Simultaneously Transmitting and Reflecting RIS (STAR-RIS) Assisted Multi-Antenna Covert Communication: Analysis and Optimization

Han Xiao, *Student Member, IEEE,* Xiaoyan Hu*, *Member, IEEE,*
Pengcheng Mu, *Member, IEEE,* Wenjie Wang, *Member, IEEE,* Tong-Xing Zheng, *Member, IEEE,*
Kai-Kit Wong, *Fellow, IEEE*, Kun Yang, *Fellow, IEEE*

***Abstract*—This paper investigates the multi-antenna covert communications assisted by a simultaneously transmitting and reflecting reconfigurable intelligent surface (STAR-RIS). In particular, to shelter the existence of covert communications between a multi-antenna transmitter and a single-antenna receiver from a warden, a friendly full-duplex receiver with two antennas is leveraged to make contributions where one antenna is responsible for receiving the transmitted signals and the other one transmits the jamming signals with a varying power to confuse the warden. Considering the worst case, the closed-form expression of the minimum detection error probability (DEP) at the warden is derived and utilized in a covert constraint to guarantee the system performance. Then, we formulate an optimization problem maximizing the covert rate of the system under the covertness constraint and quality of service (QoS) constraint with communication outage analysis. To jointly design the active and passive beamforming of the transmitter and STAR-RIS, an iterative algorithm based on semi-definite relaxation (SDR) method and Dinkelbach's algorithm is proposed to effectively solve the non-convex optimization problem. Simulation results show that the proposed STAR-RIS-assisted scheme highly outperforms the case with conventional RIS, which validates the effectiveness of the proposed algorithm as well as the superiority of STAR-RIS in guaranteeing the covertness of wireless communications.**

***Index Terms*—Covert communication, STAR-RIS, multi-antenna, full-duplex, jamming.**

Manuscript received March 06, 2023; revised August 03, 2023; accepted November 02, 2023. Date of publication 30 January 2023; date of current version 18 April 2023. This work is supported in part by the National Natural Science Foundation of China (NSFC) under Grant 62201449 and Grant 62071370, in part by the Key R&D Projects of Shaanxi Province under Grant 2023-YBGY-040, in part by the Qin Chuang Yuan High-Level Innovation and Entrepreneurship Talent Program under Grant QCYRCXM-2022-231, and in part by the "Si Yuan Scholar" Foundation. The work of T.-X. Zheng was supported in part by the Open Research Fund of National Mobile Communications Research Laboratory, Southeast University, under Grant 2021D07, in part by the China Postdoctoral Science Foundation under Grants 2021M702631, in part by the Natural Science Basic Research Plan of Shaanxi Province under Grant 2022JM-320, and in part by the Fundamental Research Funds for the Central Universities under Grant xzy012021033. This paper will be presented in part at IEEE Global Communications Conference (GLOBECOM), Kuala Lumpur, Malaysia, December 2023. The associate editor coordinating the review of this article and approving it for publication was Prof. M. Cenk Gursoy. *(Corresponding author: Xiaoyan Hu.)*

H. Xiao, X. Hu, P. Mu, W. Wang, and T.-X. Zheng are with the School of Information and Communications Engineering, Xi'an Jiaotong University, Xi'an 710049, China. (email: hanxiaonuli@stu.xjtu.edu.cn, xiaoyanhu@xjtu.edu.cn, {pcmu, wjwang, zhengtx}@mail.xjtu.edu.cn).

K.-K. Wong is with the Department of Electronic and Electrical Engineering, University College London, London WC1E 7JE, U.K. (email: kai-kit.wong@ucl.ac.uk)

K. Yang is with the School of Computer Science and Electronic Engineering, University of Essex, Colchester CO4 3SQ, U.K. (e-mail: kunyang@essex.ac.uk).

I. Introduction

With the advent of 5G era, people are becoming increasingly dependent on wireless communications driven by the advanced communications and data processing techniques. Massive important and sensitive information, e.g., ID information, confidential documents, etc., are transmitted over open wireless networks, which aggravates the eavesdropping risk. Hence, people pay more and more attention to the problem of information security. Physical layer security (PLS) as a critical technology in protecting private information from eavesdropping attacks has drawn great attention in recent years [1]–[3]. However, PLS techniques cannot perform well in scenarios with covertness requirements, e.g., secret military operations, since PLS is only able to protect the content information of wireless communications but is unable to hide the existence of communications [4], [5]. Recently, the technology of covert communications has emerged as a new security paradigm and attracted significant research interests in both civilian and military applications [6], which can shelter the existence of communications between transceivers and provide a higher level of security for wireless communication systems.

A. Related Works

As a breakthrough work, [7] first proves the fundamental limit of covert communications over additive white Gaussian noise (AWGN) channels from the perspective of information theory. It demonstrates that $O(\sqrt{n})$ bits information can be transmitted covertly and reliably from transmitters to receivers over n channel uses while the warden can achieve correct detections if the amount of transmitted information exceeds this square root law. Actually, this conclusion is pessimistic since the intrinsic uncertainty of wireless channels and the background noise are not taken into consideration [7]. For example, [8] and [9] indicate that $O(n)$ bits information can be transmitted to the receiver when eavesdroppers do not exactly know the background noise power or the channel state information (CSI). Besides, existing works also resort to other uncertainties to enhance the performance of covert comunications [10]–[12], [4]. In particular, a full-duplex receiver is adopted in [10] for generating power-varying artificial noise to obtain a decent covert rate. In [11], random transmit power is leveraged to confuse the warden on detection of covert transmissions. An uninformed jammer is introduced in [12] to assist covert

communications by actively generating jamming signals under different channel models. Later in [4], a multi-jammer scheme with uncoordinated jammer selection is studied to defeat the warden. Considering more practical scenarios, [13] evaluates the effect of imperfect CSI on system covert rate, and [14] explores the case with multiple randomly distributed wardens and maximizes the average effective covert throughput by jointly optimizing the transmit power and blocklength.

