forall m \in \mathcal{M}, \quad (36d)$$

where $\bar{\psi} = \bar{\psi}_a + \bar{\psi}_p$. It can be readily inferred that the fractional objective function (36a) and the constraints (36c), (36d) in problem (36) are all nonconvex. Similar to solving problem (15), we also consider applying the Dinkelbach's method to rewrite the fractional objective function into the following parametric subtractive form.

$$\bar{F}(\bar{\gamma}, \bar{\psi}_a, \bar{\psi}_p, \mathbf{V}) = 1 + K_{\mathbf{V}}^{-1} |c^* + \bar{\psi}^H \mathbf{b}|^2 - \bar{\gamma} (1 + (\mathbf{a} + \mathbf{G}\bar{\psi})^H \mathbf{R}_{\bar{\Psi}_a}^{-1} (\mathbf{a} + \mathbf{G}\bar{\psi})), \quad (37)$$

where $\bar{\gamma}$ is the introduced auxiliary variable. Then, by sequentially applying the same SCA techniques as (17a), (17b) and (17c), we can find a locally tight lower bound of the function $\bar{F}(\bar{\gamma}, \bar{\psi}_a, \bar{\psi}_p, \mathbf{V})$ as follows.

$$\begin{aligned} \bar{F}(\bar{\gamma}, \bar{\psi}_a, \bar{\psi}_p, \mathbf{V}) &\geq -\bar{\gamma} \sigma_e^{-2} \|\sqrt{\lambda_{\mathbf{G}}} \bar{\psi}_p + \frac{\mathbf{G}^H}{\sqrt{\lambda_{\mathbf{G}}}} (\bar{\mathbf{A}}_{\text{RE}} \bar{\psi}_a + \mathbf{a})\|_2^2 \\ &+ 2\Re\{\bar{\gamma} \sigma_e^{-2} \bar{\psi}_p^H (\bar{\lambda}_{\mathbf{G}} \mathbf{I}_M - \mathbf{G}^H \mathbf{G}) \bar{\psi}_p^{(t_{\text{out}})}\} - |\bar{w}^{(t_{\text{out}})}|^2 \sigma_r^2 \mathbf{h}_{\text{RU}}^H \mathbf{V} \mathbf{h}_{\text{RU}} \end{aligned}$$

$$\begin{aligned} & - \bar{\gamma} \sigma_e^{-2} (\bar{\mathbf{A}}_{\text{RE}} \bar{\psi}_a + \mathbf{a})^H (\mathbf{I}_{N_e} - \frac{\mathbf{G} \mathbf{G}^H}{\lambda_{\mathbf{G}}}) (\bar{\mathbf{A}}_{\text{RE}} \bar{\psi}_a + \mathbf{a}) - \bar{\gamma} \\ & + 2\Re\{\bar{w}^{(t_{\text{out}})} \bar{\psi}^H \mathbf{b}\} - \bar{\gamma} \sigma_e^{-4} \sigma_r^{-2} \|\bar{\mathbf{m}}^{(t_{\text{out}})}\|_2^2 + c_4^{(t_{\text{out}})} + \bar{\gamma} c_5^{(t_{\text{out}})} \\ & = \hat{F}(\bar{\gamma}, \bar{\psi}_a, \bar{\psi}_p, \mathbf{V}), \end{aligned} \quad (38)$$

where the involved auxiliary parameters are defined as

$$\begin{aligned} \bar{\mathbf{A}}_{\text{RE}} &= \mathbf{G} - \sigma_e^{-2} \mathbf{H}_{\text{RE}} \text{diag}(\bar{\mathbf{m}}^{(t_{\text{out}})}), \\ \bar{w}^{(t_{\text{out}})} &= (c + \mathbf{b}^H \bar{\psi}^{(t_{\text{out}})}) (K_{\mathbf{V}}^{(t_{\text{out}})})^{-1}, \\ \bar{\mathbf{m}}^{(t_{\text{out}})} &= (\sigma_r^{-2} \mathbf{I}_M + \sigma_e^{-2} (\bar{\Psi}_a^{(t_{\text{out}})})^H \mathbf{H}_{\text{RE}}^H \mathbf{H}_{\text{RE}} \bar{\Psi}_a^{(t_{\text{out}})})^{-1} \\ &\quad \times (\bar{\Psi}_a^{(t_{\text{out}})})^H \mathbf{H}_{\text{RE}}^H (\mathbf{a} + \mathbf{G} \bar{\psi}^{(t_{\text{out}})}), \\ c_4^{(t_{\text{out}})} &= 2\Re\{\bar{w}^{(t_{\text{out}})} c^*\} - |\bar{w}^{(t_{\text{out}})}|^2 \sigma_u^2 + 1, \\ c_5^{(t_{\text{out}})} &= -\sigma_e^{-2} (\bar{\psi}_p^{(t_{\text{out}})})^H (\bar{\lambda}_{\mathbf{G}} \mathbf{I}_M - \mathbf{G}^H \mathbf{G}) \bar{\psi}_p^{(t_{\text{out}})}. \end{aligned} \quad (39)$$

Thus, problem (36) can be reformulated as

$$\max_{\bar{\gamma}, \bar{\psi}_a, \bar{\psi}_p, \mathbf{V}, \Lambda_{\text{rel}}} \hat{F}(\bar{\gamma}, \bar{\psi}_a, \bar{\psi}_p, \mathbf{V}) \quad \text{s.t.} \quad (34), (36b) \sim (36d). \quad (40)$$

However, problem (40) is still nonconvex due to the coupled variables and nonconvex constraints (36c) and (36d). In order to ensure compliance with the constant-modulus constraint (36c), we introduce a unit-modulus constrained vector $\boldsymbol{\theta}$ to reexpress $\bar{\psi}_p$ as $\bar{\psi}_p = (\mathbf{I}_M - \Lambda_{\text{rel}}) \boldsymbol{\theta}$, based on which problem (40) can be recast as

$$\max_{\bar{\gamma}, \boldsymbol{\theta}, \bar{\psi}_a, \mathbf{V}, \Lambda_{\text{rel}}} \hat{F}(\bar{\gamma}, \boldsymbol{\theta}, \bar{\psi}_a, \mathbf{V}, \Lambda_{\text{rel}}) \quad (41a)$$

$$\text{s.t.} \quad (34), (36b), (36d), |[\boldsymbol{\theta}]_m| = 1, \quad \forall m \in \mathcal{M}, \quad (41b)$$

where

$$\begin{aligned} \hat{F}(\bar{\gamma}, \boldsymbol{\theta}, \bar{\psi}_a, \mathbf{V}, \Lambda_{\text{rel}}) &= -|\bar{w}^{(t_{\text{out}})}|^2 \sigma_r^2 \mathbf{h}_{\text{RU}}^H \mathbf{V} \mathbf{h}_{\text{RU}} \\ &- \bar{\gamma} \sigma_e^{-2} \|\sqrt{\lambda_{\mathbf{G}}} (\mathbf{I}_M - \Lambda_{\text{rel}}) \boldsymbol{\theta} + \frac{\mathbf{G}^H (\bar{\mathbf{A}}_{\text{RE}} \bar{\psi}_a + \mathbf{a})}{\sqrt{\lambda_{\mathbf{G}}}}\|_2^2 \\ &+ 2\Re\{\boldsymbol{\theta}^H (\mathbf{I}_M - \Lambda_{\text{rel}}) (\bar{w}^{(t_{\text{out}})} \mathbf{b})\} + 2\Re\{\bar{w}^{(t_{\text{out}})} \bar{\psi}_a^H \mathbf{b}\} \\ &- \bar{\gamma} \sigma_e^{-2} (\bar{\mathbf{A}}_{\text{RE}} \bar{\psi}_a + \mathbf{a})^H (\mathbf{I}_{N_e} - \frac{\mathbf{G} \mathbf{G}^H}{\lambda_{\mathbf{G}}}) (\bar{\mathbf{A}}_{\text{RE}} \bar{\psi}_a + \mathbf{a}) \\ &+ 2\Re\{\boldsymbol{\theta}^H (\mathbf{I}_M - \Lambda_{\text{rel}}) (\bar{\gamma} \sigma_e^{-2} (\bar{\lambda}_{\mathbf{G}} \mathbf{I}_M - \mathbf{G}^H \mathbf{G}) \bar{\psi}_p^{(t_{\text{out}})})\} \\ &- \bar{\gamma} - \bar{\gamma} \sigma_e^{-4} \sigma_r^{-2} \|\bar{\mathbf{m}}^{(t_{\text{out}})}\|_2^2 + c_4^{(t_{\text{out}})} + \bar{\gamma} c_5^{(t_{\text{out}})}. \end{aligned} \quad (42)$$

Unfortunately, the optimization variables $\{\bar{\gamma}, \boldsymbol{\theta}, \bar{\psi}_a, \mathbf{V}, \Lambda_{\text{rel}}\}$ are still strongly coupled in problem (41). Moreover, the unit-modulus constraint (41b) renders problem (41) difficult to solve. To address these issues, we next propose a SCA-based AO algorithm to solve problem (41) by alternately optimizing $\bar{\gamma}, \boldsymbol{\theta}$ and $\{\bar{\psi}_a, \mathbf{V}, \Lambda_{\text{rel}}\}$.

1) *Optimization of $\bar{\gamma}$* : Similar to (22), with any given $\boldsymbol{\theta}$ and $\{\bar{\psi}_a, \mathbf{V}, \Lambda_{\text{rel}}\}$, $\bar{\gamma}$ is updated by

$$\bar{\gamma} = \frac{1 + (K_{\mathbf{V}}^{(t_{\text{in}})})^{-1} |c^* + (\bar{\psi}^{(t_{\text{in}})})^H \mathbf{b}|^2}{1 + (\mathbf{a} + \mathbf{G} \bar{\psi}^{(t_{\text{in}})})^H (\mathbf{R}_{\bar{\Psi}_a}^{(t_{\text{in}})})^{-1} (\mathbf{a} + \mathbf{G} \bar{\psi}^{(t_{\text{in}})})}. \quad (43)$$

2) *Optimization of $\boldsymbol{\theta}$* : With the obtained $\bar{\gamma}$ and any given $\{\bar{\psi}_a, \mathbf{V}, \Lambda_{\text{rel}}\}$, we can obtain the closed-form solution of the optimal $\boldsymbol{\theta}^*$ as

$$\boldsymbol{\theta}^* = e^{j \angle \bar{\mathbf{p}}_3^{(t_{\text{in}})}}, \quad (44)$$

where $\bar{\mathbf{p}}_3^{(t_{\text{in}})} = (\mathbf{I}_M - \Lambda_{\text{rel}}^{(t_{\text{in}})}) (\bar{\gamma} \sigma_e^{-2} (\bar{\lambda}_{\mathbf{G}} \mathbf{I}_M - \mathbf{G}^H \mathbf{G}) \bar{\psi}_p^{(t_{\text{in}})} + \bar{w}^{(t_{\text{in}})} \mathbf{b} - \bar{\gamma} \sigma_e^{-2} \mathbf{G}^H (\bar{\mathbf{A}}_{\text{RE}} \bar{\psi}_a^{(t_{\text{in}})} + \mathbf{a}))$.

