Joint active and passive beamforming optimization for intelligent reflecting surface assisted proactive eavesdropping

Jie Yang | Kaizhi Huang | Xiaoli Sun | Yi Wang

1 Information Engineering University, Zhengzhou, China
2 Zhengzhou University of Aeronautics, Zhengzhou, China

Correspondence
Kaizhi Huang, Information Engineering University, Zhengzhou, 450002, China.
Email: huangkaizhi@tsinghua.org.cn

Funding information
National Natural Science Foundation of China, Grant/Award Numbers: 61801435, 61871404; The Scientific and Technological Key Project of Henan Province, Grant/Award Number: 212102210559; National Natural Science Foundation Innovative Groups Project of China, Grant/Award Number: 61521003

Abstract
This paper investigates a three-node proactive eavesdropping system with the aid of intelligent reflecting surface (IRS), where a monitor tries to wiretap information and transmit jamming signals simultaneously to interfere with the suspicious link. The IRS is deployed around the suspicious users to reconstruct propagation channel. Meanwhile, with large low-cost passive reflecting elements, it can provide more spatial degrees of freedom to moderate the link quality. In order to degrade exposure risk of eavesdropping behavior, the authors’ objective is to minimize the jamming power by carefully designing jamming beamforming vectors and the phase shifts under the constraint of reliable intercepted information. The formulated problem is non-trivial to solve due to the coupling variables and unit-modulus constraints. Fortunately, by using alternating optimization and successive convex approximation techniques, the original problem is transformed into convex form and a sub-optimal solution is achieved. Numerical results show that the proposed scheme can effectively reduce the jamming power and the number of the jamming antennas over benchmark schemes, which is quite suitable for covertly overhearing signal.

1 | INTRODUCTION

Recently, the ubiquitous accessibility of wireless communication system has greatly changed our daily life. However, due to the broadcast nature of wireless communication, safeguarding wireless communications system has become an important issue. Physical layer security technology has been ascertained as a promising approach, which could provide lightweight protection compared with traditional encryption technology. In [1], secure communication is studied in multi-user massive multiple-input multiple-output system. Mamaghani et al. [2, 3] investigate a two-way secure communication scheme where legitimate communication parties exchange confidential information via a wireless-powered untrusted relay. In [4], two cooperative unmanned aerial vehicles (UAVs) are exploited to assist secure communication and energy harvesting. In these studies, many designs have been carried out for the purpose that makes the transmitted messages indecipherable to eavesdroppers. While, in some areas, wireless communication may be used for unkind purposes, such as commercial crime. Specially, malicious users can send out secret data via infrastructure-free communication means, such as device-to-device communications, which is hard to access for authorized parties. Therefore, as a new paradigm shift in secure communication, wireless information surveillance is proposed and has garnered increased attention for its capability in dealing with malicious users[5].

Since the monitor is usually located away from the suspicious users, the eavesdropping link is inferior to the communication link and it is hard to decode the suspicious signal clearly for passive eavesdropping. Thus, jamming-assisted and relay-assisted proactive eavesdropping are widely recognized solutions [6–9]. Xu et al. [6] considers a three-node surveillance model, where the monitor moderates the suspicious communication rate by transmitting jamming signal, and derives the optimal jamming strategy for the maximum eavesdropping non-outage probability. In [7], to facilitate covertly distant eavesdropping, multiple intermediate nodes switching between jamming and monitoring mode are introduced and deployed around a pair of suspicious users. And the optimal transmit power allocation and mode selection strategy is developed to achieve the maximum eavesdropping rate. In [8], power splitting technique is utilized
by the legitimate monitor and the received signal is split into two parts for information eavesdropping and spoofing relaying, respectively. Then, the relay is designed to act as different roles, including jamming, constructive relaying and so on, to improve eavesdropping performance. Furthermore, UAVs acted as amplify-and-forward (AF) relays with the advantage of less interference air-to-ground relay channels are utilized to improve information surveillance[9]. Despite such striking performance improvements brought by these methods, there exists a risk of exposing eavesdropping behavior. Relatively high jamming power and active relay nodes are easy to be detected by the suspicious users. Therefore, it is worth paying attention to considering how to realize reliable eavesdropping on the basis of concealment.

Meanwhile, intelligent reflecting surface (IRS), which can reconstruct wireless propagation environment in real time, has been viewed as a key emerging technology for 6G network[10]. IRS is composed of lots of low-cost reflecting units, and each unit can adjust the reflection coefficient on the incident signal independently whereby the reflected signal can be strengthened or weakened at specified users. Moreover, with the characteristic of low profile and conformal geometry, it is easy to mount on surrounding objects [11]. Thus, as a novel promising solution, IRS has been investigated in many applications such as coverage extension[12], physical layer security[13, 14], energy efficiency improvement[15]. In [16], the author studies the secrecy performance gains of non-orthogonal multiple access system via IRS and derives the closed-form average secrecy capacity expressions. In [17], a novel adverse application of IRS is proposed, where the IRS controlled by malicious users is utilized to sabotage the legitimate communication system without any energy footprint.

In fact, IRS is naturally applicable for covertly eavesdropping due to its inherent features of passive enhanced communication and easy-to-deploy. Compared with the existing schemes, more design parameters brought by IRS is more challenge and usually intractable. By far, there are few work considering IRS in proactive eavesdropping system. Motivated by this, we study an IRS-aided three-node surveillance model with the goal to minimize the jamming power. Our main contributions are summarized as follows:

(1) An IRS-assisted three-node proactive eavesdropping scheme is proposed. In this model, IRS is deployed covertly around the suspicious users and is remotely controlled by the monitor. The monitor works in full-duplex mode and transmits jamming signals to moderate the suspicious communication rate. Meanwhile, more reflected channel paths are brought by the IRS and the received signals are re-analyzed. The phase shifts of the IRS are properly tuned to adjust both the eavesdropping and communication channel quality.

(2) A resource allocation problem aimed at jamming power minimization is established by jointly designing jamming beamforming and the phase shifts. The original problem is non-convex with a complicated form. To tackle it, a sub-optimal iterative algorithm is proposed. First, the problem is decomposed into two sub-problems by applying an alternating optimization algorithm. Then each sub-problem is transformed into convex with successive convex approximation (SCA) and semi-definite relaxation (SDR) techniques. Moreover, the convergence of the algorithm is analyzed. Finally, we prove that when the number of reflection units approaches infinity, the monitor can overhear the signal reliably without transmitting jamming signal.

(3) Simulation results show the impact of parameters, such as the location of the monitor, jamming antenna numbers, on system performance. Compared with the scheme without IRS and the scheme with AF relay, our proposed scheme can achieve better performance and dramatically reduce jamming power and jamming antenna numbers. Moreover, when the number of low cost reflecting units increases, the required jamming power can be further cut down, which leads to higher power efficiency and is beneficial for proactive eavesdropping.

The remainder of this paper is organized as follows. Section 2 introduces the system model and formulates the optimal problem. Section 3 focuses on the algorithms for the active and passive beamforming design. Sections 4 and 5 give simulation results and conclusions, respectively.

