Beamforming-Enabled Covert Communications for Multi-Position Warden

Mingqian Liu, *Member, IEEE*, Zhaoxi Wen, Yunfei Chen, *Senior Member, IEEE*, Jie Tang, *Senior Member, IEEE*, Kai-Kit Wong, *Fellow, IEEE*, Xiaoniu Yang

***Abstract*—In covert communications, the position of the warden has a variety of situations, which leads to different scenes that require different covert communication schemes to ensure the security. In response to this situation, in this paper, a beamforming optimization method of covert communications for multi-position warden is proposed. Firstly, we formulate a general optimization problem and optimize it to maximize the covert communication rate of the user based on Dinkelbach's transform. Subsequently, according to the optimized general optimization problem, we propose three schemes for three scenes corresponding to different fixed warden positions, using appropriate technologies for assistance in each scheme. Specifically, the intelligent reflecting surface (IRS) is used in Scene 1 and the integrated communication and jamming (ICAJ) is used in Scenes 2 and 3, and these technologies can assist the covert communication. Moreover, we propose an alternate optimization (AO) algorithm to solve the optimization problem of Scene 1 for its optimal covert communication performance. Additionally, we also propose an AO algorithm to solve the optimization problems of Scenes 2 and 3 to optimize the active beamforming. Simulation results demonstrate the effectiveness of all three proposed schemes, that outperform their respective benchmark schemes.**

***Index Terms*—Beamforming optimization, covert communications, integrated communication and jamming, intelligent reflecting surface, multi-position warden.**

I. Introduction

WIRELESS communication has brought a lot of convenience to our life for its rapid developments and wide applications [1], such as the Internet of Things (IoT) and the intelligent transportation system (ITS) [2], [3]. However, the wireless communication channel is an open channel, allowing all legal users and illegal eavesdroppers to receive information from it [4]. This situation provides illegal eavesdroppers with the opportunity to eavesdrop on confidential information sent to the legal users. Therefore, wireless communication faces serious security threats, and safeguarding confidential information in the open channel becomes a significant challenge [5]. The traditional research on wireless communication security primarily focuses on protecting communication content to prevent eavesdropping such as physical layer security (PLS) and cryptography [6]–[8]. However, these methods only secure the content of the information and do not protect the information transmission process. Consequently, if the information content is inadequately encrypted, its security will still be compromised. Furthermore, since the information can be accessed by anyone, the privacy of the information is not guaranteed. To address the limitations of traditional wireless communication security methods, the covert communication technology has been presented as a method to protect the information transmission process [9], [10]. The purpose of covert communications is to transmit the communication information without being detected by the warden when the warden is monitoring the channel, thereby concealing the communication process while ensuring that legitimate users can decode [11]. Therefore, covert communications have the advantage of not being alarmed by warden and have been applied in a variety of scenes, such as multiple access channels, interweave cognitive radio networks, and discrete memoryless channels [12]–[14].

Recently, covert communications have been extensively studied [15]. For example, reference [16] used two unmanned aerial vehicles (UAVs) to assist covert communications, one as a communicator to send information and the other as a jammer to jam the monitor. Reference [17] studied the relay selection in covert communications and proposed a safe zone scheme, which has the advantages of covert transmission and device-to-device (D2D) communication simultaneously. In the covert millimeter-wave (mmWave) communication system, reference [18] proposed a dual-decomposition successive convex approximation algorithm for the joint design of beam training and data transmission. On multi-carrier channels, finite blocklength covert communication was studied in [19] to maximize the covert communication rate under the constraint of a finite power budget. In [20], a method of using a jammer to generate intermittent artificial noise (AN) for the warden to enhance covert communications was proposed. Reference [21] proved that faster-than-Nyquist (FTN) signaling has a better maximum transmitting power and covert communication rate than Nyquist signaling in a Rayleigh block fading channel, demonstrating the advantage of FTN signaling in covert

This work was supported by the National Natural Science Foundation of China under Grant U2441250 and 62231027, Natural Science Basic Research Program of Shaanxi under Grant 2024JC-JCQN-63, Innovation Capability Support Program of Shaanxi under Grant 2024RS-CXTD-01 and in part by the Fundamental Research Funds for the Central Universities and the Innovation Fund of Xidian University under Grant YJSJ25007. *(Corresponding author: Zhaoxi Wen.)*

M. Liu, Z. Wen, and X. Yang are with the State Key Laboratory of Integrated Service Networks, Xidian University, Shaanxi, Xi'an 710071, China (e-mail: mqliu@mail.xidian.edu.cn; zx_wen@stu.xidian.edu.cn; yxn2117@126.com).

Y. Chen is with Department of Engineering, University of Durham, Durham, UK DH1 3LE (e-mail: Yunfei.Chen@durham.ac.uk).

J. Tang is with the School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510640, China (e-mail: eejtang@scut.edu.cn).

K. K. Wong is affiliated with the Department of Electronic and Electrical Engineering, University College London, Torrington Place, WC1E 7JE, United Kingdom and he is also affiliated with Yonsei Frontier Lab, Yonsei University, Seoul, Korea (e-mail: kai-kit.wong@ucl.ac.uk).

communications. In addition to the study of the optimization problem, reference [22] studied the repetitive coding of binary phase shift keying (BPSK) modulation, thus obtaining the influence of channel coding on covert communications. Reference [23] considered two scenes in which the warden knows and does not know the transmitter's position, and it used random frequency diverse array-directional modulation (RFDA-DM) to enhance covert communications.

Moreover, to enhance covert communications, intelligent reflecting surface (IRS) is proposed as a low-cost technology that can improve the performance of wireless communication in the channel [24]–[28], and it is divided into many types, such as passive IRS and active IRS. For IoT networks, reference [29] proposed a multiple-input-multiple-output (MIMO) covert communication system assisted by IRS to jointly optimize precoder and IRS passive beamforming with the goal of maximizing the minimum signal-to-noise ratio (SNR) at receivers. The robust transmission design of the covert communication system in the presence of hardware impairments (HWIs) was carried out in [30] with the assistance of IRS, where the transmitting power and IRS passive beamforming were jointly optimized. Reference [31] studied covert communications in satellite communication systems to maximize the minimum user covert communication rate under limited resource constraints in the multi-user scene. Reference [32] used IRS to aid covert communications and invoked the Gauss-Poisson process (GPP) to simulate the position of certain nodes. In [33], aiming at maximizing the effective covert throughput (ECT), two IRS-assisted covert communication schemes with finite blocklength were proposed under the condition of equal and unequal transmit prior probabilities. Different from traditional passive IRS-assisted covert communications, reference [34] studied the covert communication system based on the active IRS and solved the problems brought by active IRS. Reference [35] studied the use of simultaneously transmitting and reflecting IRS (STAR-IRS) assisted covert communications in a multi-antenna scene, and it used a friendly full-duplex receiver with two antennas to assist covert communications. Reference [36] proposed a novel covert communication scheme with the assistance of the UAV-IRS, and it used the alternate optimization (AO) algorithm to effectively achieve the maximization of the covert communication rate.

From above works, we can see that most existing researches on covert communications do not consider the diversity of the warden's position and also overlook the protection of communication content. Traditional covert communications work effectively under the condition that the covertness constraint is satisfied. Further, if the warden is near the transmitter or the warden's distance from transmitter is between near and far, integrated communication and jamming (ICAJ) can be used to protect the content of the communication information while achieving covert communications. Therefore, this paper studies a covert communication system with a multi-position warden and proposes three corresponding covert communication schemes for three scenarios corresponding to three fixed positions of warden. Furthermore, to enhance the performance of covert communications, we optimize the beamforming under appropriate constraints in each proposed scheme. Simulation results show that the three proposed schemes have better performance than that of their respective benchmark schemes.

The main contributions of this paper can be summarized as follows:

- We innovatively consider a covert communication system with a multi-position warden. The fixed positions of the warden in three scenarios are given as: the warden is far from the transmitter (scheme 1), near the transmitter (scheme 2), and the warden's distance from transmitter is between near and far (scheme 3);
- We firstly formulate a general optimization problem and optimize it based on Dinkelbach's transform, and then propose three corresponding covert communication schemes for three scenarios corresponding to three fixed positions of warden based on the optimized general optimization problem. In each scheme, we optimize the beamforming to achieve the best performance of covert communications. Specifically, in scheme 1, Alice only transmits the communication signal with the assistance of IRS, and beamformings of communication and IRS are jointly optimized for better performance of the covert communication system. In scheme 2, Alice transmits the ICAJ signal, and the beamforming of communication and jamming is optimized. Based on scheme 2, scheme 3 changes the objective function to the sum of communication and jamming performance (overall performance), and removes the lower bound of jamming tolerance in constraints for better feasibility;
- For the optimization problem of scheme 1, we optimize the active beamforming and the passive beamforming using the proposed AO algorithm under the constraints of transmitting power, jamming tolerance, minimum covert rate, and covertness. In addition, we also propose an AO algorithm to solve the optimization problems of schemes 2 and 3.

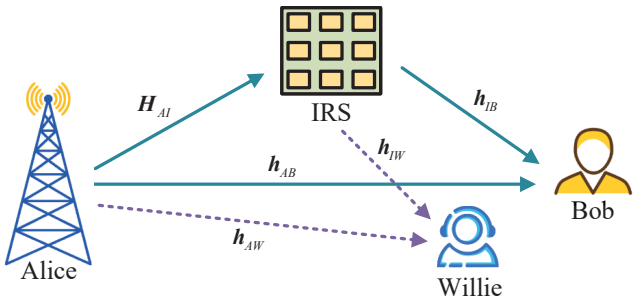
The remainder of this paper is organized as follows. In Section II, the system model is given. Section III formulates and optimizes the general optimization problem, proposes three schemes for different positions of warden, and proposes two AO algorithms to optimize the optimization problems of schemes 1, 2, and 3. Section IV gives the simulation results and analysis. Finally, Section V concludes this paper.

Notation: The symbols used in this paper are represented as follows. Scalars, vectors, and matrices are represented by letters without boldface, boldface lowercase, and boldface uppercase, respectively. $\boldsymbol{A}^H$ denotes the conjugate transpose of $\boldsymbol{A}$. $tr(\boldsymbol{A})$ stands for the trace of $\boldsymbol{A}$. $rank(\boldsymbol{A})$ denotes the rank of $\boldsymbol{A}$. diag(a) refers to a diagonal matrix where the diagonal element is a. $a \sim \mathcal{CN}(\mu, \sigma^2)$ means that a follows the complex Gaussian distribution with mean μ and variance σ^2. Pr$(\cdot)$ represents the probability of an event. $\boldsymbol{A} \succeq 0$ means that $\boldsymbol{A}$ is positive semidefinite. $\mathbb{C}^{m\times n}$ denotes the space of $m \times n$ complex matrices or vectors.

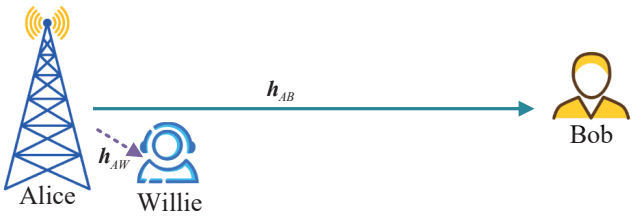
II. System Model

A. System Structure

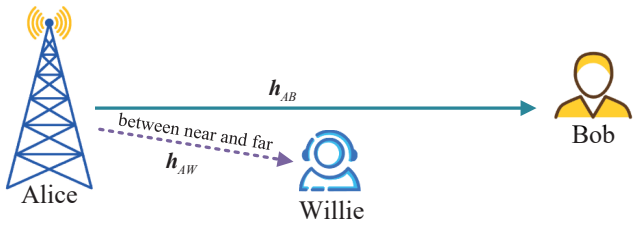
This paper studies a covert communication system with a multi-position warden, and the covert communication system



(a) Scene 1: Willie is far from Alice.



(b) Scene 2: Willie is near Alice.



(c) Scene 3: The Willie's distance from Alice is between near and far.