Algorithm 2: Proposed two-loop SCA-based iterative algorithm for solving problem (35)

Input: Initial $\bar{\Psi}^{(0)}$, $\Lambda_{\text{rel}}^{(0)}$, iteration indexes $t_{\text{out}} = 0$, $t_{\text{in}} = 0$, maximum iteration numbers T_{out} , T_{in} , convergence threshold $\epsilon = 10^{-4}$;

Output: The BS transmit covariance matrix $\mathbf{Q}$, the active sub-RIS reflection matrix $\bar{\Psi}_{\text{a}}$ and the passive sub-RIS reflection matrix $\bar{\Psi}_{\text{p}}$;

```

1 repeat
2    $t_{\text{out}} = t_{\text{out}} + 1$ ;
3   Update  $\mathbf{q}^{(t_{\text{out}})}$  based on Proposition 2;
4   repeat
5      $t_{\text{in}} = t_{\text{in}} + 1$ ;
6     Update  $\bar{\gamma}^{(t_{\text{in}})}$  based on (43);
7     Update  $\boldsymbol{\theta}$  based on (44);
8     Update  $\{\bar{\psi}_{\text{a}}, \mathbf{V}, \Lambda_{\text{rel}}\}$  by solving problem (47);
9   until  $t_{\text{in}} = T_{\text{in}}$  or  $|E_{\text{in}}^{(t_{\text{in}})} - E_{\text{in}}^{(t_{\text{in}}-1)}| < \epsilon E_{\text{in}}^{(t_{\text{in}}-1)}$ ;
10  Update  $\bar{\Psi}_{\text{a}}^{(t_{\text{out}})}$  based on  $\bar{\Psi}_{\text{a}} = \text{diag}(\bar{\psi}_{\text{a}}^{(t_{\text{in}})})$ ;
11  Update  $\bar{\Psi}_{\text{p}}^{(t_{\text{out}})}$  based on  $\bar{\Psi}_{\text{p}} = \text{diag}(\bar{\psi}_{\text{p}}^{(t_{\text{in}})})$ ;
12 until  $t_{\text{out}} = T_{\text{out}}$  or  $|E_{\text{out}}^{(t_{\text{out}})} - E_{\text{out}}^{(t_{\text{out}}-1)}| < \epsilon E_{\text{out}}^{(t_{\text{out}}-1)}$ ;
```

3) *Optimization of $\{\bar{\psi}_{\text{a}}, \mathbf{V}, \Lambda_{\text{rel}}\}$:* With the obtained $\bar{\gamma}$ and $\boldsymbol{\theta}$, the subproblem w.r.t. $\{\bar{\psi}_{\text{a}}, \mathbf{V}, \Lambda_{\text{rel}}\}$ can be written as

$$\max_{\bar{\psi}_{\text{a}}, \mathbf{V}, \Lambda_{\text{rel}}} \hat{F}(\bar{\psi}_{\text{a}}, \mathbf{V}, \Lambda_{\text{rel}}; \bar{\gamma}, \boldsymbol{\theta}) \quad \text{s.t.} \quad (34), (36\text{b}), (36\text{d}). \quad (45)$$

To address the nonconvexity of the constraint (36d), we firstly reformulate it as $\frac{|\bar{\psi}_{\text{a}}|_m^2}{|\mathbf{V}|_{m,m}} \leq [\Lambda_{\text{rel}}]_{m,m}^{\chi}, \forall m$, both sides of which are convex. Therefore, according to the idea of SCA, we construct a concave lower bound of the right-hand side function $[\Lambda_{\text{rel}}]_{m,m}^{\chi}$ by leveraging its first-order Taylor expansion at the local point $[\Lambda_{\text{rel}}]_{m,m}^{(t_{\text{in}})}$, which is expressed as

$$[\Lambda_{\text{rel}}]_{m,m}^{\chi} \geq \underbrace{(1-\chi)([\Lambda_{\text{rel}}]_{m,m}^{(t_{\text{in}})})^{\chi} + \chi([\Lambda_{\text{rel}}]_{m,m}^{(t_{\text{in}})})^{(\chi-1)}[\Lambda_{\text{rel}}]_{m,m}}_{h([\Lambda_{\text{rel}}]_{m,m})}, \forall m. \quad (46)$$

Based on (46), the subproblem (45) can be relaxed into

$$\max_{\bar{\psi}_{\text{a}}, \mathbf{V}, \Lambda_{\text{rel}}} \hat{F}(\bar{\psi}_{\text{a}}, \mathbf{V}, \Lambda_{\text{rel}}; \bar{\gamma}, \boldsymbol{\theta}) \quad (47\text{a})$$

$$\text{s.t.} \quad (34), (36\text{b}), \frac{|\bar{\psi}_{\text{a}}|_m^2}{|\mathbf{V}|_{m,m}} \leq h([\Lambda_{\text{rel}}]_{m,m}), \forall m. \quad (47\text{b})$$

It is evident that problem (47) is a strictly convex problem. Hence, we can apply the well-known CVX toolbox to obtain the optimal $\{\bar{\psi}_{\text{a}}, \mathbf{V}, \Lambda_{\text{rel}}\}$.

In summary, we develop a two-loop SCA-based iterative algorithm to tackle the nonconvex problem (35). The random initializations of $\bar{\Psi}_{\text{a}}$ and $\bar{\Psi}_{\text{p}}$ are the same as those of **Algorithm 1**. In addition, the initial values of $[\Lambda_{\text{rel}}]_{m,m}, \forall m \in \mathcal{M}$ are randomly selected from the range $(0, 1]$. Then, we iteratively update $\bar{\gamma}$, $\boldsymbol{\theta}$, $\{\bar{\psi}_{\text{a}}, \mathbf{V}, \Lambda_{\text{rel}}\}$ and $\bar{\psi}_{\text{p}}$ with the given $\mathbf{q}$ until convergence and finally recover $\bar{\Psi}_{\text{a}}$ and $\bar{\Psi}_{\text{p}}$ from the obtained

$\{\boldsymbol{\theta}, \bar{\psi}_{\text{a}}, \mathbf{V}, \Lambda_{\text{rel}}\}$. In the outer loop, we obtain a high-quality suboptimal solution to problem (35) by alternately optimizing $\mathbf{q}$, $\bar{\Psi}_{\text{a}}$ and $\bar{\Psi}_{\text{p}}$. The proposed algorithm is summarized in Algorithm 2, where E_{in} and E_{out} denote the objective values of problem (41) and problem (35), respectively.

C. Recovery of the Binary Solution Λ

After obtaining the relaxed continuous solution Λ_{rel} from **Algorithm 2**, we readily infer that the value of each $[\Lambda_{\text{rel}}]_{m,m}^{\chi}, m \in \mathcal{M}$ sufficiently approaches two discrete points of 0 and 1 according to the property of the exponent-based continuous relaxation method. Then, the corresponding m -th diagonal element of Λ can be accordingly set to 0 or 1. Ultimately, based on the determined Λ , we re-carry out **Algorithm 1** to obtain the optimal $\{\mathbf{Q}, \bar{\Psi}_{\text{a}}, \bar{\Psi}_{\text{p}}\}$ to the original SRM problem (29).

V. ANALYSIS AND EXTENSIONS

In this section, we discuss the convergence behaviors and computational complexity of the proposed **Algorithm 1** and **Algorithm 2**. Moreover, we also discuss the applicability of our proposed algorithms to the imperfect CSI case, as well as the multi-user and multi-Eve system.

A. Convergence and Complexity Analysis

Firstly, the convergence behaviours of the proposed **Algorithm 1** and **Algorithm 2** can be ensured by the following proposition.

Proposition 3. *The objective values of problems (9) and (35) are monotonically nondecreasing over the iterations by applying **Algorithm 1** and **Algorithm 2**, respectively.*

Proof. Please see Appendix E. $\square$

The complexities of our proposed **Algorithm 1** and **Algorithm 2** both mainly come from the matrix multiplication and the standard convex optimization algorithms. Specifically, the worst-case complexity of deriving the optimal transmit beamforming vector $\mathbf{q}$ is given by $\mathcal{O}((N_t^2)^{3.5} \log(1/\epsilon))$, which arises from solving the SDP problem (13) via the interior-point method embedded in the CVX toolbox [39]. For **Algorithm 1**, the complexities of updating γ and ψ_{a} in each inner iteration are respectively given by $\mathcal{O}(N_e M^2 + N_e^3)$ and $\mathcal{O}(M_a^3 + \log(B/\epsilon_1))$, where B and ϵ_1 denote the initial interval length and the desired accuracy of the bisection method, respectively. In the outer iteration, the complexity of updating ψ_{p} is calculated as $\mathcal{O}(N_e M^2)$. Hence, the total complexity of **Algorithm 1** is given by $\mathcal{O}(I_{\text{out}}((N_t^2)^{3.5} \log(1/\epsilon) + N_e M^2 + I_{\text{in}}(N_e M^2 + N_e^3 + M_a^3 + \log(B/\epsilon_1))))$, where I_{in} and I_{out} respectively denote the numbers of inner and outer iterations. Similarly, the total complexity of **Algorithm 2** is calculated as $\mathcal{O}(I_{\text{out}}((N_t^2)^{3.5} \log(1/\epsilon) + I_{\text{in}}(N_e M^2 + N_e^3 + (3M)^{3.5} \log(1/\epsilon))))$. The details are omitted here for brevity.