2 SYSTEM MODEL AND PROBLEM FORMULATION

2.1 System model

As depicted in Figure 1, we consider a three-node point-to-point surveillance model. The suspicious transmitter Alice exchanges the secret data with the suspicious receiver Bob via device-to-device communications without going through cellular infrastructures. The monitor Mon tries to overhear the suspicious signal. In order to avoid being detected, Mon is deployed far away from Alice, which leads to that the eavesdropping channel condition is worse than the communication channel.
condition. Moreover, Mon exploits full-duplex technology to simultaneously eavesdrop and jam [18]. Meanwhile, an intelligent reflecting surface, which is composed of \( N \) reflecting units, is covertly deployed on surrounding buildings near Bob and is controlled by Mon through a private wireless channel. Alice is equipped with \( M_a \) antennas, and Bob is deployed with a single antenna. While, Mon is equipped with two kinds of antennas, that is single antenna for receiving and \( M_e \) antennas for jamming.

We have considered a block Rician fading environment [6, 7], where the channel condition remains unchanged over each transmission block but changes from one block to another. Each communication link is described as follows: the suspicious communication link from Alice to Bob is composed of a direct link and a reflected link, where the direct link is denoted by \( h_{ab}^{II} \) and the reflected link from Alice to IRS and IRS to Bob are represented as \( H_{ar} \) and \( h_{rb}^{II} \), respectively. The eavesdropping link from Alice to Mon is also consisted of a direct link and a reflected link, where the direct link is denoted by \( h_{ab}^{I} \) and the reflected link from Alice to IRS and IRS to Mon are denoted by \( H_{er} \) and \( h_{rb}^{I} \). Mon performs co-channel jamming to Bob while eavesdropping. The jamming link from Mon to Bob is also composed of a direct link and a reflected link, where the direct link is denoted by \( h_{fb}^{II} \), the reflected link from Mon to IRS and from IRS to Bob are represented as \( H_{er} \) and \( h_{rb}^{II} \). Then the phase shift matrix at its \( nth \) reflecting element. Compared with the model without IRS, the signal received at each user is the combination of reflected signal and direct signal, which needs to be re-analyzed.

The transmitted suspicious signal from Alice is expressed as

\[
x = \sqrt{P_t} w_i,
\]

where \( s \sim \mathcal{CN}(0,1) \), \( P_t \) denotes the transmit power, \( w \in \mathbb{C}^{M_a \times 1} \) denotes the transmit precoding vector.

The received signal at Bob is the combination of the transmitted signal from Alice and the jamming signal from Mon, which can be expressed as

\[
y_b = \sqrt{P_t} (h_{ab}^{II} + h_{ab}^{I} \Phi H_{ar}) w_i + (h_{rb}^{II} + h_{rb}^{I} \Phi H_{er}) f_a + n_1,
\]

where \( n_1 \sim \mathcal{CN}(0,1) \) denotes the additive Gaussian white noise (AWGN) at Bob, \( a \sim \mathcal{CN}(0,1) \) denotes the jamming signal from Mon, and \( f \) denotes the jamming beamforming vector from Mon.

Then the signal-to-interference-plus noise ratio (SINR) at Bob can be given by

\[
\gamma_b = \frac{P_t |h_{ab}^{I} + h_{ab}^{II} \Phi H_{ar}|^2}{||h_{ab}^{II} + h_{ab}^{I} \Phi H_{er}||^2 + \sigma_0^2}.
\]

Let \( v^{II} = [v_1, v_2, \ldots, v_N] \) and \( v_I = e^{i\theta_i} \), then we can have

\[
h_{ab}^{I} \Phi H_{ar} = v^{II} H_{ar}^{b}, \quad h_{rb}^{II} \Phi H_{er} = v^{I} H_{er}^{b},
\]

where \( H_{ar} = \text{diag}(v_1, v_2, \ldots, v_N) H_{ar} \) and \( H_{er} = \text{diag}(v_1, v_2, \ldots, v_N) H_{er} \).

By denoting \( H_{ar} = [H_{ar}^{a}, H_{ar}^{b}], H_{er} = [H_{er}^{a}, H_{er}^{b}], v^{I} = [v^{II}, 1] \), the SINR at Bob can be simplified as

\[
\gamma_b = \frac{P_t |v^{II} H_{ar}^{b}|^2}{|v^{I} H_{er}^{b}|^2 + \sigma_0^2}.
\]

From (5), we can observe that \( \bar{v}^{II} H_{ar}^{b} \) and \( \bar{v}^{I} H_{er}^{b} \) can be seen as the equivalent communication link from Alice to Bob and the equivalent jamming link from Mon to Bob, respectively.

Thus, the suspicious communication rate at Bob is written as

\[
R_b = \log_2 (1 + \gamma_b) = \log_2 \left( 1 + \frac{P_t |v^{II} H_{ar}^{b}|^2}{|v^{I} H_{er}^{b}|^2 + \sigma_0^2} \right).
\]

Similarly, the received signal at Mon is given by

\[
y_e = \sqrt{P_t} (h_{ab}^{I} + h_{ab}^{II} \Phi H_{ar}) w_i + (h_{rb}^{I} + h_{rb}^{II} \Phi H_{er}) f_a + n_2,
\]

where \( n_2 \sim \mathcal{CN}(0,1) \) denotes the AWGN at Mon.

As shown in [19][20], the residual self-interference link is described as \( \sqrt{\rho} v^{II} \), where \( v^{II} \in \mathbb{C}^{M_e \times 1} \) is the fadong loop channel with complex Gaussian random variables (RV) with zero-mean and variance \( \lambda \). Thus, the SINR at Mon is derived as

\[
\gamma_e = \frac{P_t |h_{ab}^{I} + \eta^{I} \Phi H_{ar} w_i|^2}{|\sqrt{\rho} v^{II} + h_{rb}^{I} \Phi H_{er} f_a|^2 + \sigma_0^2}.
\]

Further, let us define the equivalent eavesdropping link as \( \bar{v}^{I} H_{ar}^{a} = \sqrt{\rho} v^{II} [H_{ar}^{a}] \), the equivalent jamming link as \( \bar{v}^{I} H_{er} = \sqrt{\rho} v^{II} [H_{er}] \), where \( H_{ar} = \text{diag}(v_1, v_2, \ldots, v_N) H_{ar} \) and \( H_{er} = \text{diag}(v_1, v_2, \ldots, v_N) H_{er} \), and then (8) can be transformed into

\[
\gamma_e = \frac{P_t |\bar{v}^{I} H_{ar}^{a} w_i|^2}{|\bar{v}^{I} H_{er} f_a|^2 + \sigma_0^2}.
\]

2.2 Problem formulation

According to [21], when the monitor has a better link than the suspicious receiver, the suspicious signal can be decoded with an
arbitrarily small error probability, and the effective eavesdropping rate is defined as $R_e = \log_2(1 + \gamma_e)$. Otherwise, the suspicious signal cannot be decoded without error, and the effective eavesdropping rate is defined as $R_e = 0$. In fact, since the monitor is usually located far away from the suspicious users, the eavesdropping link is worse than the communication link with high probability, which results in poor eavesdropping performance. Although jamming-assisted scheme can compensate this gap, excessive jamming power has a negative effect on the concealment of the monitor. Meanwhile, from (5)–(9), one can find that IRS could provide more spatial degrees of freedom with little energy.