Fig. 1. Three scenes of the covert communication system with a multi-position warden.

has three scenes, which are shown in Fig. 1. From Fig. 1, in Scene 1, the covert communication system consists of a transmitting base station (Alice), a user (Bob), a warden (Willie), and an IRS. Alice wants to transmit the communication signal to Bob without being discovered by Willie, while Willie monitors the transmission environment and detects whether Alice is transmitting the communication signal to Bob, and the IRS is deployed to enhance the covert communication rate at Bob. In Scenes 2 and 3, the covert communication system consists of Alice, Bob, and a warden. Alice wants to transmit the ICAJ signal to Bob without being detected by Willie, while Willie monitors the transmission environment and detects whether Alice is transmitting the ICAJ signal to Bob.

Assume that Alice is equipped with M antennas, the IRS is equipped with N reflecting elements, and both Bob and Willie are equipped with a single antenna. The phase shift matrix of the IRS can be given by

$$\boldsymbol{\Phi} = \mathrm{diag}(e^{j\phi_1}, e^{j\phi_2}, ..., e^{j\phi_N}), \tag{1}$$

where $\boldsymbol{\Phi} \in \mathbb{C}^{N\times N}$ and $\phi_j \in (0, 2\pi]$ denotes the phase shift of the IRS reflection element with $j \in \{1, 2, ..., N\}$.

To facilitate the processing, we assume that the channel state information (CSI) of all links in the covert communication system is known. The channel matrix between Alice and IRS is expressed as $\boldsymbol{H}_{AI} \in \mathbb{C}^{N\times M}$, and the channel vectors between Alice and Willie, Alice and Bob, IRS and Bob, and IRS and Willie are denoted by $\boldsymbol{h}_{AW} \in \mathbb{C}^{M\times 1}$, $\boldsymbol{h}_{AB} \in \mathbb{C}^{M\times 1}$, $\boldsymbol{h}_{IB} \in \mathbb{C}^{N\times 1}$, and $\boldsymbol{h}_{IW} \in \mathbb{C}^{N\times 1}$, respectively. In addition, all channels are composed of large-scale path loss and small-scale Rayleigh fading to give

$$\boldsymbol{h}_i = \sqrt{Q_0\left(\frac{d_i}{d_0}\right)^{-\alpha_i}} \cdot \widetilde{\boldsymbol{h}}_i, i \in \{AI, AW, AB, IB, IW\}, \tag{2}$$

where Q_0 stands for the path loss at the reference distance d_0, d_i represents the distance between nodes, α_i denotes the path loss exponent corresponding to d_i, $\widetilde{\boldsymbol{h}}_i \sim \mathcal{CN}(0, 1)$ is Rayleigh fading, and when $i = AI$, $\boldsymbol{h}_i$ becomes $\boldsymbol{H}_i$.

B. Signal Model

Let $\mathcal{H}_0$ denote hypothesis that Alice does not transmit the ICAJ signal, and $\mathcal{H}_1$ denote hypothesis that Alice transmits the ICAJ signal to Bob. So the ICAJ signal that Alice transmits can be written as

$$\boldsymbol{x} = \begin{cases} 0, & \mathcal{H}_0 \\ \boldsymbol{w}_c s_i + \boldsymbol{w}_p q, & \mathcal{H}_1 \end{cases} \tag{3}$$

where $\boldsymbol{w}_c \in \mathbb{C}^{M\times 1}$ and $\boldsymbol{w}_p \in \mathbb{C}^{M\times 1}$ are beamforming vector for transmitting communication data symbol $s_i \sim \mathcal{CN}(0, 1)$ and beamforming vector for transmitting jamming signal $q \sim \mathcal{CN}(0, 1)$, respectively.

The received signal at Bob can be represented as follows:

$$\begin{aligned} y_B &= \begin{cases} e_B, & \mathcal{H}_0 \\ (\boldsymbol{H}_{AI}^H \boldsymbol{\Phi}^H \boldsymbol{h}_{IB} + \boldsymbol{h}_{AB})^H \boldsymbol{x} + e_B, & \mathcal{H}_1 \end{cases} \\ &= \begin{cases} e_B, & \mathcal{H}_0 \\ (\boldsymbol{h}_{IB}^H \boldsymbol{\Phi} \boldsymbol{H}_{AI} + \boldsymbol{h}_{AB}^H)\boldsymbol{x} + e_B, & \mathcal{H}_1 \end{cases} \end{aligned} \tag{4}$$

where $e_B \sim \mathcal{CN}(0, \sigma_b^2)$ stands for the noise at Bob.

The signal to interference plus noise ratio (SINR) at Bob can be expressed as:

$$SINR_B = \frac{|(\boldsymbol{h}_{IB}^H \boldsymbol{\Phi} \boldsymbol{H}_{AI}^H + \boldsymbol{h}_{AB}^H)\boldsymbol{w}_c|^2}{|(\boldsymbol{h}_{IB}^H \boldsymbol{\Phi} \boldsymbol{H}_{AI}^H + \boldsymbol{h}_{AB}^H)\boldsymbol{w}_p|^2 + \sigma_b^2}. \tag{5}$$

Using $SINR_B$, we can obtain the covert communication rate as follows:

$$\begin{aligned} R_c &= \log_2(1 + SINR_B) \\ &= \log_2\left(1 + \frac{|(\boldsymbol{h}_{IB}^H \boldsymbol{\Phi} \boldsymbol{H}_{AI}^H + \boldsymbol{h}_{AB}^H)\boldsymbol{w}_c|^2}{|(\boldsymbol{h}_{IB}^H \boldsymbol{\Phi} \boldsymbol{H}_{AI}^H + \boldsymbol{h}_{AB}^H)\boldsymbol{w}_p|^2 + \sigma_b^2}\right). \end{aligned} \tag{6}$$

Also, the received signal at Willie can be given by

$$\begin{aligned} y_W &= \begin{cases} e_W, & \mathcal{H}_0 \\ (\boldsymbol{H}_{AI}^H \boldsymbol{\Phi}^H \boldsymbol{h}_{IW} + \boldsymbol{h}_{AW})^H \boldsymbol{x} + e_W, & \mathcal{H}_1 \end{cases} \\ &= \begin{cases} e_W, & \mathcal{H}_0 \\ (\boldsymbol{h}_{IW}^H \boldsymbol{\Phi} \boldsymbol{H}_{AI} + \boldsymbol{h}_{AW}^H)\boldsymbol{x} + e_W, & \mathcal{H}_1 \end{cases} \end{aligned} \tag{7}$$

where $e_W \sim \mathcal{CN}(0, \sigma_w^2)$ stands for the noise at Willie, and the received effective jamming power at Willie can be given by

$$P_w = |(\boldsymbol{h}_{IW}^H \boldsymbol{\Phi} \boldsymbol{H}_{AI}^H + \boldsymbol{h}_{AW}^H)\boldsymbol{w}_p|^2. \tag{8}$$

If Alice only transmits the communication signal, the term related to $\boldsymbol{w}_p$ in the ICAJ signal should be removed, and the ICAJ signal becomes the communication signal.

C. Covertness Requirement

Willie is a non-emitting warden who only monitors the received signal to determine if Alice is transmitting the signal. According to (7), we can obtain the likelihood function of the received signal at Willie under $\mathcal{H}_0$ as

$$p_0(y_W) = \frac{1}{\pi\sigma_0^2}\exp\left(-\frac{|y_W|^2}{\sigma_0^2}\right), \tag{9}$$

where $\sigma_0^2 = \sigma_w^2$. Similarly, the likelihood function of the received signal at Willie under $\mathcal{H}_1$ can be given by

$$p_1(y_W) = \frac{1}{\pi\sigma_1^2}\exp\left(-\frac{|y_W|^2}{\sigma_1^2}\right), \tag{10}$$

where

$$\begin{aligned}\sigma_1^2 =&|(\boldsymbol{h}_{IW}^H\boldsymbol{\Phi}\boldsymbol{H}_{AI}^H + \boldsymbol{h}_{AW}^H)\boldsymbol{w}_c|^2 \\ &+ |(\boldsymbol{h}_{IW}^H\boldsymbol{\Phi}\boldsymbol{H}_{AI}^H + \boldsymbol{h}_{AW}^H)\boldsymbol{w}_p|^2 + \sigma_w^2.\end{aligned} \tag{11}$$

Generally, we assume that the prior probabilities of $\mathcal{H}_0$ and $\mathcal{H}_1$ are equal. Willie wants to minimize the detection error using the likelihood ratio test as

$$\frac{p_1(y_W)}{p_0(y_W)} \underset{\mathcal{D}_0}{\overset{\mathcal{D}_1}{\gtrless}} 1, \tag{12}$$

where $\mathcal{D}_0$ and $\mathcal{D}_1$ are binary decisions that correspond to hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively. In addition, the total detection error probability of Willie is the sum of the false alarm probability and the missed detection probability, which can be expressed as:

$$\begin{aligned}\xi &= \Pr(\mathcal{D}_1|\mathcal{H}_0) + \Pr(\mathcal{D}_0|\mathcal{H}_1) \\ &= \Pr(p_1(y_W) \geq 1|\mathcal{H}_0) + \Pr(p_0(y_W) \leq 1|\mathcal{H}_1).\end{aligned} \tag{13}$$

However, (13) cannot be used in the optimization problem, so we use the lower bound of ξ, which can be expressed as [30]

$$\xi \geq 1 - \sqrt{\frac{1}{2}\mathcal{D}(p_0(y_W)||p_1(y_W))}, \tag{14}$$

where $\mathcal{D}(p_0(y_W)||p_1(y_W))$ denotes the Kullback-Leibler (KL) divergence from $p_0(y_W)$ to $p_1(y_W)$ with

$$\mathcal{D}(p_0(y_W)||p_1(y_W)) = \ln\left(\frac{\sigma_1^2}{\sigma_0^2}\right) + \frac{\sigma_0^2}{\sigma_1^2} - 1. \tag{15}$$

Property 1. *The KL divergence can be calculated using* (15).

Proof. The proof is given in Appendix A. □

In covert communications, the covertness constraint is expressed as

$$\xi \geq 1 - \varepsilon, \forall\varepsilon \geq 0, \tag{16}$$

where ε is used to determine the required covertness level with a small value. The smaller the value of ε is, the higher the required covertness level will be. According to [37], we can obtain a much stricter covertness constraint as follows:

$$\mathcal{D}(p_0(y_W)||p_1(y_W)) \leq 2\varepsilon^2. \tag{17}$$

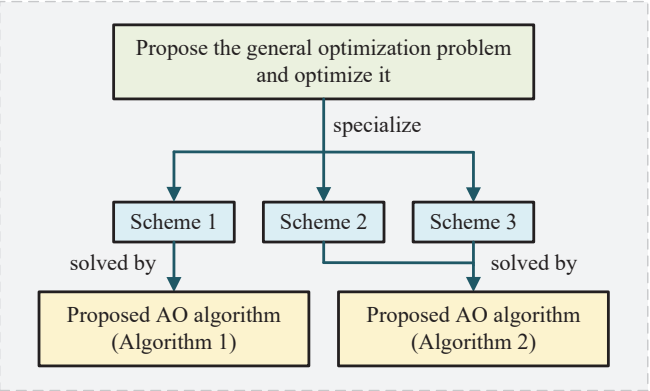


Fig. 2. Flow chart of the beamforming problem and solution.

III. Beamforming Problem and Solution

A. General Optimization Problem Formulation

In the covert communication system with a multi-position warden, since Willie's positions are different, we need to propose corresponding schemes according to different fixed positions of Willie to achieve the optimal covert communication. There are three scenes corresponding to three possible fixed positions of Willie: Willie is far from Alice, Willie is near Alice, and the Willie's distance from Alice is between near and far. We propose three corresponding schemes for the three scenes.

Therefore, we will firstly propose an optimization problem called the general optimization problem and optimize it, then specialize it to three specific schemes (three optimization problems). The flow chart of the beamforming problem and solution is shown in Fig. 2.