The aforementioned works validate the effectiveness of covert communication techniques from different perspectives, however, they only investigate the single-antenna covert communication scenarios. In fact, multi-antenna technologies are beneficial in improving the capacity and reliability of traditional wireless communications which are also conductive in enhancing covert communication performance [5], [15], [16]. Specifically, in [5], a multi-antenna transmitter and a full-duplex jamming receiver are utilized to alleviate the influence caused by the uncertainty of the warden. The authors in [15] study the potential performance gain of centralized and distributed multi-antenna transmitters in covert communication systems with random positions for wardens and interferers. Different from the above situations, a multi-antenna adversary warden is considered in [16] which indicates that a slight increase in the antenna number of the adversary warden will result in a dramatical fall of covert rate.

Although multi-antenna techniques can enhance the covertness of communications, it cannot tackle the the randomness of wireless propagation environment. To break through this limitation, reconfigurable intelligent surface (RIS) has recently emerged as a promising solution [17]–[22]. In particular, RIS is usually a two-dimensional metamaterial consisting of a large number of low-cost passive and adjustable reflecting elements. The electromagnetic properties (e.g., phase and amplitude) of signals impinged on RIS can be adaptively adjusted with the assistance of RIS elements via a smart controller. Hence, the utilization of RISs is capable of reshaping desirable wireless propagation environment, which has attracted intensive attentions and been leveraged in many wireless communication scenarios including covert communications [19]–[22]. Specifically, [19] generally summarizes the application potentials of RIS in improving covert communications. Later in [20], the authors explore the performance gain of covert communications provided by RIS and first prove that the perfect covertness can be achieved with the aid of RIS when the instantaneous CSI of the warden is available. A multi-input multi-output (MIMO) covert communication system assisted by RIS is applied in [21] to resist the multi-antenna eavesdropper. Also, [22] investigates the RIS-assisted multi-antenna covert communications by jointly optimizing the active and passive beamformers.

B. Motivation and Contributions

It is worth noting that the RISs applied by the aforementioned works only reflect the incident signals which are limited to the scenarios that the transceivers locating at the same side of RISs. However, users may be at either side of RISs in practice, and thus the flexibility and effectiveness of conventional RIS appear inadequate. To overcome this limitation, a novel technology called simultaneously transmitting and reflecting RIS (STAR-RIS) has further emerged [23]. In particular, the incident signal arriving at the STAR-RIS will be separated into two parts, where one part is reflected to the same side of the incident signal and the other part is transited to the opposite side [24]. Note that STAR-RISs are capable of adjusting the reflected and transmitted signals by controlling the reflected and the transmitted coefficients simultaneously, which can help establish a more flexible full-space smart radio environment with 360° coverage. Therefore, STAR-RIS possesses a huge application potential in wireless communications which has attracted intensive research interests from both academia and industry [23]. However, the investigation of leveraging STAR-RISs into wireless communication systems is still in its infancy stage. As for secure communication systems, only a small number of state-of-the-art works have utilized STAR-RISs to enhance the system secure performance [25], [26].

To our best knowledge, the application of STAR-RIS in covert communications has not been studied in existing works. This is the first work investigates a STAR-RIS assisted multi-antenna covert communication scenario so as to fully exploit the potentials of STAR-RIS in covert communications. Our main contributions are summarized as follows:

- ***STAR-RIS-assisted Covert Communication Architecture:*** A STAR-RIS-assisted covert communication architecture is constructed through which the legitimate users located on both sides of the STAR-RIS can be simultaneously served. Through elaborately design the reflected and transmitted coefficients of the STAR-RIS, this architecture can highly enhance the covert performance of the system though more flexible reconfigurations on the random wireless environment.
- ***Closed-form Expressions for Covert System Indicators:*** Based on the constructed covert communication system, The closed-form expressions of the minimum detection error probability (DEP) and the corresponding optimal detection threshold at the warden are analytically derived considering the worst-case scenario. Based on a lower bound of the detection threshold and the large system analytic techniques, we further derive a lower bound of the average minimum DEP which is leveraged to jointly design the active and passive beamformers. The reasonability for choosing this lower bound is further validated by simulation results.
- ***Problem Formulation under Practical Constraints:*** We formulate an optimization problem aiming at maximizing the covert rate of the considered STAR-RIS-assisted covert communication system, under the covert communication constraint and the quality of service (QoS) constraint based on communication outage analysis, by jointly optimizing the active and passive beamforming at the base station (BS) and STAR-RIS. Due to the strongly coupled optimization variables and the characteristic amplitude constraint introduced by STAR-RIS, it is challenging to solve the formulated problem directly.
- ***Alternating Algorithm with Guaranteed Convergence:*** An optimization algorithm based on alternating strategy

is proposed to solve the formulated optimization problem in an iterative manner. In particular, the original problem is divided into three subproblems which are effectively solved by the semi-definite relaxation (SDR) method and Dinkelbach's algorithm. It is verified that the convergence of the proposed algorithm can always been guaranteed.

- ***Performance Improvement:*** The effectiveness of the proposed STAR-RIS-assisted algorithm is validated by numerical results where we evaluate the average covert rate in comparison with a benchmark using conventional RIS and a baseline algorithm called globally convergent version of method of moving asymptotes (GCMMA). It is shown that the proposed algorithm can achieve great performance improvement compared with the baselines and the advantages become even more obvious with a larger number of STAR-RIS elements.