Naturally, we consider a strategy that Mon works on passive eavesdropping mode and only deploys a large number of reflection units to ensure reliable information reception. And the problem is similar to the adverse application of [13], where the closed-form solution is hard to derive. Fortunately, when the number of reflection units approaches infinity, we have the following proposition.

**Proposition 1.** When Mon works on passive eavesdropping mode and the reflection coefficients at the IRS are adjusted to enhance the received signal at Mon, the average SINR at Mon approaches infinity as the number of reflection units approaches infinity.

**Proof.** see Appendix 8.1. □

Then, based on this simple phase shifts design strategy, since the channel $H_{ab}^{\parallel}$ and $H_{ab}^{\perp}$ are independent, the phase shifts is random with respect to Bob. And it is easy to know that as the number of reflection units approaches infinity, the SINR at Mon must be higher than that of Bob, which means Mon can overhear signal reliably without transmitting jamming signal.

Thus, when the difference between the communication link and the eavesdropping link is small, Mon can deploy passive IRS to reliably overhearing, which is verified in the subsequent simulation. Otherwise, jamming power is still needed since too large units are difficult to be fabricated in practice.

Based on above analysis, our goal is to minimize the jamming power by jointly designing the reflection coefficients and jamming beamforming vector while ensuring monitoring effectively. Hence, the optimization problem can be established by

$$
\begin{align*}
\min_{\hat{\nu}^{\parallel,\perp}} & \quad \text{Tr}(f^{\parallel} f) \\
\text{s.t.} & \quad \nu^{\parallel} \geq \nu^{\perp} \\
& \quad |\nu_n| = 1, n = 1, 2 \cdots N, \\
& \quad \nu_{N+1} = 1.
\end{align*}
$$

From problem (10), channel state information (CSI) is necessary for Mon to obtain the optimal solution. In this model, the CSI of each link can be perceived as follows. First, the CSI of the equivalent links can be acquired similar to that of conventional system without IRS. The equivalent jamming link $\hat{\nu}^{\parallel} H_{b}^{\parallel}$ and eavesdropping link $\hat{\nu}^{\parallel} H_{b}^{\parallel}$ can be estimated by overhearing the public pilot signal and control signal[22]. The equivalent loop link $\hat{\nu}^{\parallel} H_{a}^{\parallel}$ can be estimated beforehand by Mon. The equivalent suspicious communication link $\hat{\nu}^{\parallel} H_{b}^{\parallel}$ is relatively difficult to obtain. Fortunately, it could be perceived by listening the channel feedback sent from Bob to Alice[8, 22]. Secondly, to obtain the CSI of the direct link, such as $H_{ab}^{\parallel}, H_{ab}^{\perp}$, the reflected link should be estimated first. The estimation approach of the reflected link at the IRS is also the hotspot of current research[23]. Semi-passive IRS and Fully-passive IRS have been recognized two estimation methods. We consider the former approach where there exists some low-power sensing devices integrated into the IRS and the reflected link can be probed by these sensors. How to recover all the CSI of the reflected link can refer to [24, 25], which is beyond the scope of our discussion. As the IRS is controlled by Mon, the phase shifts and the CSI of the reflected link are known by Mon via privacy channel. Therefore, the direct link can be easily derived from formula (4). To study the fundamental performance limit of our proposed scheme, similar to [18], we assume perfect CSI of all links are available at Mon.

Furthermore, by substituting (5) and (9) into problem (10), the original problem can be retransformed as:

$$
\begin{align*}
\min_{\hat{\nu}^{\parallel,\perp}} & \quad \text{Tr}(f^{\parallel} f) \\
\text{s.t.} & \quad P_s |\hat{\nu}^{\parallel} H_{b}^{\parallel} w|^2 \geq P_s |\hat{\nu}^{\parallel} H_{b}^{\parallel} \hat{f} - \sigma_0^2 |\hat{\nu}^{\parallel} H_{b}^{\parallel} \hat{f} + \sigma_0^2 \\
& \quad |\nu_n| = 1, n = 1, 2 \cdots N, \\
& \quad \nu_{N+1} = 1.
\end{align*}
$$

Kindly note that for problem (11), $\hat{\nu}^{\parallel}$ and $f$ are designed by Mon, while transmit beamforming vector $w$ and transmit power $P_s$ are set by Alice, rather than Mon. Fortunately, from (11b), variable $P_s$ has no effect on the problem. Different from conventional physical secure communication model, of which Alice jointly designs transmit beamforming vector $w$ and phase shifts $\hat{\nu}^{\parallel}$ to maximize the secrecy rate at Bob [26], suspicious users have no knowledge about the existence of Mon and IRS. However, it can be predicted that the design criterion of vector $w$ by Alice is still to maximize communication rate and then $w$ is set based on the Maximal Ratio Transmission criterion (MRT). The CSI of the suspicious link estimated by Alice is denoted as $H_{Alice-Bob}^{\parallel}$ from (5), we have $H_{Alice-Bob}^{\parallel} = H_{0}^{\parallel} H_{b}^{\parallel}$, where $\hat{\nu}^{\parallel}_0$ denotes the phase shifts at the IRS during channel estimation stage. Then we have:

$$
\hat{w}^{\parallel} = \frac{\hat{\nu}^{\parallel}_0 H_{b}^{\parallel}}{||\hat{\nu}^{\parallel}_0 H_{b}^{\parallel}||}.
$$

It should be emphasized that since Mon has no knowledge about the instantaneous CSI of any links during channel estimation stage, the phase shifts cannot be designed in advance. Thus, at this time, a simple scheme with low energy consumption can be considered that all the reflecting units at the IRS are
simply turned off, namely $\Bar{\phi} = 0$ and $h_{\text{Alice-Bob}} = h_{\text{Bob}}^H$. Actually, from the follow-up simulation, the performance of random phase shifts is almost equal to the case without IRS. Moreover, in the slow-varying channel, statistical CSI can be long-term observed [27, 28] and the phase shifts $\Bar{\phi}$ during channel estimation stage can be preset based on statistical CSI, which is left for future work.

Thus, substituting (12) into the constraint (11b), we have

$$\frac{\gamma_0 |\Bar{\phi}^H H_{b_2} f|^2 + 1}{\gamma_0 |\Bar{\phi}^H H_{b_2} e|^2 + 1} \geq \frac{|\Bar{\phi}^H H_{b_1} h_{a b}|^2}{|\Bar{\phi}^H H_{a b}|^2},$$

(13)

where $\gamma_0 = 1/\sigma_0^2$.

It is worth noting that problem (11) is non-convex and intractable due to the non-concave fractional constraints, unit-modulus constraints and the coupling variables. In the following section, an efficient algorithm is developed to find a sub-optimal solution.