B. Extension to the Imperfect CSI case

In this subsection, considering that it is quite challenging to acquire the perfect CSI for the Eve in practice, we assume imperfect CSI of the channels $\mathbf{H}_{BE}$ and $\mathbf{H}_{RE}$, which can be modeled as [40],

$$\mathbf{H}_i = \hat{\mathbf{H}}_i + \Delta\mathbf{H}_i, \quad i \in \{BE, RE\}, \quad (48)$$

where $\hat{\mathbf{H}}_i$ denotes the estimated channel and $\Delta\mathbf{H}_i$ denotes the corresponding CSI error bounded by Frobenius norm, i.e., $\|\Delta\mathbf{H}_i\|_F \leq \epsilon_i, i \in \{BE, RE\}$. Given the estimated $\hat{\mathbf{H}}_{BE}$, $\hat{\mathbf{H}}_{RE}$, along with the perfectly known channels $\mathbf{H}_{BR}$, $\mathbf{h}_{BU}$ and $\mathbf{h}_{RU}$, we aim to jointly design the BS transmit covariance matrix $\mathbf{Q}$ and the reflection matrix of the hybrid active-passive RIS $\{\Psi_a, \Psi_p\}$ to maximize the system secrecy rate in the worst CSI error scenario, which is formulated as

$$\max_{\mathbf{Q} \succeq \mathbf{0}, \Psi_a, \Psi_p} R_u - \max_{\Delta\mathbf{H}_{BE}, \Delta\mathbf{H}_{RE}} R'_e \quad (49a)$$

$$\text{s.t.} \quad (8b) \sim (8d),$$

$$\|\Delta\mathbf{H}_i\|_F \leq \epsilon_i, i \in \{BE, RE\}, \quad (49b)$$

where $R'_e = \log_2 \det(\mathbf{I}_{N_e} + ((\hat{\mathbf{H}}_{BE} + \Delta\mathbf{H}_{BE}) + (\hat{\mathbf{H}}_{RE} + \Delta\mathbf{H}_{RE})\Psi\mathbf{H}_{BR})\mathbf{Q}((\hat{\mathbf{H}}_{BE} + \Delta\mathbf{H}_{BE}) + (\hat{\mathbf{H}}_{RE} + \Delta\mathbf{H}_{RE})\Psi\mathbf{H}_{BR})^H \mathbf{R}_{\Psi}^{-1})$. It is worth noting that the robust max-min problem (49) is more intractable to solve than problem (8), due to the complicated max-min objective function (49a) and the extra CSI errors (49b). Then, by introducing a slack variable τ , problem (49) can be converted into

$$\max_{\mathbf{Q} \succeq \mathbf{0}, \Psi_a, \Psi_p} R_u - \log_2(\tau), \quad (50a)$$

$$\text{s.t.} \quad (8b) \sim (8d),$$

$$R'_e \leq \log_2(\tau), \|\Delta\mathbf{H}_i\|_F \leq \epsilon_i, i \in \{BE, RE\}, \quad (50b)$$

Similar to (17), the nonconvex constraint (50b) can be converted into a tractable form by introducing the Woodbury equality and the inequality $\text{Tr}(\mathbf{X}^H \mathbf{Y}^{-1} \mathbf{X}) \geq 2\Re\{\text{Tr}((\mathbf{X}^{(t_{out})})^H (\mathbf{Y}^{(t_{out})})^{-1} \mathbf{X})\} - \text{Tr}((\mathbf{Y}^{(t_{out})})^{-1} \mathbf{X}^{(t_{out})} (\mathbf{X}^{(t_{out})})^H (\mathbf{Y}^{(t_{out})})^{-1} \mathbf{Y})$. Furthermore, based on the Schur complement and sign-definiteness lemmas [39], [41], we can derive an effective reformulation of constraint (50b) without the CSI errors. Then, the proposed SCA-based iterative algorithm can still be applied to solve the reformulated problem. However, this imperfect CSI case is out of the scope of this paper and we plan to address it in the future research.

C. Extension to the Multi-User and Multi-Eve System

In this subsection, we consider a general system with K user-Eve pairs, meaning that each user is wiretapped by a specific Eve. The set of the user-Eve pairs is defined as $\mathcal{K} = \{1, \dots, K\}$. We denote the confidential message intended for the k -th user and the corresponding precoding vector as $s_k \sim \mathcal{CN}(0, 1)$ and $\mathbf{w}_k$, respectively. Thus, the transmit signal at the BS is expressed as $\mathbf{x} = \sum_{k \in \mathcal{K}} \mathbf{w}_k s_k$. Denote by $\mathbf{h}_{BU,k}$ ($\mathbf{h}_{RU,k}$) and $\mathbf{H}_{BE,k}$ ($\mathbf{H}_{RE,k}$) the baseband equivalent channels from the BS (RIS) to the user k and the Eve k , respectively. Then, the sum secrecy rate (SSR) of the considered system is expressed as [42]

$$\tilde{R}_s(\{\mathbf{w}_k\}, \Psi_a, \Psi_p) = \sum_{k \in \mathcal{K}} [\tilde{R}_{u,k} - \tilde{R}_{e,k}]^+, \quad (51)$$

where $\tilde{R}_{u,k}$ and $\tilde{R}_{e,k}$ respectively denote the achievable rate at the k -th user and the k -th Eve for the confidential message s_k , shown at the bottom of this page. Therefore, the SSR maximization problem of the considered system can be reformulated as

$$\max_{\{\mathbf{w}_k\}, \Psi_a, \Psi_p} \tilde{R}_s(\{\mathbf{w}_k\}, \Psi_a, \Psi_p) \quad (53a)$$

$$\text{s.t.} \quad (8d), \quad \sum_{k \in \mathcal{K}} \text{Tr}(\mathbf{w}_k \mathbf{w}_k^H) \leq P_t, \quad (53b)$$

$$\sum_{k \in \mathcal{K}} \text{Tr}(\Psi_a (\mathbf{H}_{BR} \mathbf{w}_k \mathbf{w}_k^H \mathbf{H}_{BR}^H + \sigma_r^2 \mathbf{I}_M) \Psi_a^H) + M_a P_{dc} \leq P_r. \quad (53c)$$

Due to the the presence of the inter-user interference, the proposed **Algorithm 1** and **Algorithm 2** cannot be directly applied to solve problem (53). By referring to [43, Proposition 1], we can obtain the following equation by introducing a series of variables $\mathbf{u} = [u_1, \dots, u_K]^T$.

$$\begin{aligned} & -\log_2 \left(\sum_{j \in \mathcal{K}, j \neq k} |(\mathbf{h}_{BU,j}^H + \mathbf{h}_{RU,j}^H \Psi \mathbf{H}_{BR}) \mathbf{w}_j|^2 \right. \\ & \quad \left. + \sigma_r^2 \mathbf{h}_{RU}^H \Psi_a \Psi_a^H \mathbf{h}_{RU} + \sigma_u^2 \right) \\ & = \max_{u_k > 0} \left[-\frac{u_k}{\ln 2} \left(\sum_{j \in \mathcal{K}, j \neq k} |(\mathbf{h}_{BU,j}^H + \mathbf{h}_{RU,j}^H \Psi \mathbf{H}_{BR}) \mathbf{w}_j|^2 \right. \right. \\ & \quad \left. \left. + \sigma_r^2 \mathbf{h}_{RU}^H \Psi_a \Psi_a^H \mathbf{h}_{RU} + \sigma_u^2 \right) + \log_2(u_k) + \frac{1}{\ln 2} \right], \forall k. \quad (54) \end{aligned}$$

Based on (54), we can convert the log-fractional objective function (53a) into a more tractable form. An efficient AO strategy can still be applied to tackle the strong coupling among the optimization variables. Specifically, we can solve

$$\tilde{R}_{u,k} = \log_2 \left(1 + |(\mathbf{h}_{BU,k}^H + \mathbf{h}_{RU,k}^H \Psi \mathbf{H}_{BR}) \mathbf{w}_k|^2 \left(\sum_{j \in \mathcal{K}, j \neq k} |(\mathbf{h}_{BU,j}^H + \mathbf{h}_{RU,j}^H \Psi \mathbf{H}_{BR}) \mathbf{w}_j|^2 + \sigma_r^2 \mathbf{h}_{RU}^H \Psi_a \Psi_a^H \mathbf{h}_{RU} + \sigma_u^2 \right)^{-1} \right), \quad (52a)$$

$$\begin{aligned} \tilde{R}_{e,k} & = \log_2 \det \left(\mathbf{I}_{N_e} + (\mathbf{H}_{BE,k} + \mathbf{H}_{RE,k} \Psi \mathbf{H}_{BR}) \mathbf{w}_k \mathbf{w}_k^H (\mathbf{H}_{BE,k} + \mathbf{H}_{RE,k} \Psi \mathbf{H}_{BR})^H \right. \\ & \quad \left. \times \left(\sum_{j \in \mathcal{K}, j \neq k} (\mathbf{H}_{BE,k} + \mathbf{H}_{RE,k} \Psi \mathbf{H}_{BR}) \mathbf{w}_j \mathbf{w}_j^H (\mathbf{H}_{BE,k} + \mathbf{H}_{RE,k} \Psi \mathbf{H}_{BR})^H + \sigma_r^2 \mathbf{H}_{RE,k} \Psi_a \Psi_a^H \mathbf{H}_{RE,k} + \sigma_e^2 \mathbf{I}_{N_e} \right)^{-1} \right). \quad (52b) \end{aligned}$$

TABLE I
SIMULATION PARAMETER SETTINGS

Parameter	Notation, Value	Parameter	Notation, Value
Number of transmit antennas	$N_t = 6$	Number of antennas at Eve	$N_e = 4$
Number of elements at RIS	$M = 36$	Fading factor	$\beta_0 = -30$ dB
BS transmission power	$P_t = 20$ dBm	Power budget at RIS	$P_r = 6$ dBm
DC signal power at RIS	$P_{dc} = -10$ dBm	Gaussian noise at user	$\sigma_u^2 = -75$ dBm
Gaussian noise at Eve	$\sigma_e^2 = -75$ dBm	Amplification noise	$\sigma_r^2 = -70$ dBm
Coordinate of the BS	$(0, 0, 5)$	Coordinate of the RIS	$(10, 0, 5)$
Coordinate of the user	$(55, 2, 0)$	Coordinate of the Eve	$(45, 2, 0)$
Path-loss exponents	$\xi_{BU} = \xi_{BE} = 3.7$ $\xi_{RU} = \xi_{RE} = 2.5$ $\xi_{BR} = 2.2$	Rician factors	$\kappa_{BU} = \kappa_{BE} = 0$ $\kappa_{RU} = \kappa_{RE} = 1$ $\kappa_{BR} = \infty$

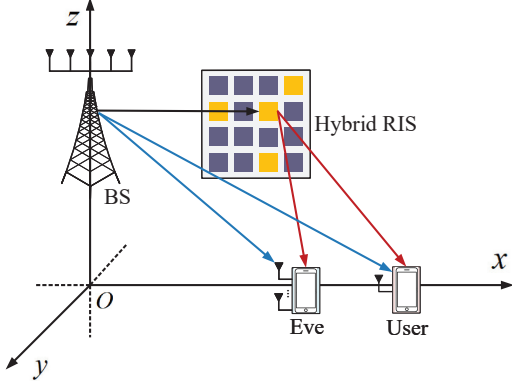


Fig. 3. A three-dimensional coordinate system.

the subproblems of optimizing $\{\mathbf{w}_k\}$ and $\{\Psi_a, \Psi_p\}$ by adopting the SDR method and the SCA method, respectively. Moreover, the exponent-based continuous relaxation method can be leveraged to optimize the active-passive element allocation matrix Λ for the case of the dynamic hybrid RIS. However, considering that problem (53) is out of the scope of this paper and we plan to investigate it in our future research works.