The rest of this paper is organized as follows. In Section II, we introduce the STAR-RIS-aided covert communication system model. The DEP of the warden and communication outage probability based on this model are derived and analyzed in Section III. Section IV formulates the optimization problem and designs an iterative algorithm for jointly optimizing the passive and active beamforming. Section V shows the simulation results to validate the effectiveness of the proposed algorithm. Finally, a conclusion is drawn in Section VI

Notation: Operator $\circ$ denotes the Hadamard product. $(\cdot)^T$, $(\cdot)^H$ and $(\cdot)^*$ represent transpose, conjugate transpose and conjugate, respectively. $\mathrm{Diag}(\mathbf{a})$ denotes a diagonal matrix with diagonal elements in vector $\mathbf{a}$, and $\mathrm{diag}(\mathbf{A})$ denotes a vector whose elements are composed of the diagonal elements of matrix $\mathbf{A}$. $|\cdot|$ and $\|\cdot\|_2$ are the complex modulus and the spectral norm. $\mathbb{C}^{N\times M}$ stands for the set of $N \times M$ complex matrices. $x \sim \mathcal{CN}(a, b)$ and $x \sim \exp(\lambda)$ denote the circularly symmetric complex Gaussian random variable with mean a and variance b and the exponential random variable with mean λ, respectively. $\mathrm{Tr}(\cdot)$ represents the trace of a matrix. $\mathbf{A} \succeq 0$ indicates that matrix $\mathbf{A}$ is a positive semidefinite matrix. $\mathbf{I}_{N\times 1}$ represents the vector with $N \times 1$ entries that are 1.

II. System Model

In this paper, we consider a STAR-RIS-assisted covert communication system model as shown in Fig.1, mainly consisting of a M-antenna BS transmitter (Alice) assisted by a STAR-RIS with N elements, a covert user (Bob) and a warden user (Willie) both equipped with a single antenna, and an assistant public user (Carol) with two antennas. Willie tries to detect the existence of data transmissions from Alice to Bob, preparing for some security attacks. It is assumed that the single-antenna Bob and Willie work at the half-duplex mode, while the two-antenna Carol operates in the full-duplex mode where one antenna receives the transmitted singles from Alice and the other one transmits jamming signals to weaken Willie's detection ability. Due to the existence of blockages, we assume that there is no direct links between Alice and all users, which is reasonable in practical environment. The STAR-RIS is deployed at the users' vicinity to enhance the end-to-end communications between Alice and the legal users Bob and Carol while confusing the detection of the warden user Willie. Without loss of generality, we consider a scenario that Bob and Carol locate on opposite sides of the STAR-RIS which can be served simultaneously by the reflected (T) and transmitted (R) signals via STAR-RIS, respectively.[1] The energy splitting protocol is adopted for STAR-RIS whose all elements can work at T&R modes simultaneously [24].

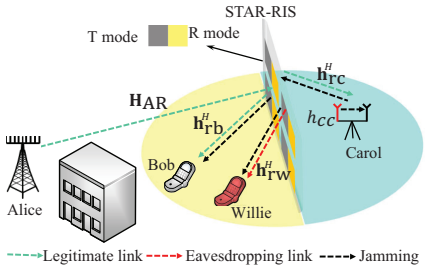


Fig. 1. System model for STAR-RIS-assisted covert communications.

The wireless communication channels from Alice to STAR-RIS, and from STAR-RIS to Bob, Carol, Willie are denoted as $\mathbf{H}_{\mathrm{AR}}=\sqrt{l_{\mathrm{AR}}}\mathbf{G}_{\mathrm{AR}}\in\ \mathbb{C}^{N\times M}$ and $\mathbf{h}_{\mathrm{rb}}=\sqrt{l_{\mathrm{rb}}}\mathbf{g}_{\mathrm{rb}}\in\ \mathbb{C}^{N\times 1}$, $\mathbf{h}_{\mathrm{rc}}=\sqrt{l_{\mathrm{rc}}}\mathbf{g}_{\mathrm{rc}}\in\ \mathbb{C}^{N\times 1}$, $\mathbf{h}_{\mathrm{rw}}=\sqrt{l_{\mathrm{rw}}}\mathbf{g}_{\mathrm{rw}}\in\ \mathbb{C}^{N\times 1}$, respectively. In particular, $\mathbf{G}_{\mathrm{AR}}$ and $\mathbf{g}_{\mathrm{rb}}$, $\mathbf{g}_{\mathrm{rc}}$, $\mathbf{g}_{\mathrm{rw}}$ are the small-scale Rayleigh fading coefficients whose entries are independent identically distributed (i.i.d.) following the complex Gaussian distribution with zero mean and unit variance. In addition, l_{AR} and l_{rb}, l_{rc}, l_{rw} are the large-scale path loss coefficients modeled as $\sqrt{\frac{\rho_0}{d^\alpha}}$, where ρ_0 denotes the reference power gain at a distance of one meter (m), α represents the path-loss exponent, and d corresponds to the node distances of d_{AR} and d_{rb}, d_{rc}, d_{rw}. We assume that the considered STAR-RIS-assisted covert communication system operates in the time division duplex (TDD) mode, so that the uplink channel estimation techniques based on STAR-RIS can be exploited to estimate the aforementioned CSI by utilizing channel reciprocity [27]. As for the full-duplex assistant user Carol, its self-interference channel can be modeled as $h_{\mathrm{cc}} = \sqrt{\phi}g_{\mathrm{cc}}$, where $g_{\mathrm{cc}} \sim \mathcal{CN}(0, \phi)$, $\phi \in [0, 1]$ is the self-interference cancellation (SIC) coefficient determined by the performing efficiency of the SIC [10], [28].