3 | JOINT DESIGN OF BEAMFORMING AND PHASE SHIFTS

In this section, alternating optimization is firstly utilized to separate the original problem into two sub-problems. Then, with SDR and SCA techniques, non-convex sub-problem is transformed to convex form approximately, which can be effectively solved.

3.1 | Optimizing $f$ for given $\Bar{\phi}^H$

When $\Bar{\phi}^H$ is fixed, the right-hand side of (13) is independent of $f$, which can be regarded as a constant. Let $\Hat{H}_{b_2}^H = \Hat{H}_{b_1}^H$, $\Hat{H}_{b_2} = \Bar{\phi}^H H_{b_2}$, $F = \Bar{\phi}^H$, and we have Rank($\Hat{F}$) = 1. Due to the non-convex constraint of Rank-1, SDR technique is utilized to drop it. Then with the properties of matrix traces, the problem can be equivalently rewritten as:

$$\min_{\Hat{F}} \text{Tr}(\Hat{F})$$

s.t. $\gamma_0 \text{Tr}(\Hat{H}_{b_2}^H \Hat{F}) + 1 \geq \lambda (\gamma_0 \text{Tr}(\Hat{H}_{b_2}^H F) + 1)$

(14a)

$$F \geq 0,$$

(14c)

$$\lambda = \frac{|\Bar{\phi}^H H_{b_1} h_{a b}|^2}{|\Bar{\phi}^H H_{a b}|^2}$$

where $\lambda$ is a constant.

Then, we can observe that problem (14) is a convex semi-definite programming (SDP) problem with linear constraints, which can be directly solved via the CVX solver. Since the constraint of Rank-1 is dropped in problem (14), it is necessary to verify the Rank of the obtained solution. Fortunately, we can conclude Proposition 1. Then, the optimal $f$ can be directly recovered based on the eigenvalue decomposition.

**Proposition 2.** The solution of $F$ always satisfies $\text{Rank}(F) \leq 1$

**Proof:** see Appendix 8.2.

3.2 | Optimizing $\Bar{\phi}^H$ for given $f$

When $f$ is fixed, the original problem is transformed into a check feasible problem. The convergence speed is slow when solving the problem directly. Thus, a slack variable $\bar{\phi} \geq \bar{\phi}$ is integrated into the right-hand side of the constraint (13), and the problem of finding feasible solution $\Bar{\phi}^H$ is equivalent to finding the maximum $\bar{\phi}$, which can accelerate the convergence[29].

Note that since the constraints (11c) and (11d) are not consistent, which cannot be dealt with in a unified form, we introduce slack variable $\bar{\phi} \in [0, 2\pi]$ and denote $\Bar{\phi}^H = e^{j\bar{\phi}}\Bar{\phi}^H$. Then the constraint (11b-11d) can be rewritten as:

$$\frac{\gamma_0 |\Bar{\phi}^H H_{b_2} f|^2 + 1}{\gamma_0 |\Bar{\phi}^H H_{b_2} e|^2 + 1} \geq 2^\alpha \frac{|\Bar{\phi}^H H_{b_1} h_{a b}|^2}{|\Bar{\phi}^H H_{a b}|^2}$$

(15a)

$|\bar{\phi}| = 1, n = 1, 2 \cdots N + 1.$

(15b)

From (15), we can see that the constraint (11d) can be relaxed without effect on the optimal solution. And the original problem can be transformed into finding the optimal $\Bar{\phi}^H$.

Next, let $\Hat{H}_{b_2} = H_{b_2} f, \Hat{H}_{b_1} = H_{b_1} f, \Hat{H}_{b_2} = \Hat{H}_{b_1} = H_{b_2} f, \Hat{H}_{a b} = H_{a b} f, V = \bar{\phi}^H, V = \bar{\phi}^H$, and we have Rank($V$) = 1. Similarly, the Rank-1 constraint is firstly discarded. Then with the properties of matrix traces, the original problem can be equivalently given by:

$$\max_\bar{\phi} \alpha$$

s.t. $\gamma_0 \text{Tr}(\Hat{H}_{b_2} \Hat{V}) + 1 \geq 2^\alpha \frac{\text{Tr}(\Hat{V} \Hat{H}_{b_1})}{\text{Tr}(\Hat{V} \Hat{H}_{a b})}$

(16b)

$$\bar{\phi} = 1 (n = 1, 2 \cdots N + 1), \bar{\phi} \geq 0$$

(16c)

$$\alpha \geq 0.$$ 

(16d)

Clearly, (16a), (16c) and (16d) are convex constraints while (16b) remains non-convex following from the fact that both sides of the constraint contain fractional operations. To tackle it, slack variables $\phi, t_1, t_2, t_3, t_4$ are introduced, and after simple exponential operations, the problem can be equivalently derived as:

$$\gamma_0 \text{Tr}(\Hat{H}_{b_2} \Hat{V}) + 1 \geq 2^\rho$$

(17a)
\[ y_0 \text{Tr}(\hat{H}_d \hat{V}) + 1 \leq 2^b \]  \hspace{1cm} (17b)

\[ t_1 - t_2 \geq \alpha + \varphi \]  \hspace{1cm} (17c)

\[ \text{Tr}(\hat{V} \hat{H}_{b_i}) \leq 2^b \]  \hspace{1cm} (17d)

\[ \text{Tr}(\hat{V} \hat{H}_{b_i}) \geq 2^b \]  \hspace{1cm} (17e)

\[ 2b - t_i \leq \varphi. \]  \hspace{1cm} (17f)

Further, (17b) and (17d) are still non-convex because both sides of the constraints are convex. To deal with it, SCA algorithm based on first-order Taylor expansion is used to approximate the exponential function of the right-hand side (17b) and (17d) in an iterative manner\[30\]. At the \( k\)th iteration, \( t_2[k], \gamma_1[k] \) are denoted as the value solved at current iteration, and \( t_2[k], \gamma_1[k] \) are denoted as a feasible point, then we can have the lower bound of the exponential function:

\[
2^{b[k]} \geq 2^{b[k]}[1 + (t_2[k] - \gamma_1[k]) \ln 2] \quad (18a)
\]

\[
2^{b[k]} \geq 2^{b[k]}[1 + (t_3[k] - \gamma_3[k]) \ln 2]. \quad (18b)
\]

With (18), (17b) and (17d) can be written as the following linear inequality

\[
y_0 \text{Tr}(\hat{H}_d \hat{V}^{(k)}) + 1 \leq 2^{b[k]}[1 + (t_2[k] - \gamma_1[k]) \ln 2] \]  \hspace{1cm} (19a)

\[
\text{Tr}(\hat{V}^{(k)} \hat{H}_{b_i}) \leq 2^{b[k]}[1 + (t_3[k] - \gamma_3[k]) \ln 2], \]  \hspace{1cm} (19b)

where \( \hat{V}^{(k)} \) is the solution at the \( k\)th iteration.