Remark 1: Considering a special case of a single-user and multi-Eve system, it can be inferred that **Proposition 1** also holds for optimizing the transmit covariance matrix, and thus the proposed two-loop SCA-based iterative algorithm is also applicable to solve the SRM problem of the considered system after some modifications.

VI. SIMULATION RESULTS

In this section, numerical simulation results are presented to validate the secrecy performance of the considered hybrid active-passive RIS assisted system. We assume that both the BS and Eve are equipped with uniform linear arrays (ULAs), while the hybrid RIS is equipped with a uniform planar array (UPA) of size $\sqrt{M} \times \sqrt{M}$. As illustrated in Fig. 3, we establish a three-dimensional coordinate system, where the BS and the hybrid RIS are assumed to be located at $(0, 0, z_B)\text{m}$ and $(x_R, 0, z_R)\text{m}$, respectively. The user and the Eve are located at $(x_U, y_U, z_U)\text{m}$ and $(x_E, y_E, z_E)\text{m}$, respectively. Moreover, each involved channel matrix is modeled as the Rician fading channel, i.e., $\mathbf{H} = \sqrt{\beta} \tilde{\mathbf{H}}$. $\beta = \beta_0 (\frac{d}{d_0})^{-\xi}$ denotes the distance-dependent large scale path-loss coefficient, where β_0 denotes the path loss at the reference distance $d_0 = 1\text{m}$; d represents the individual link distance and ξ denotes the path loss exponent. $\tilde{\mathbf{H}}$ denotes the small-scale Rician fading channel matrix, which is modeled as

$$\tilde{\mathbf{H}} = \sqrt{\frac{\kappa}{1+\kappa}} \mathbf{H}_{\text{LoS}} + \sqrt{\frac{1}{1+\kappa}} \mathbf{H}_{\text{NLoS}}, \quad (55)$$

where $\kappa \in [0, \infty)$ denotes the Rician factor. In terms of our work, the path loss exponents and the Rician factors of the BS-User, BS-Eve, RIS-User, RIS-Eve and BS-RIS channels are collected as $\{\xi_{BU}, \xi_{BE}, \xi_{RU}, \xi_{RE}, \xi_{BR}\}$ and $\{\kappa_{BU}, \kappa_{BE}, \kappa_{RU}, \kappa_{RE}, \kappa_{BR}\}$, respectively. For convenience, the specific values of the above simulation parameters are summarized in Table I. In addition, $\mathbf{H}_{\text{LoS}}$ and $\mathbf{H}_{\text{NLoS}}$ denote the LoS and NLoS components of the channel $\mathbf{H}$, respectively. Each element of $\mathbf{H}_{\text{NLoS}}$ is assumed to be Rayleigh fading, i.e., $[\mathbf{H}_{\text{NLoS}}]_{i,j} \sim \mathcal{CN}(0, 1), \forall i, j$, while the LoS component is modeled as the product of the steering vectors at the transmitter and the receiver. Specifically, the normalized transmit steering vector of a ULA is defined as $\mathbf{a}_i(\theta_i) = 1/\sqrt{N_t} [1, \dots, e^{j2\pi \frac{D}{\lambda} (n_t-1) \sin(\theta_i)}, \dots, e^{j2\pi \frac{D}{\lambda} (N_t-1) \sin(\theta_i)}]^T$, where $\mathbf{a}_i, i \in \{B, E\}$ stands for the steering vector at the BS or the Eve, and $n_t \in \{1, 2, \dots, N_t\}$ denotes the antenna index. D and λ denote the inter-antenna spacing and the wave-length, respectively. $\theta_i \in [0, 2\pi), i \in \{B, E\}$ represents the angle-of-departure (AoD) at the BS or the angle-of-arrival (AoA) at the Eve. Moreover, the normalized UPA steering vector is defined as $\mathbf{a}_R(\theta_R, \psi_R) = 1/\sqrt{M} [1, \dots, e^{j2\pi \frac{D}{\lambda} ((m_y-1) \sin(\theta_R) \sin(\psi_R) + (m_z-1) \cos(\psi_R))}, \dots, e^{j2\pi \frac{D}{\lambda} ((\sqrt{M}-1) \sin(\theta_R) \sin(\psi_R) + (\sqrt{M}-1) \cos(\psi_R))}]^T$, where $m_y (m_z) \in \{1, 2, \dots, \sqrt{M}\}$ denotes the antenna index in the horizontal (vertical) direction. $\theta_R \in [0, 2\pi)$ and $\psi_R \in [-\pi/2, \pi/2)$ denote the azimuth angle and the elevation angle at the hybrid RIS, respectively. Moreover, the BS (Eve) inter-antenna spacing of $\lambda/2$ wavelength and the RIS inter-element spacing of $\lambda/8$ wavelength are considered. Without loss of generality, we set the number of active elements at the fixed hybrid RIS and the exponent in **Algorithm 2** as $M_a = 4$ and $\chi = 2$, respectively. Unless otherwise specified, the basic simulation parameters are listed in Table I. All simulation results are obtained by averaging over 100 channel realizations. Furthermore, we compare our proposed algorithms with the following benchmark schemes.

- 1) **SDR-based fixed (dynamic) hybrid RIS:** In this scheme, we jointly utilize the Charnes-Cooper transformation and the SDR method to optimize the transmit beamforming vector and derive the optimal hybrid RIS reflection matrix based on **Algorithm 1** (**Algorithm 2**).
- 2) **CR-based dynamic hybrid RIS:** In this scheme, we derive the optimal transmit beamforming vector based on **Proposition 2** and optimize the dynamic RIS reflection matrix by employing the traditional continuous relaxation method with $\chi = 1$.

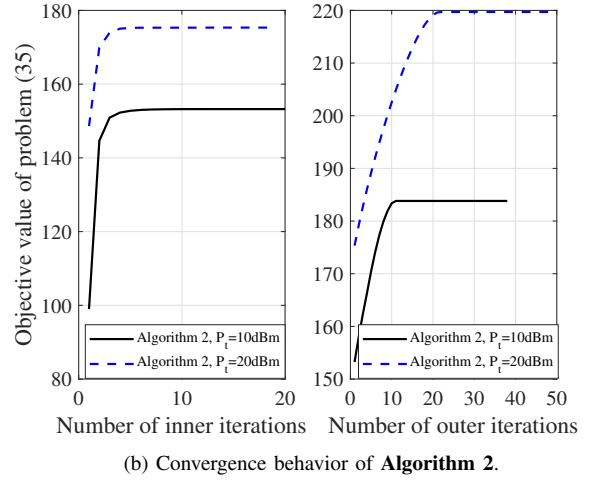
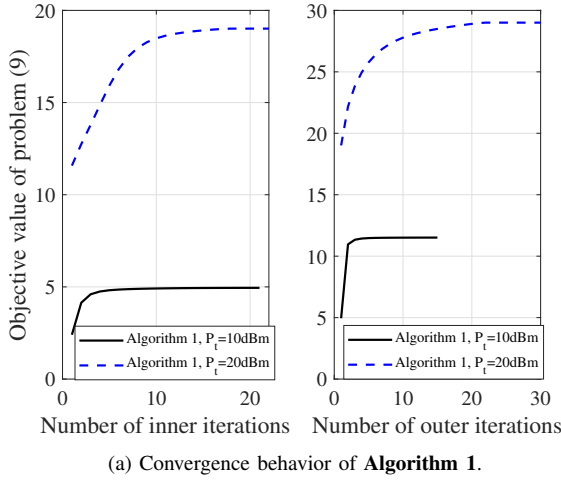


Fig. 4. Convergence behaviors of the proposed algorithms.

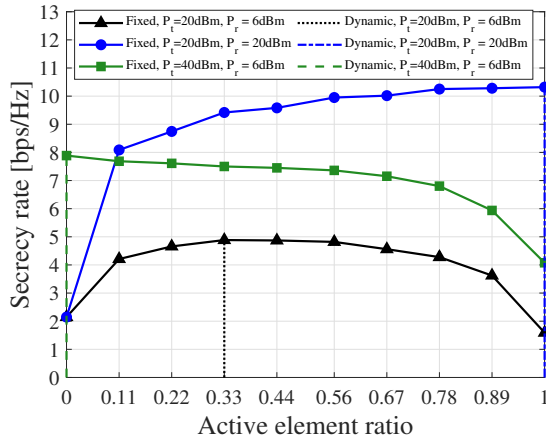


Fig. 5. Secrecy rate versus the active element ratio.

- 3) **Fully-passive (Fully-active) RIS:** In this scheme, we set $M_p = M(M_a = M)$ and optimize the RIS reflection matrix based on **Algorithm 1**.
- 4) **Random-phase RIS:** In this scheme, we consider the fully-passive RIS and randomly generate each RIS phase shift according to the distribution $\theta_m \in [0, 2\pi)$, $\forall m$.

Fig. 4(a) and Fig. 4(b) respectively show the convergence behaviors of the proposed **Algorithm 1** and **Algorithm 2** under different transmit power levels, i.e., $P_t = 10\text{dBm}$ and $P_t = 20\text{dBm}$. Specifically, the left subfigure in Fig. 4(a) (Fig. 4(b)) illustrates the convergence of the proposed **Algorithm 1**(**Algorithm 2**) in the inner loop of the first outer iteration, while the right subfigure demonstrates the convergence of the proposed algorithm in the outer loop. It can be observed from Fig. 4 that both **Algorithm 1** and **Algorithm 2** converge within 20 outer iterations under each considered P_t . Naturally, the optimal objective values achieved by these two algorithms both increase with P_t .

Fig. 5 illustrates the secrecy rates achieved by the fixed hybrid RIS scheme versus the active element ratio, defined as the ratio of the number of active elements to the total number of RIS elements under different BS and hybrid RIS transmit power budgets, i.e., $\{P_t, P_r\}$. In Fig. 5, the coordinates of the user and the Eve are set as $(75, 2, 0)$ and $(85, 2, 0)$, respectively. Moreover, the optimal active-passive ratios for

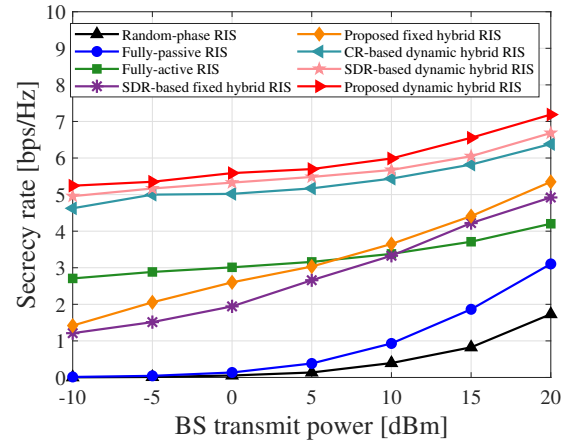


Fig. 6. Secrecy rate versus the BS transmit power P_t .

the dynamic hybrid RIS scheme under different values of P_t and P_r are marked in Fig. 5. It can be observed that when $P_t = 20\text{dBm}$ and $P_r = 6\text{dBm}$, both the fixed and dynamic hybrid RIS schemes achieve the same maximum secrecy rate at the same active element ratio of 0.33, which verifies the effectiveness of **Algorithm 2**. Furthermore, as P_t grows to 40dBm and P_r remains unchanged, the optimal active element ratio of both the fixed and dynamic hybrid RIS schemes decreases to 0, implying that the fully-passive RIS performs best among all schemes. When keeping $P_t = 20\text{dBm}$ and increasing P_r to 20dBm , the optimal active element ratios of both fixed and dynamic hybrid RIS schemes increase to 1, implying that the higher power budget on the hybrid RIS allows more elements to switch to active mode.