In this paper, we assume that the instantaneous CSI between STAR-RIS and Alice, Bob, Carol (i.e., $\mathbf{H}_{\mathrm{AR}}$, $\mathbf{h}_{\mathrm{rb}}$, $\mathbf{h}_{\mathrm{rc}}$) is available at Alice, while only the statistical CSI between STAR-RIS and the Willie ($\mathbf{h}_{\mathrm{rw}}$) is known at Alice. In contrast, it is assumed that Willie is capable to know the instantaneous CSI of all the users, i.e., $\mathbf{h}_{\mathrm{rw}}$, $\mathbf{h}_{\mathrm{rb}}$ and $\mathbf{h}_{\mathrm{rc}}$, but can only access the statistical CSI of Alice, i.e., $\mathbf{H}_{\mathrm{AR}}$. The reasonability of these assumptions can be verified as follows:

1) Alice periodically receives pilot signals sent by legitimate users Bob and Carol via the STAR-RIS, and thus it is capable to obtain the instantaneous CSI of $\mathbf{H}_{\mathrm{AR}}$, $\mathbf{h}_{\mathrm{rb}}$ and $\mathbf{h}_{\mathrm{rc}}$ through channel estimation techniques [27]. Although Willie wants to conceal itself, the inadvertent signal leakage is unavoidable in nearly all practical radiometers [29]. Hence, Alice can adopt some advanced detection

[1]Similar to [17] and [22], we ignore the signals reflected or transmitted more than once by the STAR-RIS considering the severe path losses.

tools, e.g., "Ghostbuster" in [29], [30], to capture the leakage signals and obtain Willie's suspected area so as to estimate the statistical CSI of $\mathbf{h}_{\mathrm{rw}}$.

2) With the received pilot signals from Bob and Carol, the instantaneous CSI between Bob/Carol/Willie and the STAR-RIS can be estimated by Willie, which is the worst-case scenario from the perspective of covert communication design because Willie is able to adjust its detection threshold based on the instantaneous CSI. It is more likely for Alice to keep silent during channels estimation, so the instantaneous CSI of $\mathbf{H}_{\mathrm{AR}}$ is unavailable at Willie. However, Willie possibly knows the suspected locations of Alice and the STAR-RIS as indicated in [10], and thus, Willie is able to obtain the statistical CSI of $\mathbf{H}_{\mathrm{AR}}$.

As for the jamming signals transmitted by Carol, we assume that the power of the jamming signals, denoted as P_{j}, follows the uniform distribution with $P_{\mathrm{j}}^{\max}$ being the maximum power limit [10], [22]. It is assumed that Willie can only obtain the distribution of the jamming power, and thus it is difficult for Willie to detect the existence of communications between Alice and Bob under the random jamming interference. . Under such a covert strategy, the received signals at Bob and Carol in the considered STAR-RIS-assisted covert communication system can be respectively expressed as

$$\begin{aligned} y_{\mathrm{b}}[k] =& \mathbf{h}_{\mathrm{rb}}^H \mathbf{\Theta}_{\mathrm{r}} \mathbf{H}_{\mathrm{AR}} \left(\mathbf{w}_{\mathrm{b}} s_{\mathrm{b}}[k] + \mathbf{w}_{\mathrm{c}} s_{\mathrm{c}}[k]\right) + \\ & \mathbf{h}_{\mathrm{rb}}^H \mathbf{\Theta}_{\mathrm{t}} \mathbf{h}_{\mathrm{rc}}^* \sqrt{P_{\mathrm{j}}} s_{\mathrm{j}}[k] + n_{\mathrm{b}}[k], \end{aligned} \tag{1}$$

$$\begin{aligned} y_{\mathrm{c}}[k] =& \mathbf{h}_{\mathrm{rc}}^H \mathbf{\Theta}_{\mathrm{t}} \mathbf{H}_{\mathrm{AR}} \left(\mathbf{w}_{\mathrm{b}} s_{\mathrm{b}}[k] + \mathbf{w}_{\mathrm{c}} s_{\mathrm{c}}[k]\right) + \\ & h_{\mathrm{cc}} \sqrt{P_{\mathrm{j}}} s_{\mathrm{j}}[k] + n_{\mathrm{c}}[k], \end{aligned} \tag{2}$$

where $k \in \mathcal{K} \triangleq \{1, \ldots, K\}$ denotes the index of each communication channel use with the maximum number of K in a time slot. $\mathbf{\Theta}_{\mathrm{r}} = \mathrm{Diag}\left\{\sqrt{\beta_{\mathrm{r}}^1} e^{\mathrm{j}\phi_{\mathrm{r}}^1}, \ldots, \sqrt{\beta_{\mathrm{r}}^N} e^{\mathrm{j}\phi_{\mathrm{r}}^N}\right\}$ and $\mathbf{\Theta}_{\mathrm{t}} = \mathrm{Diag}\left\{\sqrt{\beta_{\mathrm{t}}^1} e^{\mathrm{j}\phi_{\mathrm{t}}^1}, \ldots, \sqrt{\beta_{\mathrm{t}}^N} e^{\mathrm{j}\phi_{\mathrm{t}}^N}\right\}$ respectively indicate the STAR-RIS reflected and transmitted coefficient matrices, where $\beta_{\mathrm{r}}^n, \beta_{\mathrm{t}}^n \in [0,1]$, $\beta_{\mathrm{r}}^n + \beta_{\mathrm{t}}^n = 1$ and $\phi_{\mathrm{r}}^n, \phi_{\mathrm{t}}^n \in [0, 2\pi)$, for $\forall n \in \mathcal{N} \triangleq \{1, 2, \ldots, N\}$. In addition, $\mathbf{w}_{\mathrm{b}} \in \mathbb{C}^{M\times 1}$ and $\mathbf{w}_{\mathrm{c}} \in \mathbb{C}^{M\times 1}$ are the precoding vectors at Alice for Bob and Carol, respectively. $s_{\mathrm{b}}[k]$ and $s_{\mathrm{c}}[k] \sim \mathcal{CN}(0,1)$ are the signals transmitted by Alice to Bob and Carol while $s_{\mathrm{j}}[k] \sim \mathcal{CN}(0,1)$ is the jamming signal transmitted by Carol, where we try to hide the transmission of $s_{\mathrm{b}}[k]$ from the detection of Willie. Also, we use $n_{\mathrm{b}}[k] \sim \mathcal{CN}(0, \sigma_{\mathrm{b}}^2)$ and $n_{\mathrm{c}}[k] \sim \mathcal{CN}(0, \sigma_{\mathrm{c}}^2)$ to represent the AWGN noise received at Bob and Carol with σ_{b}^2 and σ_{c}^2 being the corresponding noise power.[2]