For the general optimization problem, we aim to maximize the covert communication rate under the transmitting power constraint, effective jamming power constraint, covert communication rate constraint, and covertness constraint. The general optimization problem can be expressed as follows:

$$\max_{\boldsymbol{w}_c,\boldsymbol{w}_p,\boldsymbol{\Phi}} \quad R_c \tag{18}$$

$$\text{s.t.} \quad |\boldsymbol{w}_c|^2 + |\boldsymbol{w}_p|^2 \leq P_{max}, \tag{18a}$$

$$P_w \leq P_{wu}, \tag{18b}$$

$$P_w \geq P_{wl}, \tag{18c}$$

$$R_c \geq R_{min}, \tag{18d}$$

$$\mathcal{D}(p_0(y_W)||p_1(y_W)) \leq 2\varepsilon^2, \tag{18e}$$

where (18a) is the transmitting power constraint and P_{max} stands for the maximum transmitting power threshold at Alice. (18b) is the effective jamming power upper limit constraint, (18c) denotes the effective jamming power lower limit constraint, P_{wu} and P_{wl} stand for upper and lower limits of jamming tolerance, respectively. (18d) represents the covert communication rate constraint and R_{min} is the minimum covert rate threshold. Moreover, (18e) denotes the covertness constraint.

$$|(\boldsymbol{h}_{IW}^H\boldsymbol{\Phi}\boldsymbol{H}_{AI}^H+\boldsymbol{h}_{AW}^H)\boldsymbol{w}_c|^2+|(\boldsymbol{h}_{IW}^H\boldsymbol{\Phi}\boldsymbol{H}_{AI}^H+\boldsymbol{h}_{AW}^H)\boldsymbol{w}_p|^2\le(\mathcal{R}-1)\sigma_w^2. \tag{21}$$

$$\begin{aligned}
\max_{\boldsymbol{w}_c,\boldsymbol{w}_p,\boldsymbol{\Phi}}\quad & \log_2\left(1+\frac{|(\boldsymbol{h}_{IB}^H\boldsymbol{\Phi}\boldsymbol{H}_{AI}^H+\boldsymbol{h}_{AB}^H)\boldsymbol{w}_c|^2}{|(\boldsymbol{h}_{IB}^H\boldsymbol{\Phi}\boldsymbol{H}_{AI}^H+\boldsymbol{h}_{AB}^H)\boldsymbol{w}_p|^2+\sigma_b^2}\right) && (22)\\
\text{s.t.}\quad & |\boldsymbol{w}_c|^2+|\boldsymbol{w}_p|^2\le P_{max}, && (22a)\\
& |(\boldsymbol{h}_{IW}^H\boldsymbol{\Phi}\boldsymbol{H}_{AI}^H+\boldsymbol{h}_{AW}^H)\boldsymbol{w}_p|^2\le 2|(\boldsymbol{h}_{IW}^H\boldsymbol{\Phi}\boldsymbol{H}_{AI}^H+\boldsymbol{h}_{AW}^H)\boldsymbol{w}_c|^2, && (22b)\\
& |(\boldsymbol{h}_{IW}^H\boldsymbol{\Phi}\boldsymbol{H}_{AI}^H+\boldsymbol{h}_{AW}^H)\boldsymbol{w}_p|^2\ge\frac{1}{2}|(\boldsymbol{h}_{IW}^H\boldsymbol{\Phi}\boldsymbol{H}_{AI}^H+\boldsymbol{h}_{AW}^H)\boldsymbol{w}_c|^2, && (22c)\\
& \log_2\left(1+\frac{|(\boldsymbol{h}_{IB}^H\boldsymbol{\Phi}\boldsymbol{H}_{AI}^H+\boldsymbol{h}_{AB}^H)\boldsymbol{w}_c|^2}{|(\boldsymbol{h}_{IB}^H\boldsymbol{\Phi}\boldsymbol{H}_{AI}^H+\boldsymbol{h}_{AB}^H)\boldsymbol{w}_p|^2+\sigma_b^2}\right)\ge R_{min}, && (22d)\\
& |(\boldsymbol{h}_{IW}^H\boldsymbol{\Phi}\boldsymbol{H}_{AI}^H+\boldsymbol{h}_{AW}^H)\boldsymbol{w}_c|^2+|(\boldsymbol{h}_{IW}^H\boldsymbol{\Phi}\boldsymbol{H}_{AI}^H+\boldsymbol{h}_{AW}^H)\boldsymbol{w}_p|^2\le(\mathcal{R}-1)\sigma_w^2. && (22e)
\end{aligned}$$

Using (15) in (18e), we can obtain

$$\ln\left(\frac{\sigma_1^2}{\sigma_0^2}\right)+\frac{\sigma_0^2}{\sigma_1^2}-1\le 2\varepsilon^2. \tag{19}$$

Let $\sigma_1^2/\sigma_0^2=a$ and $f(a)=\ln a+1/a-1$, one can obtain

$$f(a)\le 2\varepsilon^2. \tag{20}$$

Since $f(a)$ increases monotonically over $[1,+\infty)$, let $\mathcal{R}$ be the only solution of $f(a)=2\varepsilon^2$ on $[1,+\infty)$, so $a=\sigma_1^2/\sigma_0^2\le\mathcal{R}$, and $\sigma_1^2\le\mathcal{R}\sigma_0^2$. Thus, (18e) becomes (21) at the top of this page.

Property 2. *$f(a)$ increases monotonically over $[1,+\infty)$.*

Proof. The proof is given in Appendix B. □

Then, the general optimization problem can be rewritten as (22), which is given at the top of this page.

B. Active Beamforming Optimization at Alice

Using the given $\boldsymbol{\Phi}$, we let

$$\boldsymbol{h}_1=\boldsymbol{H}_{AI}^H\boldsymbol{\Phi}^H\boldsymbol{h}_{IB}+\boldsymbol{h}_{AB}, \tag{23}$$

where $\boldsymbol{h}_1\in\mathbb{C}^{M\times 1}$ denotes the combined channel vector between Alice and Bob. And we let $\boldsymbol{W}_c=\boldsymbol{w}_c\boldsymbol{w}_c^H$, where $\boldsymbol{W}_c\in\mathbb{C}^{M\times M}$ satisfies $\boldsymbol{W}_c\succeq 0$ and $rank(\boldsymbol{W}_c)=1$, so we can obtain

$$\begin{aligned}
&|(\boldsymbol{h}_{IB}^H\boldsymbol{\Phi}\boldsymbol{H}_{AI}+\boldsymbol{h}_{AB}^H)\boldsymbol{w}_c|^2\\
=&(\boldsymbol{h}_{IB}^H\boldsymbol{\Phi}\boldsymbol{H}_{AI}+\boldsymbol{h}_{AB}^H)\boldsymbol{w}_c\boldsymbol{w}_c^H(\boldsymbol{H}_{AI}^H\boldsymbol{\Phi}^H\boldsymbol{h}_{IB}+\boldsymbol{h}_{AB}) \quad (24)\\
=&\boldsymbol{h}_1^H\boldsymbol{W}_c\boldsymbol{h}_1.
\end{aligned}$$

Similarly, we let

$$\boldsymbol{h}_2=\boldsymbol{H}_{AI}^H\boldsymbol{\Phi}^H\boldsymbol{h}_{IW}+\boldsymbol{h}_{AW}, \tag{25}$$

where $\boldsymbol{h}_2\in\mathbb{C}^{M\times 1}$ represents the combined channel vector between Alice and Willie. And we let $\boldsymbol{W}_p=\boldsymbol{w}_p\boldsymbol{w}_p^H$, where $\boldsymbol{W}_p\in\mathbb{C}^{M\times M}$ satisfies $\boldsymbol{W}_p\succeq 0$ and $rank(\boldsymbol{W}_p)=1$, so we can obtain

$$\begin{aligned}
&|(\boldsymbol{h}_{IW}^H\boldsymbol{\Phi}\boldsymbol{H}_{AI}+\boldsymbol{h}_{AW}^H)\boldsymbol{w}_p|^2\\
=&(\boldsymbol{h}_{IW}^H\boldsymbol{\Phi}\boldsymbol{H}_{AI}+\boldsymbol{h}_{AW}^H)\boldsymbol{w}_p\boldsymbol{w}_p^H(\boldsymbol{H}_{AI}^H\boldsymbol{\Phi}^H\boldsymbol{h}_{IW}+\boldsymbol{h}_{AW}) \quad (26)\\
=&\boldsymbol{h}_2^H\boldsymbol{W}_p\boldsymbol{h}_2.
\end{aligned}$$

In addition, we can obtain the following expression:

$$\begin{aligned}
&|(\boldsymbol{h}_{IW}^H\boldsymbol{\Phi}\boldsymbol{H}_{AI}+\boldsymbol{h}_{AW}^H)\boldsymbol{w}_c|^2\\
=&(\boldsymbol{h}_{IW}^H\boldsymbol{\Phi}\boldsymbol{H}_{AI}+\boldsymbol{h}_{AW}^H)\boldsymbol{w}_c\boldsymbol{w}_c^H(\boldsymbol{H}_{AI}^H\boldsymbol{\Phi}^H\boldsymbol{h}_{IW}+\boldsymbol{h}_{AW}) \quad (27)\\
=&\boldsymbol{h}_2^H\boldsymbol{W}_c\boldsymbol{h}_2.
\end{aligned}$$

So (22) can be rewritten as

$$\begin{aligned}
\max_{\boldsymbol{W}_c,\boldsymbol{w}_p}\quad & \log_2\left(1+\frac{tr(\boldsymbol{h}_1\boldsymbol{h}_1^H\boldsymbol{W}_c)}{|\boldsymbol{h}_1^H\boldsymbol{w}_p|^2+\sigma_b^2}\right) && (28)\\
\text{s.t.}\quad & tr(\boldsymbol{W}_c)+tr(\boldsymbol{W}_p)\le P_{max}, && (28a)\\
& tr(\boldsymbol{h}_2\boldsymbol{h}_2^H\boldsymbol{W}_p)\le 2tr(\boldsymbol{h}_2\boldsymbol{h}_2^H\boldsymbol{W}_c), && (28b)\\
& tr(\boldsymbol{h}_2\boldsymbol{h}_2^H\boldsymbol{W}_p)\ge\frac{1}{2}tr(\boldsymbol{h}_2\boldsymbol{h}_2^H\boldsymbol{W}_c), && (28c)\\
& \log_2\left(1+\frac{tr(\boldsymbol{h}_1\boldsymbol{h}_1^H\boldsymbol{W}_c)}{|\boldsymbol{h}_1^H\boldsymbol{w}_p|^2+\sigma_b^2}\right)\ge R_{min}, && (28d)\\
& tr(\boldsymbol{h}_2\boldsymbol{h}_2^H\boldsymbol{W}_c)+tr(\boldsymbol{h}_2\boldsymbol{h}_2^H\boldsymbol{W}_p)\le(\mathcal{R}-1)\sigma_w^2, && (28e)\\
& \boldsymbol{W}_c\succeq 0,\boldsymbol{W}_p\succeq 0, && (28f)\\
& rank(\boldsymbol{W}_c)=1,rank(\boldsymbol{W}_p)=1. && (28g)
\end{aligned}$$