Fig. 6 plots the secrecy rates achieved by different schemes versus the BS transmit power P_t . It can be seen that the secrecy rate of each scheme naturally increases with P_t . In addition, it is clearly observed that the proposed fixed and dynamic hybrid RIS schemes respectively demonstrate superior secrecy performance over the corresponding SDR-based schemes, and the dynamic hybrid RIS scheme also achieves a higher secrecy rate than its CR-based counterpart. Owing to the additional power amplification gain, both the fixed hybrid RIS schemes and the fully-active RIS scheme show superior performance over the fully-passive RIS scheme. Naturally, the dynamic

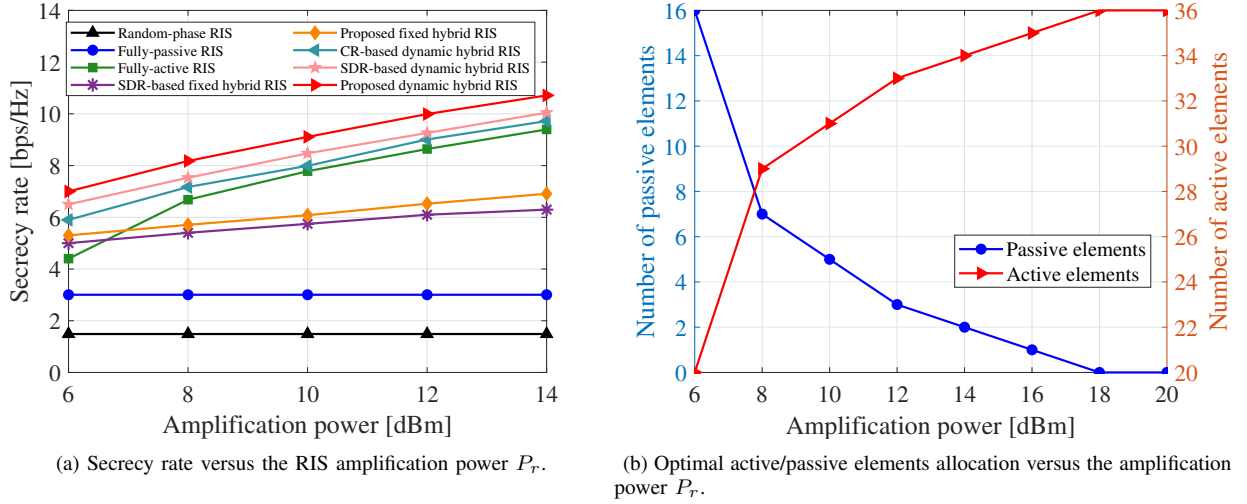


Fig. 7. Effect of amplification power P_r on the secrecy rate and the active/passive elements allocation.

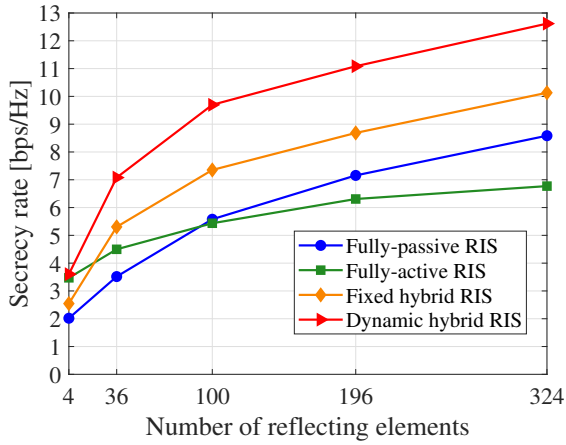


Fig. 8. Secrecy rate versus the number of RIS elements M .

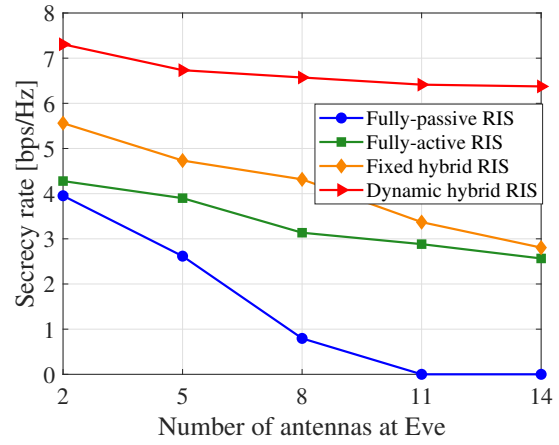


Fig. 9. Secrecy rate versus the number of antennas at Eve N_e .

hybrid RIS scheme achieves the highest secrecy rate among all schemes, which is attributed to the additional DoFs in the active-passive elements allocation. Moreover, when the transmit power is small, i.e., $P_t \in [-10, 5]$ dBm, the fully-active RIS outperforms the fixed hybrid RIS, whereas for a large P_t , i.e., $P_t \in [10, 20]$ dBm, the fixed hybrid RIS shows better secrecy performance. This implies that at a high transmit power, the passive beamforming gain becomes dominant on the achievable secrecy rate instead of the power amplification gain.

In Fig. 7, we investigate the effects of the hybrid RIS amplification power P_r on secrecy rates achieved by different schemes and the optimal active/passive elements allocation. From Fig. 7(a), it is evident that P_r doesn't make any influence to the fully-passive RIS schemes. However, the achievable secrecy rates of the fixed hybrid RIS, dynamic hybrid RIS and fully-active schemes all monotonically increase with P_r . As P_r increases, the secrecy rate of the fully-active RIS scheme approaches that of the dynamic hybrid RIS scheme, which can be attributed to the enhanced power amplification gain. The comparison results between our proposed fixed and dynamic hybrid RIS schemes and their corresponding SDR-based (CR-based) counterparts are similar to those in Fig. 6. Fig. 7(b) shows the optimal active/passive elements allocation for the

dynamic hybrid RIS scheme under different values of P_r . We readily observe that the optimal number of active elements naturally increases with the amplification power budget P_r , which coincides with the results in Fig. 7(a). This may because a larger P_r increases the DoFs of optimizing Ψ_a and Λ , thereby implying a higher power amplification gain, which is beneficial to enhance the system secrecy rate performance.

Fig. 8 illustrates the impact of the number of RIS reflecting elements M on the secrecy rates achieved by different schemes, where the active element ratio of the fixed hybrid RIS scheme is set as $\frac{1}{2\sqrt{M}}$. We readily observe that for all schemes, the secrecy rates increase with M since more RIS elements are able to provide higher spatial DoFs, thereby allowing for more precise adjustments of phases and magnitudes of the incident signals. In addition, both the fixed and dynamic hybrid RIS schemes yield a significantly higher secrecy rate than the fully-passive RIS scheme, since they are able to provide additional power amplification gains. In particular, for a large M , the fully-active RIS scheme performs the worst due to the limited amplification power budget.

Fig. 9 presents the secrecy rates achieved by different types of RISs versus the number of antennas at Eve N_e . Evidently, the secrecy rates of all considered schemes deteriorate with N_e . This is attributed to the Eve's enhanced signal interception capability induced by the increasing N_e . Owing to the addi-

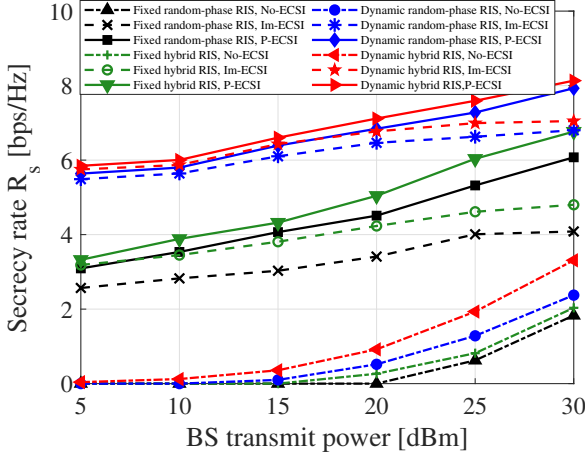


Fig. 10. Secrecy rates versus the BS transmit power under perfect ECSI, imperfect ECSI, and no ECSI.

tional power amplification gain, both the hybrid RIS and the fully-active RIS schemes perform better than the fully-passive RIS. Additionally, the secrecy rate of the dynamic hybrid RIS scheme decreases much less than those of the fixed hybrid RIS scheme and the fully-active scheme with N_e due to its sufficient flexibility in elements allocation.

Fig. 10 plots the secrecy rates achieved by the proposed fixed and dynamic hybrid RIS schemes and the random-phase RIS schemes under perfect ECSI, imperfect ECSI and no ECSI, which are respectively abbreviated as P-ECSI, Im-ECSI and No-ECSI. For the Im-ECSI case, the uncertainty region radii of both channels $\mathbf{H}_{BE}$ and $\mathbf{H}_{RE}$ is set as $\epsilon_i = \delta \|\hat{\mathbf{H}}_i\|_F, i \in \{BE, RE\}$, $\delta = 0.1$. For the No-ECSI case, we additionally introduce AN at the BS, where the power $P_{TP} = \frac{1}{4}P_t$ and $P_{AN} = \frac{3}{4}P_t$ are respectively allocated to the transmit precoder and the AN for simplicity [44]. Then, the optimal transmit precoder $\mathbf{q}^* = \sqrt{P_{TP}} \frac{\mathbf{h}}{\|\mathbf{h}\|}$ can be obtained based on the maximum-ratio transmission, where $\mathbf{h}^H = \mathbf{h}_{BU}^H + \mathbf{h}_{RU}^H \Psi \mathbf{H}_{BR}$ denotes the compound BS-RIS-user channel. With the obtained $\mathbf{q}^*$, the optimal Ψ_a^* and Ψ_p^* for maximizing legitimate user SINR are still available via the proposed SCA-based iterative algorithm, based on which the AN transmit covariance matrix $\mathbf{R}_{AN}$ is designed to be orthogonal to the BS-RIS-user channel $\mathbf{h}^H$, i.e., $\mathbf{R}_{AN} = \frac{P_{AN}}{N_t - 1} \mathbf{U}_{AN} \mathbf{U}_{AN}^H$, where the columns of the semi-unitary matrix $\mathbf{U}_{AN} \in \mathbb{C}^{N_t \times N_t - 1}$ lie in the null-space of $\mathbf{h}$. Finally, we can calculate the actual achievable secrecy rate based on the obtained $\{\mathbf{q}^*, \mathbf{R}_{AN}, \Psi_a^*, \Psi_p^*, \Lambda\}$.