III. Analysis on STAR-RIS-Assisted Covert Communications

A. Covert Communication Detection Strategy at Willie

In this section, we detail the detection strategy of Willie for STAR-RIS-assisted covert communications from Alice to Bob. In particular, Willie attempts to judge whether there exists covert transmissions based on the received signal sequence $\{y_{\mathrm{w}}[k]\}_{k\in\mathcal{K}}$ in a time slot. Thus, Willie has to face a binary hypothesis for detection, which includes a null hypothesis, $\mathcal{H}_0$, representing that Alice only transmits public signals to Carol, and an alternative hypothesis, $\mathcal{H}_1$, indicating that Alice transmits both public signals and covert signals to Coral and Bob, respectively. Furthermore, the received signals at Willie based on the two hypotheses are given by

$$\begin{aligned} \mathcal{H}_0 : y_{\mathrm{w}}[k] =& \mathbf{h}_{\mathrm{rw}}^H \mathbf{\Theta}_{\mathrm{r}} \mathbf{H}_{\mathrm{AR}} \mathbf{w}_{\mathrm{c}} s_{\mathrm{c}}[k] + \mathbf{h}_{\mathrm{rb}}^H \mathbf{\Theta}_{\mathrm{t}} \mathbf{h}_{\mathrm{rc}}^* \sqrt{P_{\mathrm{j}}} s_{\mathrm{j}}[k] \\ & + n_{\mathrm{w}}[k],\ k \in \mathcal{K}, \end{aligned} \tag{3}$$

$$\begin{aligned} \mathcal{H}_1 : y_{\mathrm{w}}[k] =& \mathbf{h}_{\mathrm{rw}}^H \mathbf{\Theta}_{\mathrm{r}} \mathbf{H}_{\mathrm{AR}} \mathbf{w}_{\mathrm{b}} s_{\mathrm{b}}[k] + \mathbf{h}_{\mathrm{rw}}^H \mathbf{\Theta}_{\mathrm{r}} \mathbf{H}_{\mathrm{AR}} \mathbf{w}_{\mathrm{c}} s_{\mathrm{c}}[k] \\ & + \mathbf{h}_{\mathrm{rb}}^H \mathbf{\Theta}_{\mathrm{t}} \mathbf{h}_{\mathrm{rc}}^* \sqrt{P_{\mathrm{j}}} s_{\mathrm{j}}[k] + n_{\mathrm{w}}[k],\ k \in \mathcal{K}, \end{aligned} \tag{4}$$

where $n_{\mathrm{w}}[k] \sim \mathcal{CN}(0, \sigma_{\mathrm{w}}^2)$ is the AWGN received at Willie. We assume that Willie utilizes a radiometer to detect the covert signals from Alice to Bob, owing to its properties of low complexity and ease of implementation [11], [31].

According to the working mechanism of the radiometer, the average power of the received signals at Willie in a time slot, i.e., $\overline{P}_{\mathrm{w}} = \frac{1}{K}\sum_{k=1}^{K} |y_{\mathrm{w}}[k]|^2$, is employed for statistical test. Similar to the existing works, (e.g., [4], [10], [22]), we assume that Willie uses infinite number of signal samples to implement binary detection, i.e., $K \to \infty$. Hence, the average received power at Willie $\overline{P}_{\mathrm{w}}$ can be asymptotically approximated as (5) which is shown at the top of the next page. Willie needs to analyze $\overline{P}_{\mathrm{w}}$ to decide whether the communication between Alice and Bod is under the hypotheses of $\mathcal{H}_0$ or $\mathcal{H}_1$, and its decision rule can be presented as $\overline{P}_{\mathrm{w}} \underset{\mathcal{D}_0}{\overset{\mathcal{D}_1}{\gtrless}} \tau_{\mathrm{dt}}$, where $\mathcal{D}_0$ (or $\mathcal{D}_1$) indicates the decision that Willie favors $\mathcal{H}_0$ (or $\mathcal{H}_1$), and $\tau_{\mathrm{dt}} > 0$ is the corresponding detection threshold. In this paper, we adopt the DEP as Willie's detection performance metric and consider the worst case scenario that Willie can optimize its detection threshold to obtain the minimum DEP. According to the Neyman-Pearson criterion, the minimum DEP of Willie is the likelihood ratio [12], [22], which can be expressed as $\Lambda(\mathbf{y}_{\mathrm{w}}) = \frac{f_{\mathbf{y}_{\mathrm{w}}|\mathcal{H}_1}(\mathbf{y}_{\mathrm{w}}|\mathcal{H}_1)}{f_{\mathbf{y}_{\mathrm{w}}|\mathcal{H}_0}(\mathbf{y}_{\mathrm{w}}|\mathcal{H}_0)}$, where $\mathbf{y}_{\mathrm{w}} = \{y_{\mathrm{w}}[1], \ldots, y_{\mathrm{w}}[K]\}$ is Willie's received signal vector, $f_{\mathbf{y}_{\mathrm{w}}|\mathcal{H}_0}$ and $f_{\mathbf{y}_{\mathrm{w}}|\mathcal{H}_1}$ are the probability density functions (PDFs) of the sampling signals when $\mathcal{H}_0$ and $\mathcal{H}_1$ are true, respectively. However, it is difficult to derive the likelihood ratio due to the fact that the instantaneous CSI of $\mathbf{H}_{\mathrm{AR}}$ is unavailable at Willie, which introduces extra randomness in received signal $\mathbf{y}_{\mathrm{w}}$.