The fraction in the objective function of (28) is difficult to solve. Since $|\boldsymbol{h}_1^H\boldsymbol{w}_p|^2$ is the jamming signal term, the objective function is maximum when it is minimum, so we let $\boldsymbol{h}_1^H\boldsymbol{w}_p=0$, which is $\boldsymbol{w}_p\in null(\boldsymbol{h}_1^H)$, so (28) can be rewritten as

$$\begin{aligned}
\max_{\boldsymbol{W}_c,\boldsymbol{w}_p}\quad & \log_2\left(1+\frac{tr(\boldsymbol{h}_1\boldsymbol{h}_1^H\boldsymbol{W}_c)}{\sigma_b^2}\right) && (29)\\
\text{s.t.}\quad & \log_2\left(1+\frac{tr(\boldsymbol{h}_1\boldsymbol{h}_1^H\boldsymbol{W}_c)}{\sigma_b^2}\right)\ge R_{min}, && (29a)\\
& \boldsymbol{h}_1^H\boldsymbol{w}_p=0, && (29b)\\
& (28a),(28b),(28c),(28e),(28f),(28g), && (29c)
\end{aligned}$$

then we can rewrite (29) as follows:

$$\max_{\boldsymbol{W}_c,\boldsymbol{W}_{pn}} \quad \log_2\left(1+\frac{tr(\boldsymbol{h}_1\boldsymbol{h}_1^H\boldsymbol{W}_c)}{\sigma_b^2}\right) \tag{30}$$

$$\text{s.t.} \quad tr(\boldsymbol{W}_c)+tr(\boldsymbol{Q}^H\boldsymbol{Q}\boldsymbol{W}_{pn}) \le P_{max}, \tag{30a}$$

$$tr(\boldsymbol{Q}^H\boldsymbol{h}_2\boldsymbol{h}_2^H\boldsymbol{Q}\boldsymbol{W}_{pn}) \le 2tr(\boldsymbol{h}_2\boldsymbol{h}_2^H\boldsymbol{W}_c), \tag{30b}$$

$$tr(\boldsymbol{Q}^H\boldsymbol{h}_2\boldsymbol{h}_2^H\boldsymbol{Q}\boldsymbol{W}_{pn}) \ge \frac{1}{2}tr(\boldsymbol{h}_2\boldsymbol{h}_2^H\boldsymbol{W}_c), \tag{30c}$$

$$(29a), \tag{30d}$$

$$tr(\boldsymbol{h}_2\boldsymbol{h}_2^H\boldsymbol{W}_c)+tr(\boldsymbol{Q}^H\boldsymbol{h}_2\boldsymbol{h}_2^H\boldsymbol{Q}\boldsymbol{W}_{pn}) \le (\mathcal{R}-1)\sigma_w^2, \tag{30e}$$

$$\boldsymbol{W}_c \succeq 0, \boldsymbol{W}_{pn} \succeq 0, \tag{30f}$$

$$rank(\boldsymbol{W}_c)=1, rank(\boldsymbol{W}_{pn})=1. \tag{30g}$$

Property 3. (29) *can be rewritten as* (30)*.*

Proof. The proof is given in Appendix C. □

From (30), only the rank-one constraint (30g) makes (30) non-convex, so we can remove the rank-one constraint (30g) and obtain a convex problem. If the obtained solution satisfies the rank-one constraint, the desired solution can be obtained by eigenvalue decomposition. If the rank-one constraint is not satisfied, the rank-one solution can be recovered by Gaussian randomization. However, this may cause significant performance loss or the obtained solution may not meet some constraints, and the computational complexity is high when the number of iterations is large. To avoid these problems, we use the method based on penalty term. First, we rewrite the rank-one constraint (30g) as follows

$$||\boldsymbol{W}_c||_* - ||\boldsymbol{W}_c||_2 = 0, \tag{31}$$

$$||\boldsymbol{W}_{pn}||_* - ||\boldsymbol{W}_{pn}||_2 = 0, \tag{32}$$

where $||\cdot||_*$ and $||\cdot||_2$ represent the nuclear norm and the spectral norm, respectively. At the same time, $||\cdot||_*$ is the sum of all singular values of the matrix and $||\cdot||_2$ is the largest singular value in the matrix. In the case of $\boldsymbol{W}_{pn}$, because the rank of a matrix is equal to the number of its non-zero singular values, so when $rank(\boldsymbol{W}_{pn})=1$, we can obtain

$$||\boldsymbol{W}_{pn}||_* = ||\boldsymbol{W}_{pn}||_2, \tag{33}$$

and we can know that (32) holds. Otherwise, we have

$$||\boldsymbol{W}_{pn}||_* > ||\boldsymbol{W}_{pn}||_2, \tag{34}$$

which is

$$||\boldsymbol{W}_{pn}||_* - ||\boldsymbol{W}_{pn}||_2 > 0. \tag{35}$$

In a maximization problem, we can introduce the penalty term into the objective function to obtain the high-performance rank-one solution. In this paper, the penalty term of $\boldsymbol{W}_{pn}$ can be given as:

$$-\frac{1}{\eta}(||\boldsymbol{W}_{pn}||_* - ||\boldsymbol{W}_{pn}||_2), \tag{36}$$

where η denotes the penalty factor and $\eta>0$. When $\eta\to 0$, $\frac{1}{\eta}\to+\infty$, so when $\boldsymbol{W}_{pn}$ is not a rank-one matrix and $\eta\to 0$, the penalty term is infinitesimal, which leads to an infinitesimal value of the objective function introducing the penalty term, so a rank-one solution satisfying (32) can be obtained. In addition, $-||\boldsymbol{W}_{pn}||_2$ is non-convex, so we can obtain the upper bound of it by using the first-order Taylor expansion of point $\boldsymbol{W}_{pn}^n$ as follows [38]:

$$\begin{aligned}&-||\boldsymbol{W}_{pn}||_2 \le \widetilde{\boldsymbol{W}}_{pn}^n \\ &= -||\boldsymbol{W}_{pn}^n||_2 - tr[\boldsymbol{\lambda}_{max}^n(\boldsymbol{\lambda}_{max}^n)^H(\boldsymbol{W}_{pn}-\boldsymbol{W}_{pn}^n)],\end{aligned} \tag{37}$$

where $\boldsymbol{\lambda}_{max}^n$ stands for the eigenvector corresponding to the largest eigenvalue of $\boldsymbol{W}_{pn}^n$. Similarly, $\boldsymbol{W}_c$ also satisfies the above analysis. (30) can be rewritten as follows:

$$\max_{\boldsymbol{W}_c,\boldsymbol{W}_{pn}} \quad f_1(\boldsymbol{W}_c,\boldsymbol{W}_{pn}) \tag{38}$$

$$\text{s.t.} \quad (30a),(30b),(30c),(30d),(30e),(30f), \tag{38a}$$

where

$$\begin{aligned}&f_1(\boldsymbol{W}_c,\boldsymbol{W}_{pn}) \\ &=\log_2\left(1+\frac{tr(\boldsymbol{h}_1\boldsymbol{h}_1^H\boldsymbol{W}_c)}{\sigma_b^2}\right)-\frac{1}{\eta}\left(||\boldsymbol{W}_c||_*+\widetilde{\boldsymbol{W}}_c^n\right) \\ &-\frac{1}{\eta}\left(||\boldsymbol{W}_{pn}||_*+\widetilde{\boldsymbol{W}}_{pn}^n\right).\end{aligned} \tag{39}$$

It is obvious that (38) is a standard convex problem, which can be effectively solved with the CVX toolbox. After solving (38), the optimal solutions of $\boldsymbol{W}_c$ and $\boldsymbol{W}_{pn}$ can be obtained. Then, through eigenvalue decomposition and $\boldsymbol{w}_p=\boldsymbol{Q}\boldsymbol{w}_{pn}$, the optimal solutions of $\boldsymbol{w}_c$ and $\boldsymbol{w}_p$ can finally be obtained. They are denoted as $\boldsymbol{w}_c^o$ and $\boldsymbol{w}_p^o$.

C. Passive Beamforming Optimization at IRS

Next, given $\boldsymbol{w}_c^o$ and $\boldsymbol{w}_p^o$, we optimize $\boldsymbol{\Phi}$. First, we let $\boldsymbol{b}=(e^{j\phi_1},e^{j\phi_2},...,e^{j\phi_N})^H$ with $\boldsymbol{b}\in\mathbb{C}^{N\times 1}$, so we can obtain

$$\boldsymbol{\Phi}=diag(e^{j\phi_1},e^{j\phi_2},...,e^{j\phi_N})=diag(\boldsymbol{b}^H). \tag{40}$$

And we let $\boldsymbol{W}_{ab}^o=\boldsymbol{w}_{ab}^o(\boldsymbol{w}_{ab}^o)^H$, where $\boldsymbol{W}_{ab}^o$ satisfies $\boldsymbol{W}_{ab}^o\succeq 0$, $rank(\boldsymbol{W}_{ab}^o)=1$, and $\{ab\}\in\{c,p\}$. Thus, we can obtain (41), which is given at the top of the next page. where

$$diag(\boldsymbol{h}_{IB}^H)\boldsymbol{H}_{AI}=\boldsymbol{I}\in\mathbb{C}^{N\times M}, \tag{42}$$

and we let

$$\overline{\boldsymbol{b}}=\begin{bmatrix}\boldsymbol{b}\\1\end{bmatrix}, \hat{\boldsymbol{V}}_{ab}=\begin{bmatrix}\boldsymbol{I}\boldsymbol{W}_{ab}^o\boldsymbol{I}^H & \boldsymbol{I}\boldsymbol{W}_{ab}^o\boldsymbol{h}_{AB}\\ \boldsymbol{h}_{AB}^H\boldsymbol{W}_{ab}^o\boldsymbol{I}^H & 0\end{bmatrix}, \tag{43}$$

where $\overline{\boldsymbol{b}}\in\mathbb{C}^{N+1\times 1}$ and $\hat{\boldsymbol{V}}_{ab}\in\mathbb{C}^{N+1\times N+1}$, so we can obtain (44), which is given at the top of the next page.

Similarly, we let

$$\boldsymbol{I}_w=diag(\boldsymbol{h}_{IW}^H)\boldsymbol{H}_{AI}\in\mathbb{C}^{N\times M}, \tag{45}$$

and we have

$$\begin{aligned}&|(\boldsymbol{h}_{IW}^H\boldsymbol{\Phi}\boldsymbol{H}_{AI}+\boldsymbol{h}_{AW}^H)\boldsymbol{w}_{ab}^o|^2 \\ =&\overline{\boldsymbol{b}}^H\hat{\boldsymbol{V}}_{ab}^w\overline{\boldsymbol{b}}+\boldsymbol{h}_{AW}^H\boldsymbol{W}_{ab}^o\boldsymbol{h}_{AW},\end{aligned} \tag{46}$$

where $\hat{\boldsymbol{V}}_{ab}^w\in\mathbb{C}^{N+1\times N+1}$ and

$$\hat{\boldsymbol{V}}_{ab}^w=\begin{bmatrix}\boldsymbol{I}_w\boldsymbol{W}_{ab}^o\boldsymbol{I}_w^H & \boldsymbol{I}_w\boldsymbol{W}_{ab}^o\boldsymbol{h}_{AW}\\ \boldsymbol{h}_{AW}^H\boldsymbol{W}_{ab}^o\boldsymbol{I}_w^H & 0\end{bmatrix}. \tag{47}$$

$$\begin{aligned}&|(\boldsymbol{h}_{IB}^H\boldsymbol{\Phi}\boldsymbol{H}_{AI}+\boldsymbol{h}_{AB}^H)\boldsymbol{w}_{ab}^o|^2\\=&(\boldsymbol{b}^H diag(\boldsymbol{h}_{IB}^H)\boldsymbol{H}_{AI}+\boldsymbol{h}_{AB}^H)\boldsymbol{W}_{ab}^o(\boldsymbol{H}_{AI}^H diag(\boldsymbol{h}_{IB})\boldsymbol{b}+\boldsymbol{h}_{AB})\\=&(\boldsymbol{b}^H\boldsymbol{I}+\boldsymbol{h}_{AB}^H)\boldsymbol{W}_{ab}^o(\boldsymbol{I}^H\boldsymbol{b}+\boldsymbol{h}_{AB})\\=&\boldsymbol{b}^H\boldsymbol{I}\boldsymbol{W}_{ab}^o\boldsymbol{I}^H\boldsymbol{b}+\boldsymbol{b}^H\boldsymbol{I}\boldsymbol{W}_{ab}^o\boldsymbol{h}_{AB}+\boldsymbol{h}_{AB}^H\boldsymbol{W}_{ab}^o\boldsymbol{I}^H\boldsymbol{b}+\boldsymbol{h}_{AB}^H\boldsymbol{W}_{ab}^o\boldsymbol{h}_{AB},\end{aligned}\tag{41}$$

$$\begin{aligned}&\overline{\boldsymbol{b}}^H\hat{\boldsymbol{V}}_{ab}\overline{\boldsymbol{b}}+\boldsymbol{h}_{AB}^H\boldsymbol{W}_{ab}^o\boldsymbol{h}_{AB}\\=&(\boldsymbol{b}^H\boldsymbol{I}\boldsymbol{W}_{ab}^o\boldsymbol{I}^H+\boldsymbol{h}_{AB}^H\boldsymbol{W}_{ab}^o\boldsymbol{I}^H)\boldsymbol{b}+\boldsymbol{b}^H\boldsymbol{I}\boldsymbol{W}_{ab}^o\boldsymbol{h}_{AB}+\boldsymbol{h}_{AB}^H\boldsymbol{W}_{ab}^o\boldsymbol{h}_{AB}\\=&|(\boldsymbol{h}_{IB}^H\boldsymbol{\Phi}\boldsymbol{H}_{AI}+\boldsymbol{h}_{AB}^H)\boldsymbol{w}_{ab}^o|^2.\end{aligned}\tag{44}$$