Finally, the approximated form of problem (16) is given by problem (20), which is a convex SDP problem and can be solved via the CVX solver.

\[
\max_{\hat{V}^{(k)}} \alpha \quad \text{s.t.} \quad (16c), (16d), (17a), (17e), (17f), (19a), (19b). \]  \hspace{1cm} (20c)

It is worth mentioning that due to relaxing the Rank-1 constraint in (16), there is no guarantee that the Rank of the obtained matrix \( \hat{V}^{(k)} \) is one and the vector \( \hat{v}^{(k)} \) may not be directly obtained. If the obtained solution is not Rank-1, Gaussian randomization is utilized for recovering vector approximately \[13\]. Then according to the definition of \( \hat{v}^{(k)} \), we can easily derive the phase shift matrix \( \Phi \).

**Convergence Analysis:** We can obtain the approximation with the current solution by iteratively updating until constraints (19a) and (19b) hold with equality, which indicates that problem (20) is optimally solved. Moreover, from (19a) and (19b), the \((k + 1)\)th iteration admits a larger feasible region than the \( k \)th iteration, since the \((k + 1)\)th iterative solution cannot be inferior to the \( k \)th iterative solution. Furthermore, since the power of jamming signal is limited and \( \hat{v}^{(k)} \) obeys unit-modulus constraint, the convergence of SCA-based algorithm can be guaranteed.

### 3.3 Overall algorithm

To summarize, the outline of solving problem (11) is listed in Algorithm 1, where \( L \) and \( K \) represent the maximum iteration number for problem (11) and problem (20), respectively.

**Convergence Analysis:** For the \( l \)th alternating iteration, \( R(\hat{v}^{(l)}, f^{(l)}) \) is the objective value, where \( (\hat{v}^{(l)}, f^{(l)}) \) is the feasible solution. Then for the \((l + 1)\)th iteration, \( (\hat{v}^{(l+1)}, f^{(l+1)}) \) is the optimal solution of problem (14), and \( (\hat{v}^{(l+1)}, f^{(l+1)}) \) is the feasible solution of problem (16). Then, we have

\[
R(\hat{v}^{(l)}, f^{(l)}) \geq R(\hat{v}^{(l)}), f^{(l+1)}), \quad (21)
\]

where inequality (a) holds owing to that problem (14) is solved optimally in Step 3, equality (b) holds since Step 4 converges and problem (16) has no relationship with \( f^{(l)} \), which has no effect on the object value. This means that the object value given in the \((l + 1)\)th iteration is not larger than that in the \( l \)th iteration. Furthermore, along with the QoS constraint, the jamming power is lower bounded and thus it must guarantee to converge after some iterations.

**Complexity Analysis:** The main complexity of proposed algorithm lies on Steps 3 and 4. For Step 3, the complexity of (14), denoted as \( \mathcal{O}_1 \), depends on the number and size of variables and constraints, which is listed in Table 1. Similarly, the complexity of Step 4 is \( I^{(l)}_{\alpha} \times \mathcal{O}_2 \), where \( I^{(l)}_{\alpha} \) denotes the SCA iteration number and \( \mathcal{O}_2 \) denotes the complexity of (20). Thus, the major complexity of proposed algorithm is given
4 NUMERICAL RESULTS

In this section, numerical results are presented to evaluate the performance of our proposed scheme. Simulation setups are depicted in Figure 2, where Alice, Bob, IRS are located at (0,0), (50,20), (50,0) in meter (m), respectively. Mon is deployed at a vertical distance of 15 m from Alice, and is moved horizontally to show the performance at different locations. The link between node i to node j is modelled as $b_{ij} = \sqrt{\rho_i d_{ij}^{-\gamma_i} g_j}$, where $\gamma_i$ represents the path loss exponent, $d_{ij}$ denotes the distance from i to j, and $\rho_i$ denotes the path loss at the reference distance $d_0 = 1m$. Besides, the small-scale fading component $g_j$ is given by

$$g_j = \sqrt{\frac{1}{1 + \beta_j} s_{ij}^{NLOS}} + \sqrt{\frac{\beta_j}{1 + \beta_j} s_{ij}^{LOS}}, \quad (22)$$

where $\beta_j$ is the Rician factor, $s_{ij}^{NLOS}$ and $s_{ij}^{LOS}$ represent the non-line-of-sight(NLoS) components and LoS components, respectively. In order not to be discovered by the suspicious users, Mon lies on the NLoS range of Alice and Bob, and then we set $\beta_{ar} = \beta_{ab} = 0$, $\epsilon_{ar} = \epsilon_{ab} = 3$. The condition of the communication link from Alice to Bob is better than that of the eavesdropping link and we set $\beta_{ar} = 10$, $\epsilon_{ab} = 3$. Considering that IRS is deployed vertically high, where exists less scatter[31], we set $\beta_{ar} = \beta_{ab} = 10$, $\epsilon_{ar} = \epsilon_{ab} = 2.5$, $\epsilon_{ma} = 3$. The rest parameters are listed as follows: $\rho_0 = -30dB$, $M_a = 4$, $\rho = 0.2$, $\lambda = 1$, $\sigma_0^2 = -70dBm$, $\epsilon = 0.001$, and $N = 8$. The following results are generated by the average of simulating over 200 randomly channel realization.

With the purpose of evaluating the advantage of our scheme, three benchmark schemes are discussed as follows:

- **Scheme 1**: Without IRS. In this case, the model without IRS is similar to [12], by setting $\hat{\mathbf{v}} = \mathbf{0}_{1 \times N}$, and solving problem (14).
- **Scheme 2**: With IRS and random reflecting coefficients (IRS random). In this case, the reflecting coefficient of each element at the IRS is randomly chosen in $(0,2\pi)$. Then, we only design the jamming beamforming vector by solving problem (14).
- **Scheme 3**: With AF relay. In this case, the IRS is replaced with an AF relay. Different from acting as the role of helper or jammer in [32], the relay is exploited to enhance the monitor’s signal. In order to make a fair comparison with our proposed scheme, the AF relay equipped with $N$ antennas forwards the received signal to Mon with amplification factor $\alpha$ as described in [33]. The jamming beamforming vector is designed similar to Scheme 1 and the AF beamforming vector is designed as shown in Appendix 8.3.

Figure 3 shows the convergence behavior of our proposed algorithm when the number of reflection elements varies from 15 to 50 and Mon is located at (45,−15) and (60,−15), respectively. We can observe that with the increase of iteration numbers, the minimum jamming power gradually decreases. After

![Figure 2](image-url) Simulation setup

![Figure 3](image-url) Convergence trajectory versus the iteration numbers of alternating optimization ($N_{IRS} = 15, M_a = 4$)
The minimum jamming power versus different horizontal distance from Mon to Alice ($N_{IRS} = 20, M_e = 4$)

![Figure 4](image-url)

**Figure 4** The minimum jamming power versus different horizontal distance from Mon to Alice ($N_{IRS} = 20, M_e = 4$)

about 5 iterations, the object values all reach a stable value, which validates the convergence analysis given in Section 3.3.