To solve this problem, we will derive the false alarm (FA) probability and the miss detection (MD) probability to obtain the minimum DEP. Specifically, FA probability represents the probability of Willie making the decision $\mathcal{D}_1$ under $\mathcal{H}_0$, i.e., $P_{\mathrm{FA}} = \Pr(\mathcal{D}_1 \mid \mathcal{H}_0)$, while MD probability is the probability of Willie making the decision $\mathcal{D}_0$ under $\mathcal{H}_1$, i.e., $P_{\mathrm{MD}} = \Pr(\mathcal{D}_0 \mid \mathcal{H}_1)$. Hence, the DEP of Willie can be written as

$$P_{\mathrm{e}} = P_{\mathrm{FA}} + P_{\mathrm{MD}} = \Pr(\mathcal{D}_1 \mid \mathcal{H}_0) + \Pr(\mathcal{D}_0 \mid \mathcal{H}_1). \tag{6}$$

It is easy to note that $0 \le P_{\mathrm{e}} \le 1$ where $P_{\mathrm{e}} = 0$ indicates that Willie can always correctly detect the existence of the cover communications between Alice and Bob, while $P_{\mathrm{e}} = 1$ means that Willie is never able to make correct detections. By choosing a reasonable detection threshold τ_{dt}, the minimum DEP denoted as P_{e}^*, can be obtained at Willie. Also, in order to achieve the covertness of communications, it is necessary

[2]It is worth noting that we ignore the jamming signals reflected by STAR-RIS at Carol mainly due to the fact that it is negligible compared with the self-interference jamming signals received by Carol.

$$\overline{P}_{\rm w}=\lim_{K\to+\infty}\frac{1}{K}\sum_{k=1}^{K}|y_{\rm w}[k]|^2=\begin{cases}\left|\mathbf{h}_{\rm rw}^H\mathbf{\Theta}_{\rm r}\mathbf{H}_{\rm AR}\mathbf{w}_{\rm c}\right|^2+\left|\mathbf{h}_{\rm rw}^H\mathbf{\Theta}_{\rm t}\mathbf{h}_{\rm rc}^*\right|^2P_{\rm j}+\sigma_{\rm w}^2, & \mathcal{H}_0,\\ \left|\mathbf{h}_{\rm rw}^H\mathbf{\Theta}_{\rm r}\mathbf{H}_{\rm AR}\mathbf{w}_{\rm b}\right|^2+\left|\mathbf{h}_{\rm rw}^H\mathbf{\Theta}_{\rm r}\mathbf{H}_{\rm AR}\mathbf{w}_{\rm c}\right|^2+\left|\mathbf{h}_{\rm rw}^H\mathbf{\Theta}_{\rm t}\mathbf{h}_{\rm rc}^*\right|^2P_{\rm j}+\sigma_{\rm w}^2, & \mathcal{H}_1,\end{cases}\tag{5}$$

to guarantee $P_{\rm e}^*\geq 1-\epsilon$, where $\epsilon\in(0,1)$ is a quite small value required by the system performance indicators.

B. Analysis on Detection Error Probability

In this section, we first derive the analytical expressions for $P_{\rm FA}$ and $P_{\rm MD}$ in closed form, based on the distribution of $\overline{P}_{\rm w}$ under $\mathcal{H}_0$ and $\mathcal{H}_1$. Then, the optimal detection threshold $\tau_{\rm dt}^*$ and the minimum DEP $P_{\rm e}^*$ are obtained by analyzing the analytical expression of DEP. Considering the fact that the instantaneous CSI of channel $\mathbf{h}_{\rm rw}$ is not available at Alice, thus we derive a lower bound for the average the minimum DEP over $\mathbf{h}_{\rm rw}$. In particular, the analytical expressions for $P_{\rm FA}$ and $P_{\rm MD}$ can be obtained through Theorem 1.

Theorem 1. *The analytical expressions for FA probability $P_{\rm FA}$ and MD probability $P_{\rm MD}$ are respectively given as*

$$P_{\rm FA}=\begin{cases}1, & \tau_{\rm dt}<\sigma_{\rm w}^2,\\ 1-\frac{\left(\tau_{\rm dt}-\sigma_{\rm w}^2\right)+\lambda e^{-\frac{\tau_{\rm dt}-\sigma_{\rm w}^2}{\lambda}}-\lambda}{\gamma P_{\rm j}^{\max}}, & \sigma_{\rm w}^2\leq\tau_{\rm dt}<\sigma_{\rm w}^2+\gamma P_{\rm j}^{\max},\\ \frac{e^{-\frac{\tau_{\rm dt}-\sigma_{\rm w}^2}{\lambda}}\left(e^{\frac{\gamma P_{\rm j}^{\max}}{\lambda}}-1\right)\lambda}{\gamma P_{\rm j}^{\max}}, & \tau_{\rm dt}\geq\sigma_{\rm w}^2+\gamma P_{\rm j}^{\max},\end{cases}\tag{7}$$