$$\max_{\overline{\boldsymbol{B}}}\quad \log_2\left(1+\frac{tr(\hat{\boldsymbol{V}}_c\overline{\boldsymbol{B}})+\boldsymbol{h}_{AB}^H\boldsymbol{W}_c^o\boldsymbol{h}_{AB}}{tr(\hat{\boldsymbol{V}}_p\overline{\boldsymbol{B}})+\boldsymbol{h}_{AB}^H\boldsymbol{W}_p^o\boldsymbol{h}_{AB}+\sigma_b^2}\right)\tag{48}$$

$$\text{s.t.}\quad tr(\hat{\boldsymbol{V}}_p^w\overline{\boldsymbol{B}})+\boldsymbol{h}_{AW}^H\boldsymbol{W}_p^o\boldsymbol{h}_{AW}\le 2(tr(\hat{\boldsymbol{V}}_c^w\overline{\boldsymbol{B}})+\boldsymbol{h}_{AW}^H\boldsymbol{W}_c^o\boldsymbol{h}_{AW}),\tag{48a}$$

$$tr(\hat{\boldsymbol{V}}_p^w\overline{\boldsymbol{B}})+\boldsymbol{h}_{AW}^H\boldsymbol{W}_p^o\boldsymbol{h}_{AW}\ge\frac{1}{2}(tr(\hat{\boldsymbol{V}}_c^w\overline{\boldsymbol{B}})+\boldsymbol{h}_{AW}^H\boldsymbol{W}_c^o\boldsymbol{h}_{AW}),\tag{48b}$$

$$\log_2\left(1+\frac{tr(\hat{\boldsymbol{V}}_c\overline{\boldsymbol{B}})+\boldsymbol{h}_{AB}^H\boldsymbol{W}_c^o\boldsymbol{h}_{AB}}{tr(\hat{\boldsymbol{V}}_p\overline{\boldsymbol{B}})+\boldsymbol{h}_{AB}^H\boldsymbol{W}_p^o\boldsymbol{h}_{AB}+\sigma_b^2}\right)\ge R_{min},\tag{48c}$$

$$tr(\hat{\boldsymbol{V}}_c^w\overline{\boldsymbol{B}})+\boldsymbol{h}_{AW}^H\boldsymbol{W}_c^o\boldsymbol{h}_{AW}+tr(\hat{\boldsymbol{V}}_p^w\overline{\boldsymbol{B}})+\boldsymbol{h}_{AW}^H\boldsymbol{W}_p^o\boldsymbol{h}_{AW}\le(\mathcal{R}-1)\sigma_w^2,\tag{48d}$$

$$\overline{\boldsymbol{B}}\succeq 0,\tag{48e}$$

$$rank(\overline{\boldsymbol{B}})=1,\tag{48f}$$

$$\overline{\boldsymbol{B}}_{n,n}=1,\forall n\in\{1,2,..,N+1\}.\tag{48g}$$

$$\log_2(1+tr(\hat{\boldsymbol{V}}_c\overline{\boldsymbol{B}})+\boldsymbol{h}_{AB}^H\boldsymbol{W}_c^o\boldsymbol{h}_{AB}-\beta_k'(tr(\hat{\boldsymbol{V}}_p)+\boldsymbol{h}_{AB}^H\boldsymbol{W}_p^o\boldsymbol{h}_{AB}+\sigma_b^2)),\tag{49}$$

In addition, we let $\overline{\boldsymbol{b}}\,\overline{\boldsymbol{b}}^H=\overline{\boldsymbol{B}}\in\mathbb{C}^{N+1\times N+1}$, where $\overline{\boldsymbol{B}}$ satisfies $\overline{\boldsymbol{B}}\succeq 0$ and $rank(\overline{\boldsymbol{B}})=1$. Thus, we can rewrite (22) as (48), which is given at the top of this page.

It is clear that (48) is a single-ratio fractional programming, so we use Dinkelbach's transform to rewrite the objective function of (48) as (49), which is given at the top of this page. where β_k' stands for the introduced auxiliary variable and it follows

$$(\beta_k')_{m+1}=\frac{[tr(\hat{\boldsymbol{V}}_c\overline{\boldsymbol{B}})+\boldsymbol{h}_{AB}^H\boldsymbol{W}_c^o\boldsymbol{h}_{AB}]_m}{[tr(\hat{\boldsymbol{V}}_p\overline{\boldsymbol{B}})+\boldsymbol{h}_{AB}^H\boldsymbol{W}_p^o\boldsymbol{h}_{AB}+\sigma_b^2]_m},\tag{50}$$

where m represents the index of the iteration. When m reaches a certain number of the iteration or the modulus of the difference between two iterations of β_k' is less than a certain threshold, the optimal solution of $\overline{\boldsymbol{B}}$ can be obtained. Similar to the active beamforming vector optimization at Alice, we rewrite (48) using the penalty-based method as follows:

$$\max_{\overline{\boldsymbol{B}}}\quad f_2(\overline{\boldsymbol{B}})\tag{51}$$

$$\text{s.t.}\quad (48a),(48b),(53),(48d),(48e),(48g),\tag{51a}$$

where $f_2(\overline{\boldsymbol{B}})$ satisfies (52), which is given at the top of the next page. And (53) is expressed as follows:

$$\begin{aligned}&\log_2(1+tr(\hat{\boldsymbol{V}}_c\overline{\boldsymbol{B}})+\boldsymbol{h}_{AB}^H\boldsymbol{W}_c^o\boldsymbol{h}_{AB}-\\&\beta_k'(tr(\hat{\boldsymbol{V}}_p)+\boldsymbol{h}_{AB}^H\boldsymbol{W}_p^o\boldsymbol{h}_{AB}+\sigma_b^2))\ge R_{min}.\end{aligned}\tag{53}$$

We can know that (51) is a standard convex problem, which can be effectively solved with the CVX toolbox. After solving (51), the optimal solution of $\overline{\boldsymbol{B}}$ can be obtained, and through eigenvalue decomposition, (43), and (40), the optimal solution of $\boldsymbol{\Phi}$ can eventually be obtained.

Remark: The scenario corresponding to the proposed general optimization problem uses the ICAJ signal and IRS simultaneously, but this scenario is not within the scope of this paper's research, so it is not studied in detail in this paper. However, subsections A, B, and C of Section III and the idea of AO can provide a feasible method for the realization of this scenario, which is also one of the contributions of this paper.

D. Beamforming Optimizations in Different Covert Communications Scenes

From the previous two subsections, we can know that the beamforming optimization problems given in schemes 1, 2, and 3 below are standard convex problems, which can be effectively solved by CVX toolbox. Specifically as follows.

1) Scheme 1: When Willie is far from Alice, Alice transmits the communication signal to achieve covert communications. Respectively, the active beamforming and passive

$$f_2(\overline{\boldsymbol{B}}) = \log_2(1 + tr(\hat{\boldsymbol{V}}_c\overline{\boldsymbol{B}}) + \boldsymbol{h}_{AB}^H\boldsymbol{W}_c^o\boldsymbol{h}_{AB} - \beta_k'(tr(\hat{\boldsymbol{V}}_p\overline{\boldsymbol{B}}) + \boldsymbol{h}_{AB}^H\boldsymbol{W}_p^o\boldsymbol{h}_{AB} + \sigma_b^2)) - \frac{1}{\eta}\left(||\overline{\boldsymbol{B}}||_* + \widetilde{\overline{\boldsymbol{B}}^n}\right). \tag{52}$$

$$\max_{\overline{\boldsymbol{B}}} \quad \log_2\left(1 + \frac{tr(\hat{\boldsymbol{V}}_c\overline{\boldsymbol{B}}) + \boldsymbol{h}_{AB}^H\boldsymbol{W}_c^o\boldsymbol{h}_{AB}}{\sigma_b^2}\right) - \frac{1}{\eta}\left(||\overline{\boldsymbol{B}}||_* + \widetilde{\overline{\boldsymbol{B}}^n}\right) \tag{55}$$

$$\text{s.t.} \quad \log_2\left(1 + \frac{tr(\hat{\boldsymbol{V}}_c\overline{\boldsymbol{B}}) + \boldsymbol{h}_{AB}^H\boldsymbol{W}_c^o\boldsymbol{h}_{AB}}{\sigma_b^2}\right) \geq R_{min}, \tag{55a}$$

$$tr(\hat{\boldsymbol{V}}_c^w\overline{\boldsymbol{B}}) + \boldsymbol{h}_{AW}^H\boldsymbol{W}_c^o\boldsymbol{h}_{AW} \leq (\mathcal{R}-1)\sigma_w^2, \tag{55b}$$

$$(48e), (48g). \tag{55c}$$

beamforming optimization subproblems of this scheme are given as (54) in the following and (55) at the top of this page.

$$\max_{\boldsymbol{W}_c} \quad \log_2\left(1 + \frac{tr(\boldsymbol{h}_1\boldsymbol{h}_1^H\boldsymbol{W}_c)}{\sigma_b^2}\right) - \frac{1}{\eta}\left(||\boldsymbol{W}_c||_* + \widetilde{\boldsymbol{W}}_c^n\right) \tag{54}$$

$$\text{s.t.} \quad tr(\boldsymbol{W}_c) \leq P_{max}, \tag{54a}$$

$$(29a), \tag{54b}$$

$$tr(\boldsymbol{h}_2\boldsymbol{h}_2^H\boldsymbol{W}_c) \leq (\mathcal{R}-1)\sigma_w^2, \tag{54c}$$

$$\boldsymbol{W}_c \succeq 0. \tag{54d}$$

To solve the optimization problem of scheme 1, we propose an AO algorithm. By solving (54) and (55) alternately, we can finally obtain the optimal solutions of $\boldsymbol{w}_c$ and $\boldsymbol{\Phi}$ for the optimization problem of scheme 1. The procedure of the proposed AO algorithm for solving the optimization problem of scheme 1 is summarized in Algorithm 1.

Algorithm 1 The proposed AO algorithm for solving the optimization problem of scheme 1

Require: $\boldsymbol{\Phi}^{(0)}$: the initial IRS phase shift matrix; δ : the iterative threshold.

Ensure: The optimal solution of the beamforming vector for transmitting communication data symbol and the phase shift matrix of the IRS.

1: Let $l = 0$;
2: **repeat**
3: Use $\boldsymbol{\Phi}^{(l)}$ to solve (54) to obtain the optimal solutions of $\boldsymbol{w}_c^{(l+1)}$;
4: Use $\boldsymbol{w}_c^{(l+1)}$ to solve (55) to obtain the optimal solution of $\boldsymbol{\Phi}^{(l+1)}$;
5: Let $l = l + 1$;
6: **until** l reaches the maximum number or the modulus of the difference between the objective function values of two adjacent iterations is less than δ.
7: **return** The optimal solution of the beamforming vector for transmitting communication data symbol $\boldsymbol{w}_c^{(l)}$ and the phase shift matrix of the IRS $\boldsymbol{\Phi}^{(l)}$.

2) Scheme 2: When Willie is near Alice, Alice can transmit the ICAJ signal to maximize the covert communication rate while ensuring that the effective jamming power is sufficient. At the same time, the covertness constraint is also satisfied. In this way, we can realize covert communications while protecting the content of communication information, which improves the security of covert communications. The active beamforming optimization problem of this scheme is the problem obtained by removing IRS-related terms in $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$ from problem (38).

3) Scheme 3: When Willie's distance from Alice is between near and far, we propose the overall performance to balance the communication performance and jamming performance. In addition, we introduce a scale factor k to make the communication performance and jamming performance dimensions consistent, and the overall performance is calculated as the sum of the covert communication rate and the effective jamming power at Willie after scaling. And we define the overall performance as the objective function to be maximized. Additionally, we eliminate the lower bound of the jamming tolerance for better feasibility. We firstly give an optimization problem as follows:

$$\max_{\boldsymbol{W}_c, \boldsymbol{W}_{pn}} \quad f_1(\boldsymbol{W}_c, \boldsymbol{W}_{pn}) + k \cdot tr(\boldsymbol{Q}^H\boldsymbol{h}_2\boldsymbol{h}_2^H\boldsymbol{Q}\boldsymbol{W}_{pn}) \tag{56}$$

$$\text{s.t.} \quad (30a), (30b), (30d), (30e), (30f), \tag{56a}$$

so the optimization problem for this scheme's active beamforming is the problem obtained by removing IRS-related terms in $\boldsymbol{h}_1$ and $\boldsymbol{h}_2$ from problem (56). To solve the optimization problems of schemes 2 and 3, we also propose an AO algorithm. By solving the optimization problems of schemes 2 and 3, we can finally obtain the optimal solutions of $\boldsymbol{w}_c$ and $\boldsymbol{w}_p$ for the optimization problems of schemes 2 and 3. The procedure of the proposed AO algorithm for solving the optimization problems of schemes 2 and 3 is summarized in Algorithm 2.