It is clear from Fig. 10 that all considered schemes under P-ECSI naturally outperform their corresponding counterparts under Im-ECSI and No-ECSI. For the Im-ECSI case, our proposed fixed and dynamic hybrid RIS schemes both achieve higher secrecy rates than their random-phase counterparts, which demonstrates the robustness of our proposed schemes in a certain extent. Furthermore, when no ECSI is available, the proposed two schemes both exhibit the worst secrecy performance among all three considered cases, which may be attributed to the significantly reduced DoFs in the joint design of transmit precoder and hybrid active-passive reflection

matrices.

VII. CONCLUSION

In this paper, we investigated a novel hybrid active-passive RIS aided secure MISO system with the presence of a multi-antenna Eve, where both fixed and dynamic RIS architectures were considered. In order to maximize the secrecy rate, we jointly optimized the transmit covariance matrix and the hybrid RIS reflection matrix, subject to the transmit power budget at the BS and the amplification power budget at the active sub-RIS. To tackle this intractable problem, we firstly explored the inherent rank-1 property of the optimal transmit covariance matrix to reexpress the original SRM problem in a tractable form, and then proposed an efficient two-loop iterative algorithm to obtain a high-quality suboptimal solution. Also, we investigated the dynamic allocation of active and passive elements for further improving system security. Numerical results illustrated the superior performance of our proposed fixed and dynamic hybrid RIS schemes over the existing fully-passive and fully-active RIS schemes. Moreover, it was hinted that the dynamic hybrid RIS could strike a flexible balance between the square-order passive beamforming gain and the power amplification gain by adjusting the active/passive element allocation according to the varying propagation environment.

APPENDIX

A. Proof of Proposition 1

Firstly, by introducing an auxiliary variable τ and employing the Charnes–Cooper transformation, we can equivalently transform the relaxed SRM problem (9) into

$$\max_{\tilde{\mathbf{Q}} \succeq \mathbf{0}, \Psi_a, \Psi_p, \tau} \tau + \text{Tr}(\tilde{\mathbf{Q}} \mathbf{A}_\Psi), \quad (56a)$$

$$\text{s.t.} \quad (8d), \tau + \text{Tr}(\tilde{\mathbf{Q}} \mathbf{B}_\Psi) \leq 1, \quad (56b)$$

$$\text{Tr}(\tilde{\mathbf{Q}}) \leq P_t \tau, \quad (56c)$$

$$\text{Tr}(\Psi_a (\mathbf{H}_{BR} \tilde{\mathbf{Q}} \mathbf{H}_{BR}^H + \sigma_r^2 \tau \mathbf{I}_M) \Psi_a^H) \leq (P_r - M_a P_{dc}) \tau. \quad (56d)$$

It can be readily observed that the relaxed problem (56) is jointly concave on $\{\tilde{\mathbf{Q}}, \tau\}$ with any given Ψ_a and Ψ_p . Moreover, the equivalence between problems (9) and (56) can be established since tight (56b) always holds at the optimum of problem (56), which can be easily proved by contradiction.

Secondly, we consider the following transmit power minimization problem w.r.t. $\tilde{\mathbf{Q}}$ to explore the inherent characteristics of the optimal $\tilde{\mathbf{Q}}^*$ to problem (56) with any given Ψ_a , Ψ_p and τ .

$$\min_{\tilde{\mathbf{Q}} \succeq \mathbf{0}} \text{Tr}(\tilde{\mathbf{Q}}), \quad (57a)$$

$$\text{s.t.} \quad \tau + \text{Tr}(\tilde{\mathbf{Q}} \mathbf{B}_\Psi) \leq 1, \quad (57b)$$

$$\text{Tr}(\Psi_a \mathbf{H}_{BR} \tilde{\mathbf{Q}} \mathbf{H}_{BR}^H \Psi_a^H) \leq P_q \tau, \quad (57c)$$

$$\tau + \text{Tr}(\tilde{\mathbf{Q}} \mathbf{A}_\Psi) \geq \bar{f}, \quad (57d)$$

where $\bar{f}$ denotes the objective value of problem (56). Denote the optimal solutions to problems (56) and (57) as $\tilde{\mathbf{Q}}_s^*$ and $\tilde{\mathbf{Q}}_p^*$, respectively. We readily find that $\tilde{\mathbf{Q}}_s^*$ is also a feasible solution

to problem (57). Moreover, we infer from the constraint (56c) that $\text{Tr}(\tilde{\mathbf{Q}}_p^*) \leq \text{Tr}(\tilde{\mathbf{Q}}_s^*) \leq P_t \tau$. Similarly, since $\tilde{\mathbf{Q}}_p^*$ is also feasible in the maximization problem (56), we obtain $\tau + \text{Tr}(\tilde{\mathbf{Q}}_p^* \mathbf{A}_\Psi) \leq \tau + \text{Tr}(\tilde{\mathbf{Q}}_s^* \mathbf{A}_\Psi) = \bar{f}$. Hence, the optimal $\tilde{\mathbf{Q}}_p^*$ of problem (57) satisfies $\tau + \text{Tr}(\tilde{\mathbf{Q}}_p^* \mathbf{A}_\Psi) = \bar{f}$ due to the constraint (57d). Thus, $\tilde{\mathbf{Q}}_p^*$ is also the optimal solution to problem (56) with given Ψ_a , Ψ_p and τ . Then, we strive to explore the inherent characteristics of the optimal $\tilde{\mathbf{Q}}^*$ to problem (57) based on the KKT conditions listed as follows.

$$\mathbf{Z}^* = \mathbf{I}_{N_t} + \beta_1^* \mathbf{B}_\Psi + \beta_2^* \mathbf{H}_{\text{BR}}^H \Psi_a^H \Psi_a \mathbf{H}_{\text{BR}} - \beta_3^* \mathbf{A}_\Psi, \quad (58a)$$

$$\mathbf{Z}^* \tilde{\mathbf{Q}}^* = \mathbf{0}, \quad \mathbf{Z}^* \succeq \mathbf{0}, \quad \tilde{\mathbf{Q}}^* \succeq \mathbf{0}, \quad (58b)$$

where $\beta_1 \in \mathbb{R}_+$, $\beta_2 \in \mathbb{R}_+$, $\beta_3 \in \mathbb{R}_+$ and $\mathbf{Z} \in \mathbb{H}_+^{N_t}$ denote the dual variables associated with constraints (57b), (57c), (57d) and $\tilde{\mathbf{Q}} \succeq \mathbf{0}$, respectively. Postmultiplying (58a) by $\tilde{\mathbf{Q}}^*$ and referring to (58b) yield $(\mathbf{I}_{N_t} + \beta_1^* \mathbf{B}_\Psi + \beta_2^* \mathbf{H}_{\text{BR}}^H \Psi_a^H \Psi_a \mathbf{H}_{\text{BR}}) \tilde{\mathbf{Q}}^* = \beta_3^* \mathbf{A}_\Psi \tilde{\mathbf{Q}}^*$. Then, we obtain $\text{rank}((\mathbf{I}_{N_t} + \beta_1^* \mathbf{B}_\Psi + \beta_2^* \mathbf{H}_{\text{BR}}^H \Psi_a^H \Psi_a \mathbf{H}_{\text{BR}}) \tilde{\mathbf{Q}}^*) = \text{rank}(\beta_3^* \mathbf{A}_\Psi \tilde{\mathbf{Q}}^*)$. Based on the fact that $\mathbf{I}_{N_t} + \beta_1^* \mathbf{B}_\Psi + \beta_2^* \mathbf{H}_{\text{BR}}^H \Psi_a^H \Psi_a \mathbf{H}_{\text{BR}} \succ \mathbf{0}$ and $\text{rank}(\beta_3^* \mathbf{A}_\Psi \tilde{\mathbf{Q}}^*) \leq \text{rank}(\mathbf{A}_\Psi) = 1$ hold, it's verified that $\text{rank}(\tilde{\mathbf{Q}}^*) \leq 1$. Thus, we obtain that the optimal $\tilde{\mathbf{Q}}^*$ to problem (56) or problem (57) is rank one.

Finally, based on the equivalence between problem (9) and (56), the optimal solution to problem (9) is also rank one. Moreover, by recalling that problem (9) is an upper-bound approximation of problem (8), where the determinant equality holds only if $\text{rank}(\mathbf{A}) = 1$, we deduce that problem (9) is a tight relaxation of problem (8). This completes the proof.

B. Proof of Lemma 1

We prove Lemma 1 by contradiction. First, we suppose that the constraints (10b) and (10c) are both inactive at the optimum. Denote the optimal solution to problem (10) by $(\mathbf{q}_1^*, \Psi_{a,1}^*, \Psi_{p,1}^*)$ satisfying $(\mathbf{q}_1^*)^H \mathbf{q}_1^* < P_t$ and $\text{Tr}(\Psi_{a,1}^* \mathbf{H}_{\text{BR}} \mathbf{q}_1^* (\mathbf{q}_1^*)^H \mathbf{H}_{\text{BR}}^H (\Psi_{a,1}^*)^H) < P_q$, where P_q refers to $P_r - M_a P_{dc} - \sigma_r^2 \text{Tr}(\Psi_{a,1}^* (\Psi_{a,1}^*)^H)$ for brevity. Then, we can always find another set of feasible solutions $(\mathbf{q}_2^*, \Psi_{a,2}^*, \Psi_{p,2}^*)$ satisfying active (10b) or active (10c), where $\mathbf{q}_2^* = \sqrt{c} \mathbf{q}_1^*$, $\Psi_{a,2}^* = \Psi_{a,1}^*$, $\Psi_{p,2}^* = \Psi_{p,1}^*$ and $c = \min\{\frac{P_t}{(\mathbf{q}_1^*)^H \mathbf{q}_1^*}, \frac{P_q}{\text{Tr}(\Psi_{a,1}^* \mathbf{H}_{\text{BR}} \mathbf{q}_1^* (\mathbf{q}_1^*)^H \mathbf{H}_{\text{BR}}^H (\Psi_{a,1}^*)^H)}\} > 1$. Moreover, substituting $(\mathbf{q}_2^*, \Psi_{a,2}^*, \Psi_{p,2}^*)$ into the objective function of problem (10) yields $\frac{1 + (\mathbf{q}_2^*)^H \mathbf{A}_\Psi \mathbf{q}_2^*}{1 + (\mathbf{q}_2^*)^H \mathbf{B}_\Psi \mathbf{q}_2^*} = \frac{1 + c(\mathbf{q}_1^*)^H \mathbf{A}_\Psi \mathbf{q}_1^*}{1 + c(\mathbf{q}_1^*)^H \mathbf{B}_\Psi \mathbf{q}_1^*}$. In order to realize the correct decoding by the legitimate user and failed eavesdropping, we should guarantee that the achievable secrecy rate $R_s > 0$. That is to say, $1 + (\mathbf{q}_1^*)^H \mathbf{A}_\Psi \mathbf{q}_1^* > 1 + (\mathbf{q}_1^*)^H \mathbf{B}_\Psi \mathbf{q}_1^*$ always holds. Thus, the objective function monotonically increases with the increasing c . This property results in that the objective value generated by the feasible solution $(\mathbf{q}_2^*, \Psi_{a,2}^*, \Psi_{p,2}^*)$ is larger than that generated by $(\mathbf{q}_1^*, \Psi_{a,1}^*, \Psi_{p,1}^*)$, which conflicts with the assumption of the optimal solution. Hence, there is at least one active power constraint, i.e., (10b) or (10c). This completes the proof.