$$P_{\rm MD}=\begin{cases}0, & \tau_{\rm dt}<\sigma_{\rm w}^2,\\ \frac{\left(\tau_{\rm dt}-\sigma_{\rm w}^2\right)+\tilde{\lambda}e^{-\frac{\tau_{\rm dt}-\sigma_{\rm w}^2}{\tilde{\lambda}}}-\tilde{\lambda}}{\gamma P_{\rm j}^{\max}}, & \sigma_{\rm w}^2\leq\tau_{\rm dt}<\sigma_{\rm w}^2+\gamma P_{\rm j}^{\max},\\ 1-\frac{e^{-\frac{\tau_{\rm dt}-\sigma_{\rm w}^2}{\tilde{\lambda}}}\left(e^{\frac{\gamma P_{\rm j}^{\max}}{\tilde{\lambda}}}-1\right)\tilde{\lambda}}{\gamma P_{\rm j}^{\max}}, & \tau_{\rm dt}\geq\gamma P_{\rm j}^{\max}+\sigma_{\rm w}^2,\end{cases}\tag{8}$$

which are shown in closed form with $\lambda=\left\|\mathbf{h}_{\rm rw}^H\mathbf{\Theta}_{\rm r}\right\|_2^2\mathbf{w}_{\rm c}^H\mathbf{w}_{\rm c}$, $\tilde{\lambda}=\left\|\mathbf{h}_{\rm rw}^H\mathbf{\Theta}_{\rm r}\right\|_2^2\left(\mathbf{w}_{\rm b}^H\mathbf{w}_{\rm b}+\mathbf{w}_{\rm c}^H\mathbf{w}_{\rm c}\right)$, and $\gamma=\left|\mathbf{h}_{\rm rw}^H\mathbf{\Theta}_{\rm t}\mathbf{h}_{\rm rc}^*\right|^2$.

Proof: The proof is given in Appendix A. ∎

Note that $\sigma_{\rm w}^2$ and $\sigma_{\rm w}^2+\gamma P_{\rm j}^{\max}$ are two important boundaries affecting the values of $P_{\rm FA}$ and $P_{\rm MD}$ in (7) and (8). Specifically, when $\tau_{\rm dt}\leq\sigma_{\rm w}^2$, complete FAs will be performed by Willie and the MDs can be totally avoided. Furthermore, with the increase of detection threshold $\tau_{\rm dt}$ from $\sigma_{\rm w}^2$ to $+\infty$, $P_{\rm FA}$ will experience a decrease from 1 to 0, while $P_{\rm MD}$ will have an opposite trend. Based on the analytical expressions of (7) and (8), the analytical closed-form expressions of DEP can be given as (9) which is shown at the top of the next page.

It is important to point out that we consider the worst case scenario that Willie can optimize its detection threshold to minimize the DEP. Here, the closed-form solution of the optimal $\tau_{\rm dt}^*$ is provided in the following theorem.

Theorem 2. *The optimal detection threshold $\tau_{\rm dt}^*$ to minimize the DEP of Willie in the considered STAR-RIS-assisted communication system is given by*

$$\tau_{\rm dt}^*=\frac{\tilde{\lambda}\lambda}{\tilde{\lambda}-\lambda}\ln\Delta+\sigma_{\rm w}^2\in\left[\sigma_{\rm w}^2+\gamma P_{\rm j}^{\max},\infty\right),\tag{10}$$

where $\Delta=\frac{e^{\frac{\gamma P_{\rm j}^{\max}}{\lambda}}-1}{e^{\frac{\gamma P_{\rm j}^{\max}}{\tilde{\lambda}}}-1}$ is a function of λ, $\tilde{\lambda}$ and γ.

Proof: The proof is given in Appendix B. ∎

Substituting (10) into (9) and adopting some algebraic manipulations, the analytical closed-form expression of the minimum DEP can be obtained as

$$P_{\rm e}^*=1-\frac{\tilde{\lambda}\left(e^{\frac{\gamma P_{\rm j}^{\max}}{\lambda}}-1\right)(\Delta)^{\frac{\lambda}{\lambda-\tilde{\lambda}}}-\lambda\left(e^{\frac{\gamma P_{\rm j}^{\max}}{\lambda}}-1\right)(\Delta)^{\frac{\tilde{\lambda}}{\lambda-\tilde{\lambda}}}}{\gamma P_{\rm j}^{\max}}.\tag{11}$$

Since Alice only knows the statistical CSI of channel $\mathbf{h}_{\rm rw}$, the average minimum DEP over $\mathbf{h}_{\rm rw}$, denoted as $\overline{P}_{\rm e}^*=\mathbb{E}_{\mathbf{h}_{\rm rw}}\left(P_{\rm e}^*\right)$, is usually utilized to evaluate the covert communications between Alice and Bob [5], [31]. In (11), λ, $\tilde{\lambda}$ and γ are all random variables including $\mathbf{h}_{\rm rw}$, and thus they are coupled to each other, which makes it challenging to calculate the $\overline{P}_{\rm e}^*$ directly. To solve this problem, large system analytic techniques are utilized to remove the couplings among λ, $\tilde{\lambda}$ and γ, which are widely adopted to analyze the performance limitation of wireless communication systems (e.g., [22], [32]–[35]). By assuming that the STAR-RIS is equipped with a large number of low-cost elements, then we can obtain the asymptotic analytic result of $P_{\rm e}^*$. In particular, we first apply the large system analytic technique on λ, then the asymptotic equality about λ can be given as

$$\begin{aligned}\lim_{N\to\infty}\frac{\left\|\mathbf{h}_{\rm rw}^H\mathbf{\Theta}_{\rm r}\right\|_2^2\varpi_{\rm c}}{N}&=\lim_{N\to\infty}\frac{\mathbf{h}_{\rm rw}^H\mathbf{\Theta}_{\rm r}\mathbf{\Theta}_{\rm r}^H\mathbf{h}_{\rm rw}\varpi_{\rm c}}{N}\\&\overset{(a)}{\to}\frac{l_{\rm rw}\varpi_{\rm c}}{N}\,{\rm tr}\left(\mathbf{\Theta}_{\rm r}\mathbf{\Theta}_{\rm r}^H\right)\\&=\frac{l_{\rm rw}\varpi_{\rm c}}{N}\theta_{\rm r}=\frac{\lambda_{\rm a}}{N},\end{aligned}\tag{12}$$