IV. Simulation Results and Analysis

In this section, we conduct MATLAB simulations to evaluate the performances of the three proposed schemes for the covert communication system with a multi-position warden. For scheme 1, IRS and Willie are located at (50,30) m and (80,-10) m, respectively. For scheme 2, Willie is located at (10,-10) m. For scheme 3, Willie has 6 positions, which is located at (20,-10) m, (30,-10) m, (40,-10) m, (50,-10) m, (60,-10) m, and (70,-10) m, respectively. In addition, the common

Algorithm 2 The proposed AO algorithm for solving the optimization problems of schemes 2 and 3

Require: δ : the iterative threshold; Pb : the optimization problem.

Ensure: The optimal solution of the beamforming vector for transmitting communication data symbol and the beamforming vector for transmitting deception jamming signal.

1: Let $l = 0$;
2: **repeat**
3: Solve Pb to obtain the optimal solutions of $\boldsymbol{w}_c^{(l+1)}$ and $\boldsymbol{w}_p^{(l+1)}$;
4: Let $l = l + 1$;
5: **until** l reaches the maximum number or the modulus of the difference between the objective function values of two adjacent iterations is less than δ.
6: **return** The optimal solution of the beamforming vector for transmitting communication data symbol $\boldsymbol{w}_c^{(l)}$ and the beamforming vector for transmitting deception jamming signal $\boldsymbol{w}_p^{(l)}$.

TABLE I
COMMON PARAMETER SETTINGS FOR SCHEMES 1, 2, AND 3

Parameters	Values
Position of Alice	(0,0) m
Position of Bob	(100,0) m
Reference distance	$d_0 = 1$ m
Path loss at the reference distance	$Q_0 = -30$ dB
Path loss exponent corresponding to d_{AI}	$\alpha_{AI} = 2.4$
Path loss exponent corresponding to d_{AW}	$\alpha_{AW} = 4.2$
Path loss exponent corresponding to d_{AB}	$\alpha_{AB} = 4.2$
Path loss exponent corresponding to d_{IB}	$\alpha_{IB} = 3$
Path loss exponent corresponding to d_{IW}	$\alpha_{IW} = 3$
Penalty factor	$\eta = 0.1$
Variance of the noise at Bob	$\sigma_b^2 = -90$ dBm
Variance of the noise at Willie	$\sigma_w^2 = -90$ dBm
Minimum covert rate threshold	$R_{min} = 0.5$ bps/Hz
Required covertness level	$\varepsilon = 0.001$
Scale factor	$k = 10^{12}$
Iterative threshold	$\delta = 10^{-4}$

parameter settings for schemes 1, 2, and 3 are listed in Table I. Since the channel vector between Alice and Bob is related to the distance between them, if the positions of Alice and Bob are not fixed, the results of the optimization problem for each time slot need to be discussed in each time slots. To simplify the optimization problem, this paper does not discuss the case of multiple time slots. Therefore, we assume that the positions of Alice and Bob are fixed. If the positions of Alice and Bob are varying, it is necessary to assume that the different positions of Alice and Bob in different time slots are fixed, then the proposed method can be used in each time slot.

All benchmark schemes in this paper are expressed as follows.

1) ***Scheme 1 benchmark***: Similar to scheme 1 but the IRS assistance is removed.

2) ***Scheme 2 benchmark***: Similar to scheme 2, the separate design is adopted, that is, the number of antennas at Alice for communication and jamming is $M/2$, respectively. And the transmitting power constraint becomes two constraints: the transmitting power of communication and jamming at Alice should be less than or equal to $P_{max}/2$, respectively.

3) ***Scheme 3 benchmark***: Similar to scheme 3, the separate design is adopted.

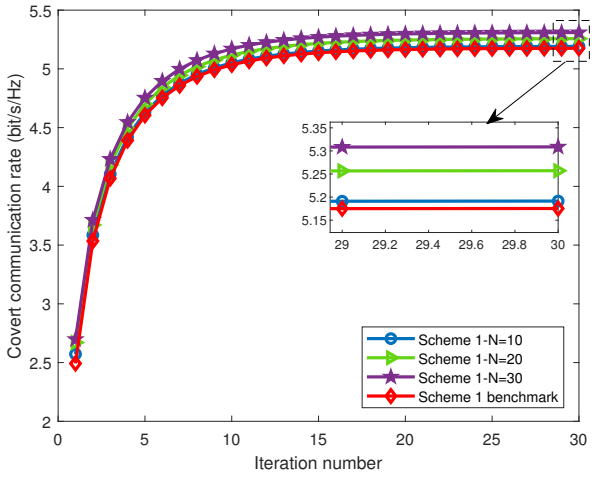


Fig. 3. Convergences of scheme 1 under different N and scheme 1 benchmark.

A. Covert Communication Performance in Scene 1

In Fig. 3, we show the convergence of scheme 1 and scheme 1 benchmark. From Fig. 3, we can see that scheme 1 can fully converge when the number of iterations is about 25, which verifies the effectiveness of scheme 1. In addition, the covert communication rate of scheme 1 is higher than that of scheme 1 benchmark, indicating that the covert communication system can have higher covert communication rate with IRS assistance than without IRS assistance when Willie is far from Alice. Moreover, with the increase of N, the covert communication rate of scheme 1 also increases. This is because the increase in N improves the IRS's ability to use space resources to be able to reflect the signal more focused to Bob. However, the increase of N will lead to the increase of the cost of the system, so the number of N can not be too large in practical applications.

In Fig. 4, we demonstrate the convergence of scheme 1 and scheme 1 benchmark under different M. From Fig. 4, it can be seen that as the number of Alice's antennas increases, the covert communication rates of scheme 1 and scheme 1 benchmark both increase, indicating that the increase of M can improve the performance of covert communications when Willie is far from Alice. This is because more M improves the utilization of space and increases the gain of transmitting beamforming, thus improving the performance of covert communications. In addition, under the same M, the covert communication rate of scheme 1 is always better than that of scheme 1 benchmark.

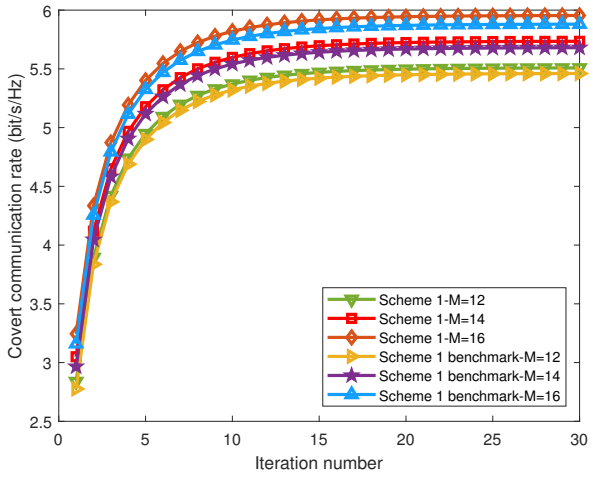


Fig. 4. Convergences of scheme 1 and scheme 1 benchmark under different M.

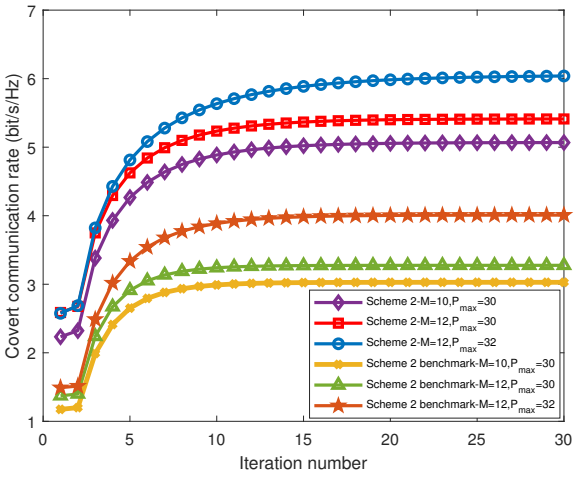


Fig. 6. Convergences of scheme 2 and scheme 2 benchmark under different M and P_{max}.

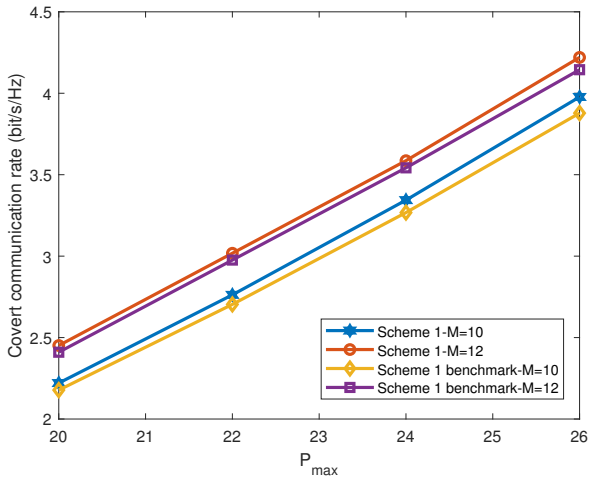


Fig. 5. Covert communication rates of scheme 1 and scheme 1 benchmark under different M and different P_{max}.

In Fig. 5, we show the covert communication rate of scheme 1 and scheme 1 benchmark under different M and different P_{max}. From Fig. 5, we can see that with the increase of P_{max}, the covert communication rates of both scheme 1 and scheme 1 benchmark increase. The reason is that when Willie is far from Alice, increasing P_{max} can raise the upper limit of Alice's transmitting power, thereby increasing Alice's transmitting power while ensuring that the transmitting signal is not detected by Willie, and thus improving the covert communication rate. At the same time, the covert communication rate of scheme 1 is always better than that of scheme 1 benchmark, which indicates the effectiveness of scheme 1.

B. Covert Communication Performance in Scene 2

In Fig. 6, we demonstrate the convergences of scheme 2 and scheme 2 benchmark under different M and P_{max}. From Fig. 6, we can know that when $P_{max} = 30$ dBm, scheme 2 can converge completely when the number of iterations is about 25, demonstrating the effectiveness of scheme 2. When $P_{max} = 32$ dBm, scheme 2 can basically converge when the number of iterations is about 30. This is because the increase of P_{max} leads to a larger optimization space, so the convergence of scheme 2 will be a little bit slower. In addition, the covert communication rate of scheme 2 is always significantly higher than that of scheme 2 benchmark, so the ICAJ signal has better covert communication performance than that of the separate design signal when Willie is near Alice, which proves the superiority of the ICAJ design.

In Fig. 7, we study the covert communication rates of scheme 2 and scheme 2 benchmark under different M when $P_{max} = 30$ dBm and $P_{max} = 32$ dBm. From Fig. 7, it can be seen that with the increase of M, the covert communication rate of scheme 2 and scheme 2 benchmark also increases, indicating that when Willie is near Alice, increasing M can enhance the covert communication performance, and the reason is the same as the explanation of Fig. 4. Additionally, under the same P_{max}, the covert communication performance of the ICAJ signal is better than that of the separate design signal, which shows the excellence of scheme 2.