C. Proof of Proposition 2

We obtain the optimal solution to problem (11) in three cases. Considering the tight constraint (10b) and loose (10c),

the problem is reformulated as $\max_{\mathbf{q}} \frac{1 + \mathbf{q}^H \mathbf{A}_\Psi \mathbf{q}}{1 + \mathbf{q}^H \mathbf{B}_\Psi \mathbf{q}}$ with tight (10b). Then, substitute the normalized vector variable $\tilde{\mathbf{q}} = \frac{\mathbf{q}}{\sqrt{P_t}}$ into the objective function $\frac{1 + \mathbf{q}^H \mathbf{A}_\Psi \mathbf{q}}{1 + \mathbf{q}^H \mathbf{B}_\Psi \mathbf{q}}$, yielding a generalized Rayleigh quotient problem, i.e., $\max_{\tilde{\mathbf{q}}} \frac{\tilde{\mathbf{q}}^H (\mathbf{I}_{N_t} + P_t \mathbf{A}_\Psi) \tilde{\mathbf{q}}}{\tilde{\mathbf{q}}^H (\mathbf{I}_{N_t} + P_t \mathbf{B}_\Psi) \tilde{\mathbf{q}}}$. Thus, the optimal solution of problem (11) in the case of tight (10b) and loose (10c) is given by $\mathbf{q}^* = \sqrt{P_t} \mathbf{u}(\mathbf{I}_{N_t} + P_t \mathbf{A}_\Psi, \mathbf{I}_{N_t} + P_t \mathbf{B}_\Psi)$. Similarly, for the tight (10c) and loose (10b), problem (11) can also be converted into a generalized Rayleigh quotient problem, which can be recast as $\max_{\tilde{\mathbf{q}}} \frac{\tilde{\mathbf{q}}^H (\mathbf{C}_\Psi + P_q \mathbf{A}_\Psi) \tilde{\mathbf{q}}}{\tilde{\mathbf{q}}^H (\mathbf{C}_\Psi + P_q \mathbf{B}_\Psi) \tilde{\mathbf{q}}}$. It can be deduced that the matrix $\mathbf{C}_\Psi + P_q \mathbf{B}_\Psi$ in the denominator has full rank, otherwise the secrecy rate will become infinity and the transmit power constraint (10b) will be tight. Thus, the optimal solution in this case is given by $\mathbf{q}^* = \sqrt{P_c} \mathbf{u}(\mathbf{C}_\Psi + P_q \mathbf{A}_\Psi, \mathbf{C}_\Psi + P_q \mathbf{B}_\Psi)$, where the scalar factor P_c equals $P_q / (\mathbf{u}^H (\mathbf{C}_\Psi + P_q \mathbf{A}_\Psi, \mathbf{C}_\Psi + P_q \mathbf{B}_\Psi) \mathbf{C}_\Psi \mathbf{u}(\mathbf{C}_\Psi + P_q \mathbf{A}_\Psi, \mathbf{C}_\Psi + P_q \mathbf{B}_\Psi))$ to guarantee the tight constraint (10c). Finally, we consider the case of two tight constraints, i.e., (10b) and (10c). In order to simplify the double-constraint problem, we substitute $\mathbf{q} = \sqrt{\frac{P_t}{\tilde{\mathbf{q}}^H \tilde{\mathbf{q}}}} \tilde{\mathbf{q}}$ into problem (11), yielding the problem $\max_{\tilde{\mathbf{q}}} \frac{\tilde{\mathbf{q}}^H \tilde{\mathbf{q}} + P_t \tilde{\mathbf{q}}^H \mathbf{A}_\Psi \tilde{\mathbf{q}}}{\tilde{\mathbf{q}}^H \tilde{\mathbf{q}} + P_t \tilde{\mathbf{q}}^H \mathbf{B}_\Psi \tilde{\mathbf{q}}}$ with a single constraint $\tilde{\mathbf{q}}^H (P_t \mathbf{C}_\Psi - P_q \mathbf{I}_{N_t}) \tilde{\mathbf{q}} = 0$, which is still a CFQP. Thus, we apply the Charnes-Cooper transform and obtain that

$$\max_{\tilde{\mathbf{q}}} \tilde{\mathbf{q}}^H (\mathbf{I}_{N_t} + P_t \mathbf{A}_\Psi) \tilde{\mathbf{q}}, \quad (59a)$$

$$\text{s.t. } \tilde{\mathbf{q}}^H (\mathbf{I}_{N_t} + P_t \mathbf{B}_\Psi) \tilde{\mathbf{q}} = 1, \quad (59b)$$

$$\tilde{\mathbf{q}}^H (P_t \mathbf{C}_\Psi - P_q \mathbf{I}_{N_t}) \tilde{\mathbf{q}} = 0. \quad (59c)$$

It can be readily observed that problem (59) is actually a QCQP problem. However, due to the convex objective function (59a) and the nonconvex constraint (59c), problem (59) is nonconvex. To tackle this problem, we convert problem (59) into the SDP problem (13) by introducing $\tilde{\mathbf{Q}} = \tilde{\mathbf{q}} \tilde{\mathbf{q}}^H$. Thus, the rank-one solution $\tilde{\mathbf{Q}}^*$ can be solved by the famous optimization toolbox CVX. Then, the optimal $\mathbf{q}^*$ can be obtained through eigenvalue decomposition on $\tilde{\mathbf{Q}}^*$. Specifically, $\mathbf{q}^*$ is given by $\mathbf{q}^* = \sqrt{\frac{P_t \bar{\lambda}_{\mathbf{q}}}{\text{Tr}(\tilde{\mathbf{Q}}^*)}} \mathbf{v}_{\mathbf{q}}$ where $\mathbf{v}_{\mathbf{q}}$ is the normalized eigenvector associated with the maximum eigenvalue $\bar{\lambda}_{\mathbf{q}}$ of the matrix $\tilde{\mathbf{Q}}^*$. This completes the proof.

D. Proof of Lemma 2

It can be readily inferred by contradiction that the constraint (32b) is tight at the optimum. Otherwise, we can always decrease $[\mathbf{V}]_{m,m}$ for any m to achieve a larger objective value for problem (32). This completes the proof.

E. Proof of Proposition 3

Considering a fixed hybrid RIS assisted secure communication system, we propose a two-loop SCA-based iterative algorithm (**Algorithm 1**) to solve the nonconvex fractional problem (9). Define the objective value of problem (9) as $f_{\text{obj}}(\mathbf{Q}, \Psi_a, \Psi_p)$. We can obtain the nondecreasing sequences $f_{\text{obj}}(\mathbf{Q}, \Psi_a, \Psi_p)$'s, i.e., $f_{\text{obj}}(\mathbf{Q}^{(t_{\text{out}})}, \Psi_a^{(t_{\text{out}})}, \Psi_p^{(t_{\text{out}})}) \stackrel{(a)}{\leq} f_{\text{obj}}(\mathbf{Q}^{(t_{\text{out}}+1)}, \Psi_a^{(t_{\text{out}})}, \Psi_p^{(t_{\text{out}})}) \stackrel{(b)}{\leq}$

$f_{\text{obj}}(\mathbf{Q}^{(t_{\text{out}}+1)}, \Psi_a^{(t_{\text{out}}+1)}, \Psi_p^{(t_{\text{out}}+1)})$, where (a) holds since the optimal $\mathbf{Q}^*$ obtained from **Proposition 2** maximizes the objective value of problem (9) during each iteration. (b) holds since the joint optimization of Ψ_a and Ψ_p can be formulated as the convex problem (25) by leveraging the Dinkelbach's method and the SCA method. In addition, due to the closed feasible region, the objective value of problem (9) is also upper-bounded. Similarly, for the dynamic hybrid RIS case, the objective value of problem (35) is also monotonically nondecreasing and upper-bounded. This completes the proof.