Figure 4 shows the required minimum jamming power versus different horizontal distance from Alice to Mon. As Mon moves away from Alice, the jamming power increases gradually. The reason lies in that the farther the distance from Mon to Alice, the higher the large scale path loss and thus the average SINR at Mon is lower than that at Bob. Meanwhile, under the same condition, the proposed scheme significantly outperforms Scheme 1 and Scheme 2, which can be concluded from (11). By carefully designing reflection coefficients, the SINR at Mon can be improved, while the SINR at Bob can be degraded. Thus the required jamming power is reduced. Besides, the performance of Scheme 3 with amplification factor $\alpha$ varying from 1 to 5 is plotted. We can see that as the amplification factor increases, the required jamming power decreases gradually. When the AF relay forwards signal without amplification, the performance is worse than that of our proposed scheme. The reason is that AF relay introduces additional noise although it forwards the signal to Mon with beamforming technique. However, no additional noise is brought by the IRS[34]. When the amplification factor is set 5, the performance of the scheme with AF relay is slightly better than the scheme with IRS, while the IRS is passive and is easy to be conformal to the surrounding objects. Thus, in proactive eavesdropping system, there exists a trade-off between deploying the AF relay or the IRS. If the channel condition of the monitor is poor and there exists a concealed location, deploying relay is a good choice. Otherwise, the IRS should be considered. Moreover, we can also see that the performance of Scheme 2 is nearly equal to Scheme 1. It means that IRS, without well designed phase shifts, may not enhance the performance, although the number of channel paths increases.

The performance gains versus the number of jamming antennas at Mon is plotted in Figure 5 when Mon is located at (45, $-15$). It can be observed that as the number of jamming antennas increases, the required jamming power decreases gradually. Moreover, under the same constraints, the proposed scheme achieves better performance than the other three schemes. Moreover, we can see that the performance of our proposed scheme with reducing one RF link is nearly consistent with that of Scheme 1, which illustrates that the proposed scheme can effectively reduce the complexity of the monitor.

In Figure 6, we further show the performance gains versus the number of reflecting units when Mon is located near Alice (45, $-15$) and near IRS (60, $-15$), respectively. We can observe that as the number of reflection units increases, the required jamming power both decrease gradually. The reason lies in that IRS with larger units can provide more adjustable dimension for inequality (15a). In addition, simulations also show that in some cases, the channel quality can be moderated only with passive IRS to ensure reliable reception, which further illustrates the advantages of our scheme in covertly overhearing.

![Figure 5](image-url)

**Figure 5** The minimum jamming power versus the number of jamming antennas at Mon ($N_{IRS} = 20$)

![Figure 6](image-url)

**Figure 6** The minimum jamming power versus the number of reflecting units ($M_e = 4$)
5 | CONCLUSION AND FUTURE WORK

We have presented an investigation of IRS-assisted proactive eavesdropping system. The jamming power minimization problem under the constraint of QoS is established by jointly designing the reflection coefficients and the jamming beamforming vector. To tackle this non-convex problem, an algorithm with alternate optimization and SCA techniques is developed to find a sub-optimal solution. Numerical results validate the advantages of our proposed scheme compared with existing schemes in various system setups. In addition, the assumption of global CSI may be overly optimistic. Therefore, it is worthwhile to consider the optimal resource allocation solution under the non-ideal CSI or the statistical CSI in future study.

ACKNOWLEDGMENTS

This work was supported in part by the National Natural Science Foundation of China under Grant 61871404, in part by the Scientific and Technological Key Project of Henan Province under Grant 212102210559, and in part by National Natural Science Foundation Innovative Groups Project of China under Grant 61521003.

ORCID

Jie Yang https://orcid.org/0000-0002-5848-0954

REFERENCES

1. AlNahari, A., et al.: Energy and spectral efficiency of secure massive mimo downlink systems. IET Communications 13(10), 1364–1372 (2019)
2. Tatar Mamaghanii, M., Kühestani, A., Wong, K.: Secure two-way transmission via wireless-powered untrusted relay and external jammer. IEEE Transactions on Vehicular Technology 67(9), 8451–8465 (2018)
3. Mamaghani, M.T., Abbas, R.: Security and reliability performance analysis for two-way wireless energy harvesting based untrusted relaying with cooperative jamming. IET Communications 13(4), 449–459 (2018)
4. Mamaghani, M.T., Hong, Y.: Improving phy-security of uav-enabled transmission with wireless energy harvesting: Robust trajectory design and communications resource allocation. IEEE Transactions on Vehicular Technology 69(8), 8586–8600 (2020)
5. Xu, J., Duan, L., Zhang, R.: Surveillance and intervention of infrastructure-free mobile communications: A new wireless security paradigm. IEEE Wireless Communications 24(4), 152–159 (2017)
6. Xu, J., Duan, L., Zhang, R.: Proactive eavesdropping via jamming for rate maximization over rayleigh fading channels. IEEE Wireless Communications Letters 5(1), 80–83 (2015)
7. Moon, J., et al.: Proactive eavesdropping with jamming and eavesdropping mode selection. IEEE Transactions on Wireless Communications 18(7), 3726–3738 (2019)
8. Zeng, Y., Zhang, R.: Wireless information surveillance via proactive eavesdropping with spoofing relay. IEEE Journal of Selected Topics in Signal Processing 10(8), 1449–1461 (2016)
9. Hu, G., Cai, Y.: Uavs-assisted proactive eavesdropping in af multi-relay system. IEEE Communications Letters 24(3), 501–505 (2020)
10. Zhao, J., Liu, Y.: A survey of intelligent reflecting surfaces (irs): Towards 6g wireless communication networks. arXiv:1907.04789 (2019)
11. Wu, Q., Zhang, R.: Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming. IEEE Transactions on Wireless Communications 18(11), 5394–5409 (2019)
12. Guo, H., et al.: Weighted sum-rate maximization for intelligent reflecting surface enhanced wireless networks. In: 2019 IEEE Global Communications Conference, Big Island (2019)
13. Cui, M., Zhang, G., Zhang, R.: Secure wireless communication via intelligent reflecting surface. IEEE Wireless Communications Letters 8(5), 1410–1414 (2019)
14. Dong, L., Wang, H.M.: Secure mimo transmission via intelligent reflecting surface. IEEE Wireless Communications Letters 9(6), 787–790 (2020)
15. Huang, C., et al.: Reconfigurable intelligent surfaces for energy efficiency in wireless communication. IEEE Transactions on Wireless Communications 18(8), 4157–4170 (2019)
16. Tang, Z., et al.: Physical layer security of intelligent reflective surface aided noma networks. arXiv:2011.03417 (2020)
17. Lyu, B., et al.: Irs-based wireless jamming attacks: When jammers can attack without power. IEEE Wireless Communications Letters 9(10), 1663–1667 (2020)
18. Zhong, C., et al.: Multi-antenna wireless legitimate surveillance systems: Design and performance analysis. IEEE Transactions on Wireless Communications 16(7), 4585–4599 (2017)
19. Zheng, G., et al.: Improving physical layer secrecy using full-duplex jamming receivers. IEEE Transactions on Signal Processing 61(20), 4962–4974 (2013)
20. Ma, R., et al.: Improving physical layer security jointly using full-duplex jamming receiver and multi-antenna jammer in wireless networks. IET Communications 13(10), 1530–1536 (2019)
21. Moon, J., et al.: Proactive eavesdropping with full-duplex relay and cooperative jamming. IEEE Transactions on Wireless Communications 17(10), 6707–6719 (2018)
22. Zhang, H., Duan, L., Zhang, R.: Jamming-assisted proactive eavesdropping over two suspicious communication links. IEEE Transactions on Wireless Communications 19(7), 4817–4830 (2020)
23. Yang, Y., et al.: Intelligent reflecting surface meets ofdm: Protocol design and rate maximization. IEEE Transactions on Communications 68(7), 4522–4535 (2020)
24. Taha, A., Albabieh, M., Alkhalteeb, A.: Enabling large intelligent surfaces with compressive sensing and deep learning. arXiv:191041036 (2019)
25. Taha, A., et al.: Deep reinforcement learning for intelligent reflecting surfaces: Towards stand-alone operation. arXiv:2004.00706 (2020)
26. Huang, C., et al.: Reconfigurable intelligent surfaces for energy efficiency in wireless communication networks. arXiv:2011.03417 (2020)
27. Hu, X., Wang, J., Zhong, C.: Statistical csi based design for intelligent reflecting surface assisted mimo systems. Science China Information Sciences, Springer, Berlin 63(12), 1–10 (2020)
28. Peng, Z., et al.: Analysis and optimization for irs-aided multi-pair communications relying on statistical csi. arXiv:2007.11704 (2020)
29. Zhang, L., et al.: Robust beamforming design for intelligent reflecting surface aided cognitive radio systems with imperfect cascaded csi. arXiv:2004.04595 (2020)
30. Chen, D., et al.: Both worst-case and chance-constrained robust secure swipt in miso interference channels. IEEE Transactions on Information Forensics and Security 15(2), 306–317 (2018)
31. Guan, X., Wu, Q., Zhang, R.: Intelligent reflecting surface assisted secrecy communication: Is artificial noise helpful or not? IEEE Wireless Communications Letters 9(6), 778–782 (2020)
32. Jiang, X., et al.: Proactive eavesdropping in relaying systems. IEEE signal processing letters 24(6), 917–921 (2017)
33. Han, Y., et al.: Large intelligent surface-aided cognitive radio systems with imperfect cascaded csi. IEEE Communications 19(6), 6823–6824 (2019)
34. DiRenzo, M., et al.: Reconfigurable intelligent surfaces vs. relaying: Differences, similarities, and performance comparison. IEEE Open Journal of the Communications Society 1, 798–807 (2020)
APPENDICES