In Fig. 8, the covert communication rates of scheme 2 and scheme 2 benchmark under different P_{max} when $M = 10$ and $M = 12$ is shown. From Fig. 8, we can see that when $M = 10$, with the increase of P_{max}, the covert communication rates of both scheme 2 and scheme 2 benchmark increase, indicating that when Willie is near Alice, increasing P_{max} can improve the covert communication performance. And when $M = 12$, scheme 2 benchmark shows the same trend as its $M = 10$ case while the performance of scheme 2 rises first and then remains stable. The reason of the above increase in performance as P_{max} increases is the same as the explanation of Fig. 5, and the reason for the above stability is that for scheme 2, when $M = 12$ and after P_{max} reaches 36 dBm, the increase of P_{max} can only increase the upper limit of Alice's transmitting power, but Alice cannot increase the actual transmitting power due to the limitation of the covertness constraint, so the performance of scheme 2 keeps stable. Moreover, the covert communication

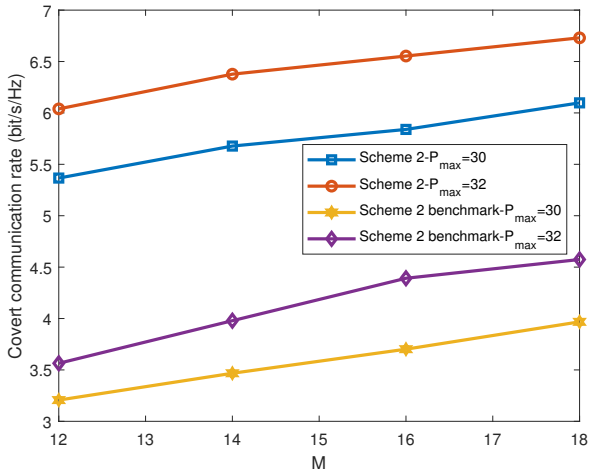


Fig. 7. Covert communication rates of scheme 2 and scheme 2 benchmark under different M when $P_{max} = 30$ dBm and $P_{max} = 32$ dBm.

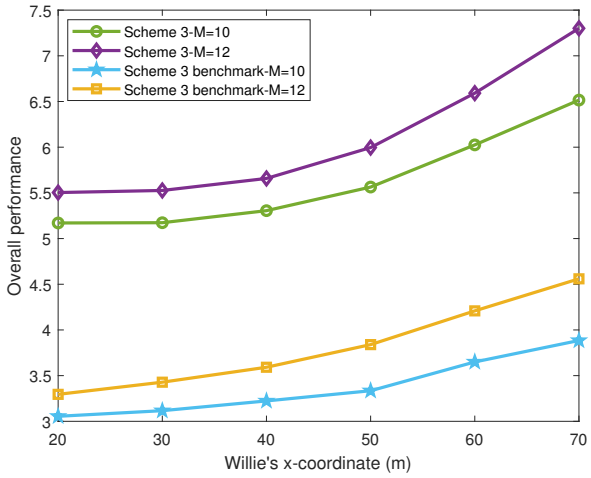


Fig. 9. Overall performances of scheme 3 and scheme 3 benchmark under different Willie's x-coordinates when $M = 10$ and $M = 12$.

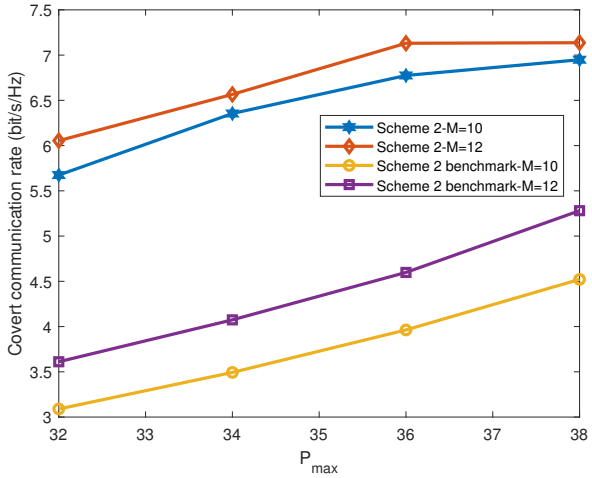


Fig. 8. Covert communication rates of scheme 2 and scheme 2 benchmark under different P_{max} when $M = 10$ and $M = 12$.

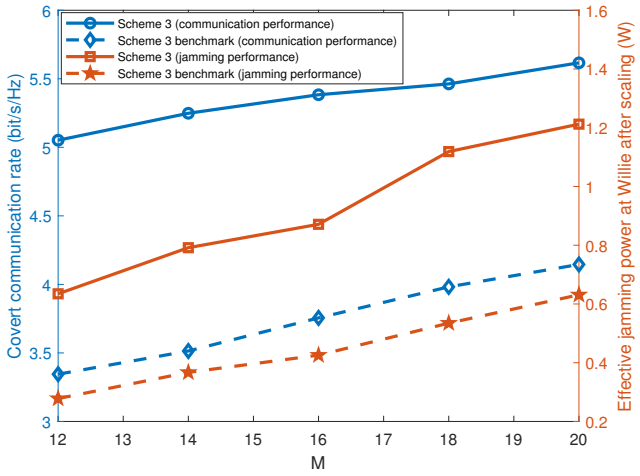


Fig. 10. Covert communication rates and the effective jamming power at Willie after scaling under different M of scheme 3 and scheme 3 benchmark.

performance of scheme 2 consistently surpasses that of scheme 2 benchmark, which proves that scheme 2 is superior to scheme 2 benchmark.

C. Covert Communication Performance in Scene 3

In Fig. 9, we analyze the relationship between Willie's x-coordinates and the overall performances of scheme 3 and scheme 3 benchmark under different M. From Fig. 9, it can be seen that the overall performance of both scheme 3 and scheme 3 benchmark is getting higher as Willie gets farther from Alice. This is because the farther Willie is from Alice, the looser the constraints on the beamforming of communication and jamming under the constraint of covertness, so the higher the transmitting power at Alice (when the transmitting power does not reach P_{max}) or more freedom in optimizing the beamforming of the communication and jamming, or both, thus finally leads to the improvement of the overall performance. Moreover, the overall performance of scheme 3 is always far better than that of scheme 3 benchmark when Willie's distance from Alice is between near and far, which reflects the effectiveness of scheme 3.

In Fig. 10, we study the impact of M on the communication performance and jamming performance of scheme 3 and scheme 3 benchmark. From Fig. 10, we can see that the communication and jamming performances of scheme 3 are always better than that of scheme 3 benchmark, and with the increase of M, the covert communication rate and the effective jamming power after scaling of both scheme 3 and scheme 3 benchmark increase, which indicates that when Willie's distance from Alice is between near and far, increasing M not only effectively improves communication performance, but also improves the jamming performance, so the security and the performance of the covert communication are enhanced. Therefore, scheme 3 has good validity.

In Fig. 11, the impact of P_{max} on the covert communication rate and the effective jamming power of scheme 3 and scheme

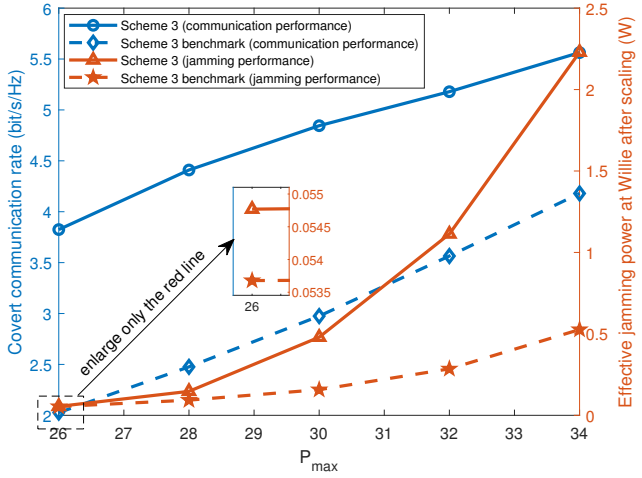


Fig. 11. Covert communication rates and the effective jamming power at Willie after scaling under different P_{max} of scheme 3 and scheme 3 benchmark.

3 benchmark is presented. From Fig. 11, we can see that the communication and jamming performances of scheme 3 are always greater than that of scheme 3 benchmark, and with the increase of P_{max}, the communication and jamming performances of both scheme 3 and scheme 3 benchmark increase, indicating that when Willie's distance from Alice is between near and far, increasing P_{max} is beneficial to both the communication and jamming performances. However, a higher P_{max} will bring more resource consumption, so the balance between the resource consumption and the performance improvement should be considered in practical applications.

V. Conclusion

This paper studies the beamforming optimization for covert communications with a multi-position warden. Firstly, a general optimization problem is proposed and optimized based on Dinkelbach's transform to maximize the covert communication rate under the constraints of transmitting power, effective jamming power, covert communication rate, and covertness. Then, based on the optimized general optimization problem, we propose three corresponding schemes for warden's three kinds of fixed positions, and we propose their own optimization problems. Scheme 1 is assisted by IRS to improve the covert communication performance. Schemes 2 and 3 use the ICAJ signal to enhance security for covert communications. Furthermore, three optimization problems corresponding to three schemes are solved using the proposed two AO algorithms. Simulation results demonstrate that all three proposed schemes have good effectiveness in their respective scenes, and their covert communication performances are better than that of their respective benchmark schemes.

Appendix A
Proof of Property 1

From [34], we know that the KL divergence can be written in the following form:

$$\begin{aligned}&\mathcal{D}(p_0(y_W)||p_1(y_W))\\&=\int p_0(y_w)\ln\left(\frac{p_0(y_w)}{p_1(y_w)}\right)dy_w\\&=\int p_0(y_w)\ln(p_0(y_w))dy_w-\int p_0(y_w)\ln(p_1(y_w))dy_w.\end{aligned}\tag{57}$$

First, we know $p_0(y_w)\sim\mathcal{CN}(0,\sigma_0^2)$, and we know that the variance of the likelihood function satisfies

$$Var=\int x^2f(x)dx-\mu^2,\tag{58}$$

where $f(x)$ stands for the likelihood function, μ denotes the mean of the likelihood function. In the meantime, since we have already obtained $p_0(y_W)$ as $p_0(y_W)=\frac{1}{\pi\sigma_0^2}\exp\left(-\frac{|y_W|^2}{\sigma_0^2}\right)$, so we can calculate the first term of (57) as follows:

$$\begin{aligned}&\int p_0(y_w)\ln(p_0(y_w))dy_w\\&=\int p_0(y_w)\ln\left(\frac{1}{\pi\sigma_0^2}\exp\left(-\frac{|y_w|^2}{\sigma_0^2}\right)\right)dy_w\\&=\int p_0(y_w)\left(\ln\left(\frac{1}{\pi\sigma_0^2}\right)+\ln\left(\exp\left(-\frac{|y_w|^2}{\sigma_0^2}\right)\right)\right)dy_w\\&=\ln\left(\frac{1}{\pi\sigma_0^2}\right)\int p_0(y_w)dy_w-\frac{1}{\sigma_0^2}\int y_w^2p_0(y_w)dy_w\\&=\ln\left(\frac{1}{\pi\sigma_0^2}\right)\cdot1-\frac{1}{\sigma_0^2}\cdot\sigma_0^2\\&=\ln\left(\frac{1}{\pi\sigma_0^2}\right)-1.\end{aligned}\tag{59}$$

In addition, we have already obtained $p_1(y_W)$ as $p_1(y_W)=\frac{1}{\pi\sigma_1^2}\exp\left(-\frac{|y_W|^2}{\sigma_1^2}\right)$. Therefore, we can obtain the second term of (57) as

$$\begin{aligned}&\int p_0(y_w)\ln p_1(y_w)dy_w\\&=\int p_0(y_w)\ln\left(\frac{1}{\pi\sigma_1^2}\exp\left(-\frac{|y_w|^2}{\sigma_1^2}\right)\right)dy_w\\&=\int p_0(y_w)\left(\ln\left(\frac{1}{\pi\sigma_1^2}\right)+\ln\left(\exp\left(-\frac{|y_w|^2}{\sigma_1^2}\right)\right)\right)dy_w\\&=\ln\left(\frac{1}{\pi\sigma_1^2}\right)\int p_0(y_w)dy_w-\frac{1}{\sigma_1^2}\int y_w^2p_0(y_w)dy_w\\&=\ln\left(\frac{1}{\pi\sigma_1^2}\right)-\frac{\sigma_0^2}{\sigma_1^2}.\end{aligned}\tag{60}$$

Substitute (59) and (60) into (57), the KL divergence can be given by

$$\begin{aligned}&\mathcal{D}(p_0(y_W)||p_1(y_W))\\&=\ln\left(\frac{1}{\pi\sigma_0^2}\right)-1-\ln\left(\frac{1}{\pi\sigma_1^2}\right)+\frac{\sigma_0^2}{\sigma_1^2}\\&=\ln\left(\frac{\sigma_1^2}{\sigma_0^2}\right)+\frac{\sigma_0^2}{\sigma_1^2}-1.\end{aligned}\tag{61}$$

Therefore, Property 1 is proved.