REFERENCES

- [1] Y. Ge and J. Fan, "Robust secure beamforming for intelligent reflecting surface assisted full-duplex MISO systems," *IEEE Trans. Inf. Forensics Security*, vol. 17, pp. 253–264, 2022.
- [2] A. Almohamad, A. M. Tahir, A. Al-Kababji, H. M. Furqan, and H. Arslan, "Smart and secure wireless communications via reflecting intelligent surfaces: A short survey," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 1442–1456, Sep. 2020.
- [3] E. Illi *et al.*, "Physical layer security for authentication, confidentiality, and malicious node detection: A paradigm shift in securing IoT networks," *IEEE Commun. Surveys Tuts.*, vol. 26, no. 1, pp. 347–388, 1st Quart., 2024.
- [4] J. Zhu, R. Schober, and V. K. Bhargava, "Secure transmission in multicell massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 4766–4781, Sep. 2014.
- [5] D. W. K. Ng, E. S. Lo, and R. Schober, "Robust beamforming for secure communication in systems with wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4599–4615, Aug. 2014.
- [6] Y. Sun *et al.*, "Energy-efficient hybrid beamforming for multilayer RIS-assisted secure integrated terrestrial-aerial networks," *IEEE Trans. Commun.*, vol. 70, no. 6, pp. 4189–4210, Jun. 2022.
- [7] Y. Liu *et al.*, "Reconfigurable intelligent surfaces: Principles and opportunities," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 3, pp. 1546–1577, 3rd Quart., 2021.
- [8] L. Ruan and H. Zhu, "Leveraging RIS to assist communication and against CSI-based passive sensing," *IEEE Wireless Commun. Lett.*, vol. 13, no. 6, pp. 1750–1754, Jun. 2024.
- [9] B. Zheng, X. Xiong, J. Tang, and R. Zhang, "Intelligent reflecting surface-aided electromagnetic stealth against radar detection," *IEEE Trans. Signal Process.*, vol. 72, pp. 3438–3452, Jun. 2024.
- [10] M. Cui, G. Zhang, and R. Zhang, "Secure wireless communication via intelligent reflecting surface," *IEEE Wireless Commun. Lett.*, vol. 8, no. 5, pp. 1410–1414, Oct. 2019.
- [11] J. Qiao and M.-S. Alouini, "Secure transmission for intelligent reflecting surface-assisted mmWave and Terahertz systems," *IEEE Wireless Commun. Lett.*, vol. 9, no. 10, pp. 1743–1747, Oct. 2020.
- [12] Z. Li, Q. Lin, Y.-C. Wu, D. W. K. Ng, and A. Nallanathan, "Enhancing physical layer security with RIS under multi-antenna eavesdroppers and spatially correlated channel uncertainties," *IEEE Trans. Commun.*, vol. 72, no. 3, pp. 1532–1547, Mar. 2024.
- [13] Y. Liu, Z. Su, C. Zhang, and H.-H. Chen, "Minimization of secrecy outage probability in reconfigurable intelligent surface-assisted MIMOME system," *IEEE Trans. Wireless Commun.*, vol. 22, no. 2, pp. 1374–1387, Feb. 2023.
- [14] Z. Zhang *et al.*, "Active RIS vs. passive RIS: Which will prevail in 6G?" *IEEE Trans. Commun.*, vol. 71, no. 3, pp. 1707–1725, Mar. 2023.
- [15] Y. Zhang *et al.*, "Secure wireless communication in active RIS-assisted DFRC system," *IEEE Trans. Veh. Technol.*, early access, doi=10.1109/TVT.2024.3438151.
- [16] L. Dong, H.-M. Wang, and J. Bai, "Active reconfigurable intelligent surface aided secure transmission," *IEEE Trans. Veh. Technol.*, vol. 71, no. 2, pp. 2181–2186, Feb. 2022.
- [17] K. Zhi *et al.*, "Active RIS versus passive RIS: Which is superior with the same power budget?" *IEEE Commun. Lett.*, vol. 26, no. 5, pp. 1150–1154, May 2022.
- [18] Y. Ju *et al.*, "Beamforming optimization for hybrid active-passive RIS assisted wireless communications: A rate-maximization perspective," *IEEE Trans. Commun.*, vol. 72, no. 9, pp. 5428–5442, Sept. 2024.
- [19] C. Liao, F. Wang, G. Han, Y. Huang, and V. K. N. Lau, "Beamforming design for hybrid active-passive RIS assisted integrated sensing and communications," *IEEE Commun. Lett.*, vol. 27, no. 11, pp. 2938–2942, Nov. 2023.
- [20] N. P. Le and M.-S. Alouini, "Distributed hybrid active-passive RIS-assisted THz wireless systems: Performance analysis and optimization," *IEEE Trans. Commun.*, early access, doi=10.1109/TCOMM.2024.3450791.
- [21] A. Huang, X. Mu, L. Guo, and G. Zhu, "Hybrid active-passive RIS transmitter enabled energy-efficient multi-user communications," *IEEE Trans. Wireless Commun.*, vol. 23, no. 9, pp. 10653–10666, Sept. 2024.
- [22] Z. Kang, C. You, and R. Zhang, "Active-passive IRS aided wireless communication: New hybrid architecture and elements allocation optimization," *IEEE Trans. Wireless Commun.*, vol. 23, no. 4, pp. 3450–3464, Apr. 2024.
- [23] Q. Peng, Q. Wu, G. Chen, R. Liu, S. Ma, and W. Chen, "Hybrid active-passive IRS assisted energy-efficient wireless communication," *IEEE Commun. Lett.*, vol. 27, no. 8, pp. 2202–2206, Aug. 2023.
- [24] R. S. P. Sankar and S. P. Chepuri, "Channel-aware placement of active and passive elements in hybrid RIS-assisted MISO systems," *IEEE Wireless Commun. Lett.*, vol. 12, no. 7, pp. 1229–1233, Jul. 2023.
- [25] H. Xie, D. Li, and B. Gu, "Exploring hybrid active-passive RIS-aided MEC systems: From the mode-switching perspective," *IEEE Trans. Wireless Commun.*, vol. 23, no. 9, pp. 11291–11308, Sept. 2024.
- [26] Y. Ma, K. Liu, Y. Liu, and L. Zhu, "Secure transmission via hybrid active-passive RIS against eavesdropping and jamming," *IEEE Wireless Commun. Lett.*, vol. 13, no. 7, pp. 1978–1982, Jul. 2024.
- [27] E. N. Egashira, D. P. M. Osorio, N. T. Nguyen, and M. Juntti, "Secrecy capacity maximization for a hybrid relay-RIS scheme in mmWave MIMO networks," in *Proc. IEEE 95th Veh. Technol. Conf.*, 2022, pp. 1–6.
- [28] Z. Chen, Y. Guo, P. Zhang, H. Jiang, Y. Xiao, and L. Huang, "Physical layer security improvement for hybrid RIS-assisted MIMO communications," *IEEE Commun. Lett.*, early access, doi=10.1109/LCOMM.2024.3427010.
- [29] H. Chen, N. Li, R. Long, and Y.-C. Liang, "Channel estimation and training design for active RIS aided wireless communications," *IEEE Wireless Commun. Lett.*, vol. 12, no. 11, pp. 1876–1880, Nov. 2023.
- [30] D. V. Q. Rodrigues and T. Singh, "Efficient power allocation strategies in hybrid active-passive reconfigurable intelligent surfaces," *IEEE Commun. Lett.*, vol. 28, no. 1, pp. 113–117, Jan. 2024.
- [31] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas I: The MISOME wiretap channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3088–3104, Jul. 2010.
- [32] M. Bashar, K. Cumanan, A. G. Burr, M. Debbah, and H. Q. Ngo, "On the uplink max-min SINR of cell-free massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 18, no. 4, pp. 2021–2036, Apr. 2019.
- [33] A. Ranjha and G. Kaddoum, "Quasi-optimization of distance and blocklength in URLLC aided multi-hop UAV relay links," *IEEE Wireless Commun. Lett.*, vol. 9, no. 3, pp. 306–310, Mar. 2020.
- [34] M. Grant, "CVX : Matlab software for disciplined convex programming, version 1.21," <http://cvxr.com/cvx>, 2011.
- [35] L. Dong and H.-M. Wang, "Secure MIMO transmission via intelligent reflecting surface," *IEEE Wireless Commun. Lett.*, vol. 9, no. 6, pp. 787–790, Jun. 2020.
- [36] K. Shen and W. Yu, "Fractional programming for communication systems—Part I: Power control and beamforming," *IEEE Trans. Signal Process.*, vol. 66, no. 10, pp. 2616–2630, May 2018.
- [37] C. Feng, W. Shen, J. An, and L. Hanzo, "Weighted sum rate maximization of the mmWave cell-free MIMO downlink relying on hybrid precoding," *IEEE Trans. Wireless Commun.*, vol. 21, no. 4, pp. 2547–2560, Apr. 2022.
- [38] O. Tervo *et al.*, "Energy-efficient multicell multigroup multicasting with joint beamforming and antenna selection," *IEEE Trans. Signal Process.*, vol. 66, no. 18, pp. 4904–4919, Sep. 2018.
- [39] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex Optimization*. Cambridge university press, 2004.
- [40] Z. Chu, K. Cumanan, Z. Ding, M. Johnston, and S. Y. Le Goff, "Secrecy rate optimizations for a MIMO secrecy channel with a cooperative jammer," *IEEE Trans. Veh. Technol.*, vol. 64, no. 5, pp. 1833–1847, May 2015.
- [41] E. A. Gharavol and E. G. Larsson, "The sign-definiteness lemma and its applications to robust transceiver optimization for multiuser mimo systems," *IEEE Trans. Signal Process.*, vol. 61, no. 2, pp. 238–252, Jan. 2013.

- [42] P. Zhao, M. Zhang, H. Yu, H. Luo, and W. Chen, "Robust beamforming design for sum secrecy rate optimization in MU-MISO networks," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 9, pp. 1812–1823, Sep. 2015.
- [43] R. Feng, Q. Li, Q. Zhang, and J. Qin, "Robust secure beamforming in MISO full-duplex two-way secure communications," *IEEE Trans. Veh. Technol.*, vol. 65, no. 1, pp. 408–414, Jan. 2016.
- [44] H.-M. Wang, J. Bai, and L. Dong, "Intelligent reflecting surfaces assisted secure transmission without eavesdropper's CSI," *IEEE Signal Process. Lett.*, vol. 27, pp. 1300–1304, 2020.



Chengwen Xing (Member, IEEE) received the B.Eng. degree from Xidian University, Xi'an, China, in 2005, and the Ph.D. degree from the University of Hong Kong, Hong Kong, China, in 2010. Since September 2010, he has been with the School of Information and Electronics, Beijing Institute of Technology, Beijing, China, where he is currently a Full Professor. From September 2012 to December 2012, he was a Visiting Scholar at the University of Macau, Macau SAR, China. His current research interests include machine learning, statistical signal processing, convex optimization, multivariate statistics, and array signal processing.



Shiqi Gong received the B.S. and Ph.D. degrees in electronic engineering from Beijing Institute of Technology, China, in 2014 and 2020, respectively. She is currently an associate professor with the School of Cyberspace Science and Technology, Beijing Institute of Technology, Beijing, China. Her research interests are in the area of intelligent reflecting surface, physical-layer security, resource allocation, and convex optimization. She was a recipient of the Best Ph.D. Thesis Award of Beijing Institute of Technology in 2020.



Yue Ju received the B.E. degree from Beijing Institute of Technology, China, in 2021. She is currently pursuing the Ph.D. degree in the School of Information and Electronics at Beijing Institute of Technology, China. Her main research interests include signal processing, reconfigurable intelligent surface and convex optimization.



Heng Liu received the B.E. degree and the Ph.D. degree from the Beijing Institute of Technology, China, in 2017 and 2022, respectively. From 2022 to 2024, he was a postdoctoral fellow with the School of Cyberspace Science and Technology, Beijing Institute of Technology. He is currently an assistant professor with the School of Information and Electronics, Beijing Institute of Technology. His main research interests include reconfigurable intelligent surface, integrated sensing and communication, satellite communication and convex optimization.



Liang Liu (Senior Member, IEEE) is currently an Associate Professor with the Department of Electrical and Electronic Engineering, The Hong Kong Polytechnic University. He was a recipient of the 2021 IEEE Signal Processing Society Best Paper Award, the 2017 IEEE Signal Processing Society Young Author Best Paper Award, the Best Student Award of 2022 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), and the Best Paper Award of the 2011 International Conference on Wireless Communications and Signal

Processing. He was recognized by Clarivate Analytics as a Highly Cited Researcher in 2018. He is an Editor of IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. He is a co-author of the book "Next Generation Multiple Access" published at Wiley-IEEE Press.