Proof of Proposition 1

When Mon works on passive eavesdropping mode, the received signal at Mon is expressed as:

\[ y_r = \sqrt{P_r} \left( h_{rl}^f + h_{rl}^s \Phi H_{ar} \right) w + n_2. \]  

(A1)

In order to facilitate the subsequent derivation, a series of scalars are introduced to represent each channel vector in (23). For example, the channel vector \( h_{rl} \) is expanded as \( h_{rl} = (h_{rl,1}, h_{rl,2}, \ldots, h_{rl,N}, h_{rl,M}) \). Then, (23) can be rewritten as:

\[ y_r = \sqrt{P_r} \sum_{n=1}^N (h_{rl,n} w_n + \sum_{m=1}^M h_{rl,m} e^{\theta_n}) s + n_2. \]  

(A2)

Next, with (12), the SINR at Mon is expressed as:

\[ \mathcal{Y}_r = \frac{\left( \sum_{m=1}^M h_{rl,m} e^{\phi_m} \right)^2}{\sum_{n=1}^N |h_{rl,n}|^2}. \]  

(A3)

Let \( Z = \left| \sum_{m=1}^M h_{rl,m} e^{\phi_m} \right|^2 \), and the limit of \( Z \) as \( N \to \infty \) is firstly discussed.

Specifically, the following inequality can be given by:

\[ Z = \left| \sum_{n=1}^N \sum_{m=1}^M h_{rl,n,m} e^{\phi_m} \right|^2 \leq \left( \sum_{n=1}^N |h_{rl,n}| \right)^2 \left( \sum_{m=1}^M |h_{rl,m}|^2 \right). \]  

(A4)

where the inequality (a) holds due to the triangle inequality and we can easily obtain that the equality holds if the reflection coefficients are adjusted to satisfy \( \phi_m = \arg(\sum_{n=1}^N h_{rl,n,m} e^{\phi_m}) - \arg(h_{rl,n}), \forall n \in N, \phi_0 \) is an arbitrary phase.

Then let \( X_n = h_{rl,n} \), \( Y_n = \sum_{m=1}^M h_{rl,n,m} e^{\phi_m} \). Similar to [11], consider \( h_{rl,n} \sim \mathcal{CN}(0, \rho_{rl}^2) \), \( h_{rl,n,m} \sim \mathcal{CN}(0, \rho_{rl,m}^2), \rho_{rl,m} \sim \mathcal{CN}(0, \rho_{rl}^2) \).

According to the properties of Gaussian distribution, \( |X_n|^2 \) is a exponential RV with parameter \( \lambda_{Y_n} = \rho_{rl}^2 \). By invoking the central limit theorem, \( Y_n \) can be approximated as a Gaussian RV with zero-mean and variance \( \sigma_{Y_n}^2 = \lambda_{Y_n} \rho_{rl}^2 \). Then \( |Y_n|^2 \) can be approximated as an exponential RV with parameter \( \lambda_{Y_n} = \rho_{rl}^2 \).

Then let \( Z' = \sum_{n=1}^N |X_n|^2 |Y_n|^2 \), and we can have \( E(Z') = N|\lambda_{Y_n}| |\rho_{rl}|^2 \). Then as \( N \to \infty \), we have \( E(Z') \to \infty \).

Due to \( Z' \leq Z \) and (25), we can easily have as \( N \to \infty \), \( \mathcal{Y}_r \to \infty \), which means that by simply designing the phase shifts to enhance the signal at Mon, the average SINR at Mon approaches infinity as \( N \to \infty \). Thus, the proof is completed.