APPENDIX B
PROOF OF PROPERTY 2

From the definitions of σ_0^2 and σ_1^2, we can obtain

$$\sigma_0^2>0,\sigma_1^2>0,\tag{62}$$

and we know

$$a=\frac{\sigma_1^2}{\sigma_0^2},\tag{63}$$

so we can obtain $a>0$. Then, take the derivative of $f(a)$, we can know

$$f'(a)=\frac{1}{a}-\frac{1}{a^2}=\frac{a-1}{a^2}.\tag{64}$$

From (64), we know that when $a\in[1,+\infty)$, $f'(a)\geq 0$, so $f(a)$ increases monotonically over $[1,+\infty)$, Property 2 is proved.

APPENDIX C
PROOF OF PROPERTY 3

First, we set

$$null(\boldsymbol{h}_1^H)=\boldsymbol{Q}\in\mathbb{C}^{M\times M-1},\tag{65}$$

thus we can obtain

$$\boldsymbol{w}_p=\boldsymbol{Q}\boldsymbol{w}_{pn},\tag{66}$$

where $\boldsymbol{w}_{pn}\in\mathbb{C}^{M-1\times 1}$. In addition, we assume that

$$\boldsymbol{w}_{pn}\boldsymbol{w}_{pn}^H=\boldsymbol{W}_{pn}\in\mathbb{C}^{M-1\times M-1},\tag{67}$$

where $\boldsymbol{W}_{pn}$ satisfies $\boldsymbol{W}_{pn}\succeq 0$ and $rank(\boldsymbol{W}_{pn})=1$. Thus, (29) can be given as (30), Property 3 is proved.

REFERENCES

[1] M. Liu, Z. Zhang, Y. Chen, J. Ge and N. Zhao, "Adversarial Attack and Defense on Deep Learning for Air Transportation Communication Jamming," *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 1, pp. 973-986, Jan. 2024.

[2] M. V. Jamali and H. Mahdavifar, "Covert Millimeter-Wave Communication: Design Strategies and Performance Analysis," *IEEE Trans. Wireless Commun.*, vol. 21, no. 6, pp. 3691-3704, June 2022.

[3] D. Wang, P. Qi, Y. Zhao, C. Li, W. Wu and Z. Li, "Covert Wireless Communication With Noise Uncertainty in Space-Air-Ground Integrated Vehicular Networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 3, pp. 2784-2797, March 2022.

[4] Y. Qian, W. Li, Y. Lin, L. Shi, X. Zhou, J. Li and F. Shu, "Antenna Coding and Rate Optimization for Covert Wireless Communications," *IEEE Internet Things J.*, vol. 10, no. 3, pp. 2459-2472, Feb. 2023.

[5] L. Lv, Z. Li, H. Ding, N. Al-Dhahir and J. Chen, "Achieving Covert Wireless Communication With a Multi-Antenna Relay," *IEEE Trans. Inf. Forensics Secur.*, vol. 17, pp. 760-773, Feb. 2022.

[6] C. Wang, Z. Li and D. W. K. Ng, "Covert Rate Optimization of Millimeter Wave Full-Duplex Communications," *IEEE Trans. Wireless Commun.*, vol. 21, no. 5, pp. 2844-2861, May 2022.

[7] R. Ma, W. Yang, X. Guan, X. Lu, Y. Song and D. Chen, "Covert mmWave Communications With Finite Blocklength Against Spatially Random Wardens," *IEEE Internet Things J.*, vol. 11, no. 2, pp. 3402-3416, Jan. 2024.

[8] Y. Ding, H. Han, W. Lu, Y. Wang, N. Zhao, X. Wang and X. Yang, "DDQN-Based Trajectory and Resource Optimization for UAV-Aided MEC Secure Communications," *IEEE Trans. Veh. Technol.*, vol. 73, no. 4, pp. 6006-6011, April 2024.

[9] B. Che, C. Gao, R. Ma, X. Zheng and W. Yang, "Covert Wireless Communication in Multichannel Systems," *IEEE Wireless Commun. Lett.*, vol. 11, no. 9, pp. 1790-1794, Sept. 2022.

[10] M. Wang, Z. Xu, B. Xia, Y. Guo and Z. Chen, "DF Relay Assisted Covert Communications: Analysis and Optimization," *IEEE Trans. Veh. Technol.*, vol. 72, no. 3, pp. 4073-4078, March 2023.

[11] T. -X. Zheng, Z. Yang, C. Wang, Z. Li, J. Yuan and X. Guan, "Wireless Covert Communications Aided by Distributed Cooperative Jamming Over Slow Fading Channels," *IEEE Trans. Wireless Commun.*, vol. 20, no. 11, pp. 7026-7039, Nov. 2021.

[12] K. -H. Cho and S. -H. Lee, "Covert Communication Over Gaussian Multiple-Access Channels With Feedback," *IEEE Wireless Commun. Lett.*, vol. 11, no. 9, pp. 1985-1989, Sept. 2022.

[13] R. Xu, D. Guo, B. Zhang and G. Ding, "Finite Blocklength Covert Communications in Interweave Cognitive Radio Networks," *IEEE Commun. Lett.*, vol. 26, no. 9, pp. 1989-1993, Sept. 2022.

[14] W. Gao, Y. Chen, C. Han and Z. Chen, "Distance-Adaptive Absorption Peak Modulation (DA-APM) for Terahertz Covert Communications," *IEEE Trans. Wireless Commun.*, vol. 20, no. 3, pp. 2064-2077, March 2021.

[15] W. Ma, Z. Niu, W. Wang, S. He and T. Jiang, "Covert Communication With Uninformed Backscatters in Hybrid Active/Passive Wireless Networks: Modeling and Performance Analysis," *IEEE Trans. Commun.*, vol. 70, no. 4, pp. 2622-2634, April 2022.

[16] G. Yang, Y. Qian, K. Ren, Z. Mei, F. Shu, X. Zhou and W. Wu, "Covert Wireless Communications for Augmented Reality Systems With Dual Cooperative UAVs," *IEEE J. Sel. Top. Signal Process.*, vol. 17, no. 5, pp. 1119-1130, Sept. 2023.

[17] H. Rao, M. Wu, J. Wang, W. Tang, S. Xiao and S. Li, "D2D Covert Communications With Safety Area," *IEEE Syst. J.*, vol. 15, no. 2, pp. 2331-2341, June 2021.

[18] J. Zhang, M. Li, S. Yan, C. Liu, X. Chen, M. Zhao and P. Whiting, "Joint Beam Training and Data Transmission Design for Covert Millimeter-Wave Communication," *IEEE Trans. Inf. Forensics Secur.*, vol. 16, pp. 2232-2245, Jan. 2021.

[19] M. Lin and W. Wang, "Multi-Carrier Wireless Covert Communication With Delay Constraint," *IEEE Trans. Veh. Technol.*, vol. 72, no. 4, pp. 5350-5355, April 2023.

[20] W. He, J. Chen, G. Li, H. Wang, X. Chu, R. He, Y. Xu and Y. Jiao, "Optimal Transmission Probabilities of Information and Artificial Noise in Covert Communications," *IEEE Commun. Lett.*, vol. 26, no. 12, pp. 2865-2869, Dec. 2022.

[21] Y. Li, Y. Zhang, J. Wang, W. Xiang, S. Xiao, L. Chang and W. Tang, "Performance Analysis for Covert Communications Under Faster-Than-Nyquist Signaling," *IEEE Commun. Lett.*, vol. 26, no. 6, pp. 1240-1244, June 2022.

[22] J. Cui, S. Yan, J. Hu, N. Li, R. Chen and J. Li, "How Does Repetition Coding Enable Reliable and Covert Communications?," *IEEE Wireless Commun. Lett.*, vol. 10, no. 3, pp. 639-643, March 2021.

[23] L. Li, Z. Chen, R. Chen, L. Yang and S. Yan, "Covert Wireless Communication With Random Frequency Diverse Array," *IEEE Trans. Veh. Technol.*, vol. 73, no. 1, pp. 1473-1478, Jan. 2024.

[24] Z. Sun and Y. Jing, "On the Performance of Multi-Antenna IRS-Assisted NOMA Networks With Continuous and Discrete IRS Phase Shifting," *IEEE Trans. Wireless Commun.*, vol. 21, no. 5, pp. 3012-3023, May 2022.

[25] Z. Sun and Y. Jing, "On the Performance of Training-Based IRS-Assisted Communications Under Correlated Rayleigh Fading," *IEEE Trans. Commun.*, vol. 71, no. 5, pp. 3117-3131, May 2023.

[26] M. Fu, W. Mei and R. Zhang, "Multi-Active/Passive-IRS Enabled Wireless Information and Power Transfer: Active IRS Deployment and Performance Analysis," *IEEE Commun. Lett.*, vol. 27, no. 8, pp. 2217-2221, Aug. 2023.

[27] J. Xu, Z. Zhu, Z. Chu, H. Niu, P. Xiao and I. Lee, "Sum Secrecy Rate Maximization for IRS-Aided Multi-Cluster MIMO-NOMA Terahertz Systems," *IEEE Trans. Inf. Forensics Secur.*, vol. 18, pp. 4463-4474, July 2023.

[28] H. Wei and H. Zhang, "An Equivalent Model for Handover Probability Analysis of IRS-Aided Networks," *IEEE Trans. Veh. Technol.*, vol. 72, no. 10, pp. 13770-13774, Oct. 2023.

[29] Z. Yang, P. Yue, S. Wang, G. Pan and J. An, "Energy-Efficient Optimization for RIS-Aided MIMO Covert Communications," *IEEE Internet Things J.*, vol. 10, no. 21, pp. 18993-19003, Nov. 2023.

[30] C. Chen, M. Wang, B. Xia, Y. Guo and J. Wang, "Performance Analysis and Optimization of IRS-Aided Covert Communication With Hardware Impairments," *IEEE Trans. Veh. Technol.*, vol. 72, no. 4, pp. 5463-5467, April 2023.

[31] D. Song, Z. Yang, G. Pan, S. Wang and J. An, "RIS-Assisted Covert Transmission in Satellite–Terrestrial Communication Systems," *IEEE Internet Things J.*, vol. 10, no. 22, pp. 19415-19426, Nov. 2023.

[32] J. Zhao, G. Li, G. Chen, Q. Wu, X. Li and J. Zhang, "Covert Communication of IRS-Assisted Wireless Networks With Stochastic Geometry Approach," *IEEE Wireless Commun. Lett.*, vol. 13, no. 4, pp. 974-978, April 2024.

[33] Y. Wu, X. Chen, M. Liu, L. Xu, N. Zhao, X. Wang and D. W. Kwan Ng, "IRS-Assisted Covert Communication With Equal and Unequal Transmit Prior Probabilities," *IEEE Trans. Commun.*, vol. 72, no. 5, pp. 2897-2912, May 2024.

[34] M. Wang, Z. Xu, B. Xia and Y. Guo, "Active Intelligent Reflecting Surface Assisted Covert Communications," *IEEE Trans. Veh. Technol.*, vol. 72, no. 4, pp. 5401-5406, April 2023.

[35] H. Xiao, X. Hu, P. Mu, W. Wang, T. -X. Zheng, K. -K. Wong and K. Yang, "Simultaneously Transmitting and Reflecting RIS (STAR-RIS) Assisted Multi-Antenna Covert Communication: Analysis and Optimization," *IEEE Trans. Wireless Commun.*, vol. 23, no. 6, pp. 6438-6452, June 2024.

[36] C. Wang, X. Chen, J. An, Z. Xiong, C. Xing, N. Zhao and D. Niyato, "Covert Communication Assisted by UAV-IRS," *IEEE Trans. Commun.*, vol. 71, no. 1, pp. 357-369, Jan. 2023.

[37] J. Jiang, J. Jing, R. Ma, B. Che and W. Yang, "Covert Relay Communications With Finite Block-Length Against Combining Detection," *IEEE Wireless Commun. Lett.*, vol. 11, no. 11, pp. 2450-2454, Nov. 2022.

[38] Z. Wang, Y. Liu, X. Mu, Z. Ding and O. A. Dobre, "NOMA Empowered Integrated Sensing and Communication," *IEEE Commun. Lett.*, vol. 26, no. 3, pp. 677-681, March 2022.