Proof of Proposition 2

We first write the Lagrangian multiplier function for problem (14) as

\[ L(F; \varepsilon_1, Y) = \text{Tr}(F) + \varepsilon_1 (\lambda (Y) \text{Tr}(H_{r_2}^f F) + 1) - (Y_0 \text{Tr}(H_{r_2}^f F) + 1) - \text{Tr}(Y F), \]  

(A5)

where \( \varepsilon_1 \geq 0, Y \geq 0 \) is the matrix dual variable associated with \( F \geq 0 \). Since problem (14) is convex and Slater condition holds, the Karush-Kuhn-Tucker(KKT) conditions are necessary and sufficient for establishing the optimal solution. Inspired by the definition of KKT conditions, we have

\[ \text{constraints}(14b) - (14c) \]  

(A6)

\[ \varepsilon_1 (\lambda (Y_0 \text{Tr}(H_{r_2}^f F) + 1) - (Y_0 \text{Tr}(H_{r_2}^f F) + 1)) = 0 \]  

(A7)

\[ \text{Tr}(Y F) = 0 \to Y F = 0 \]  

(A8)

\[ \frac{\partial L}{\partial F} = I + \varepsilon_1 \lambda Y_0 H_{r_2}^f - \varepsilon_1 Y_0 H_{r_2}^f - Y = 0. \]  

(A9)

With \( Y F = 0 \) and multiplying (28d) by \( F \), then we have

\[ F(I + \varepsilon_1 \lambda Y_0 H_{r_2}^f) = \varepsilon_1 Y_0 F H_{r_2}^f. \]  

(A10)

It is easy to derive that \( \varepsilon_1 > 0 \). The reason is that if \( \varepsilon_1 = 0 \), we have \( I = Y \), which is contradict to (28c). Meanwhile, we can have \( I + \varepsilon_1 \lambda Y_0 H_{r_2}^f \) is a full-rank matrix. Hence, according to the nature of full-rank matrix, we can have

\[ \text{Rank}(F(I + \varepsilon_1 \lambda Y_0 H_{r_2}^f)) = \text{Rank}(F) \]  

(A11)

Further, for the right hand side of (29), the following inequality can be derived

\[ \text{Rank}(\varepsilon_1 Y_0 F H_{r_2}^f) \leq \min(\text{Rank}(F), \text{Rank}(H_{r_2}^f)) \leq 1. \]  

(A12)

The inequality (a) holds due to the property of rank of matrix. The inequality (a) holds due to the fact that the rank of matrix
\( \hat{H}_{b2} \) is one. Then, with (30) and (31), we have

\[
\text{Rank}(F) \leq 1. \tag{A13}
\]

Therefore, the optimal solution \( F \) satisfies \( \text{Rank}(F) \leq 1. \)

**Beamforming design of AF relay**

In this appendix, the optimal beamforming vector of AF relay is designed. With AF relay, the signal received by Mon is the combination of the forwarding signal \( y_{\text{relay}} \) and the direct signal, which is expressed as

\[
r = \sqrt{P_{t}}h_{w}^{H}w + \alpha h_{r}^{H}w_{\text{relay}} + \sqrt{\rho}h_{w}^{H}f_{a} + n_{2}, \tag{A14}
\]

where \( y_{\text{relay}} = \sqrt{P_{t}}h_{w}^{H}w + n_{3}, \) \( h_{w}^{H} \) is the first row of \( H_{w}^{H}, \) \( n_{3} \sim \mathcal{CN}(0, \sigma_{0}^{2}) \) denotes the AWGN at the relay. \( w \) denotes the transmit beamforming vector sent by Alice, which is consistent with (12), \( w_{\text{relay}} \) denotes the forward beamforming vector sent by the relay. Then, the SINR at Mon is given by

\[
\gamma = \frac{P_{t}\left| (h_{w}^{H} + \alpha h_{r}^{H}w_{\text{relay}}h_{w}^{H})w \right|^{2}}{\sqrt{\rho}h_{w}^{H}f_{a} + \sigma_{0}^{2}\left( \left| \alpha h_{r}^{H}w_{\text{relay}} \right|^{2} + 1 \right) \left\| h_{w} \right\|^{2}}. \tag{A15}
\]

It is easy to know that the introduction of the relay has little effect on \( \gamma_{b} \). To simplify the problem, the optimal problem is transformed to solve the optimal \( w_{\text{relay}} \) to maximize \( \gamma_{b} \). And substituting the expression (12) into (34), we have

\[
\gamma = \frac{P_{t}\left| (h_{w}^{H} + \alpha h_{r}^{H}w_{\text{relay}}h_{w}^{H})h_{ab} \right|^{2}}{\left( \sqrt{\rho}h_{w}^{H}f_{a} + \sigma_{0}^{2}\left( \left| \alpha h_{r}^{H}w_{\text{relay}} \right|^{2} + 1 \right) \right) \left\| h_{w} \right\|^{2}} \tag{A16}
\]

Then let \( d = \alpha h_{r}^{H}w_{\text{relay}} = d_{e}e^{j\phi}, a = h_{w}^{H}h_{ab}, b = h_{w}^{H}h_{ab}, \epsilon = \rho |h_{w}^{H}f_{a}|^{2}, \lambda = \lambda e^{j\theta}. \) We can observe that variables \( d, \lambda \) are scalars, and the following inequality can be derived

\[
\gamma_{s}(d) = \frac{P_{t}|a(1 + d\lambda)|^{2}}{(\epsilon + \sigma_{0}^{2}(|d|^{2} + 1)) \left\| h_{w} \right\|^{2}},
\]

\[
\gamma_{s}(d) \leq \frac{P_{t}|a(1 + d\lambda)|^{2}}{(\epsilon + \sigma_{0}^{2}(\lambda^{2} + 1)) \left\| h_{w} \right\|^{2}} \tag{A17}
\]

\[
\gamma_{s} = \gamma_{s}(d) \leq \alpha^{2} \left\| h_{s} \right\|^{2}, \tag{A18}
\]

where the equality in (36a) holds if \( \phi = -\theta \), and the equality in (36b) holds if \( w_{\text{relay}} = h_{w}^{H}c_{\eta} / \| h_{w} \| \), where \( \eta \in (0, 2\pi) \).

Then maximizing \( \gamma_{s} \) is equal to find the optimal \( \lambda \), and by computing the derivative of (36a), we have

\[
\gamma'_{s}(d) = \frac{-P_{t}|d|^{2}(1 + \lambda^{2})}{\sigma_{0}^{2}\left\| h_{s} \right\|^{2}(\epsilon + \sigma_{0}^{2}(d^{2} + 1))}, \quad 0 < \lambda \leq \alpha \left\| h_{s} \right\|. \tag{A19}
\]

Regardless of the range of \( \lambda \), it is easy to know the function \( \gamma_{s}(d) \) is monotonically increasing at the interval \( -\frac{1}{\lambda} < \lambda < (1 + \frac{\lambda}{\sigma_{0}^{2}}) \). Then consider the interval ranges of \( \lambda \), we have the optimal \( \lambda_{\text{opt}} = \min\{\alpha \left\| h_{s} \right\|, (1 + \frac{\lambda}{\sigma_{0}^{2}}) \} \).

Taking into account the above discussion, there are many possible solutions for \( w_{\text{relay}}^{\text{opt}} \), which don’t affect the optimal solution. And one possible solution can be derived as

\[
w_{\text{relay}}^{\text{opt}} = \begin{cases} \frac{h_{w}^{H}e^{-j\phi}}{\| h_{w} \|} & \text{if } \left\| h_{w} \right\| \leq \lambda_{\text{opt}} \\ \frac{h_{w}^{H}e^{-j\phi}}{\| h_{w} \|} \left( 1 + \frac{\lambda_{\text{opt}}}{\sigma_{0}^{2}} \right) & \text{if } \left\| h_{w} \right\| > \lambda_{\text{opt}} \end{cases} \tag{A20}
\]