where the convergence (a) is due to [32, Corollary 1]. Here, $\varpi_{\rm c}=\mathbf{w}_{\rm c}^H\mathbf{w}_{\rm c}$, $\theta_{\rm r}={\rm diag}(\mathbf{\Theta}_{\rm r})^H\,{\rm diag}(\mathbf{\Theta}_{\rm r})$, and $\lambda_{\rm a}=l_{\rm rw}\varpi_{\rm c}\theta_{\rm r}$ is the asymptotic result of λ. Similarly, the asymptotic result of $\tilde{\lambda}$ can be expressed as $\tilde{\lambda}_{\rm a}=l_{\rm rw}\theta_{\rm r}(\varpi_{\rm b}+\varpi_{\rm c})$, where we define $\varpi_{\rm b}=\mathbf{w}_{\rm b}^H\mathbf{w}_{\rm b}$.

With the results of $\lambda_{\rm a}$ and $\tilde{\lambda}_{\rm a}$ based on the large system analytic technique, the uncertainty of λ and $\tilde{\lambda}$ can be removed from the perspective of Alice. Substituting $\lambda_{\rm a}$ and $\tilde{\lambda}_{\rm a}$ into (11), we obtain the asymptotic analytical result of the minimum DEP $P_{\rm e}^*$ with respect to (w.r.t.) the random variable γ

$$P_{\rm ea}^*=1-\frac{l_{\rm rw}\theta_{\rm r}\varpi_{\rm b}}{\gamma P_{\rm j}^{\max}}(\Delta_{\rm a}(\gamma))^{\frac{-\varpi_{\rm c}}{\varpi_{\rm b}}}\left(e^{\frac{\gamma P_{\rm j}^{\max}}{l_{\rm rw}\theta_{\rm r}(\varpi_{\rm b}+\varpi_{\rm c})}}-1\right).\tag{13}$$

$$P_{\mathrm{e}}=\begin{cases}1, & \tau_{\mathrm{dt}}<\sigma_{\mathrm{w}}^2,\\ 1+\frac{\tilde{\lambda}\left(e^{-\frac{\tau_{\mathrm{dt}}-\sigma_{\mathrm{w}}^2}{\tilde{\lambda}}}-1\right)-\lambda e^{-\frac{\tau_{\mathrm{dt}}-\sigma_{\mathrm{w}}^2}{\lambda}}+\lambda}{\gamma P_{\mathrm{j}}^{\max}}, & \sigma_{\mathrm{w}}^2\le\tau_{\mathrm{dt}}<\sigma_{\mathrm{w}}^2+\gamma P_{\mathrm{j}}^{\max},\\ 1+\frac{\tilde{\lambda}e^{-\frac{\tau_{\mathrm{dt}}-\sigma_{\mathrm{w}}^2}{\tilde{\lambda}}}\left(1-e^{\frac{\gamma P_{\mathrm{j}}^{\max}}{\tilde{\lambda}}}\right)+\lambda e^{-\frac{\tau_{\mathrm{dt}}-\sigma_{\mathrm{w}}^2}{\lambda}}\left(e^{\frac{\gamma P_{\mathrm{j}}^{\max}}{\lambda}}-1\right)}{\gamma P_{\mathrm{j}}^{\max}}, & \tau_{\mathrm{dt}}\ge\gamma P_{\mathrm{j}}^{\max}+\sigma_{\mathrm{w}}^2,\end{cases}\tag{9}$$

It is easy to verify that $\gamma=\left|\mathbf{h}_{\mathrm{rw}}^H\mathbf{\Theta}_{\mathrm{t}}\mathbf{h}_{\mathrm{rc}}^*\right|^2\sim\exp(\lambda_{\mathrm{rw}})$ where $\lambda_{\mathrm{rw}}=\|\mathbf{\Theta}_{\mathrm{t}}\mathbf{h}_{\mathrm{rc}}^*\|_2^2$. By averaging P_{ea}^* over γ, we can get the average asymptotic analytical result of the minimum DEP as

$$\begin{aligned}\overline{P}_{\mathrm{ea}}^*&=\mathbb{E}_\gamma\left(P_{\mathrm{ea}}^*\right)\\&=\int_0^{+\infty}\left(1-\frac{l_{\mathrm{rw}}\theta_{\mathrm{r}}\varpi_{\mathrm{b}}}{\gamma P_{\mathrm{j}}^{\max}}(\Delta_{\mathrm{a}}(\gamma))^{\frac{-\varpi_{\mathrm{c}}}{\varpi_{\mathrm{b}}}}\right.\\&\quad\left.\left(e^{\frac{\gamma P_{\mathrm{j}}^{\max}}{l_{\mathrm{rw}}\theta_{\mathrm{r}}(\varpi_{\mathrm{b}}+\varpi_{\mathrm{c}})}}-1\right)\right)\times\frac{1}{\lambda_{\mathrm{rw}}}e^{-\frac{\gamma}{\lambda_{\mathrm{rw}}}}d\gamma.\end{aligned}\tag{14}$$

Due to the existence of $\Delta_{\mathrm{a}}(\gamma)$ in P_{ea}^*, the integral in (14) for calculating $\overline{P}_{\mathrm{ea}}^*$ over the random variable γ is non-integrable. Therefore, the exact analytical expression for $\overline{P}_{\mathrm{ea}}^*$ is mathematically intractable. In order to guarantee the covert constraint $\overline{P}_{\mathrm{ea}}^*\ge1-\epsilon$ always holds, we further adopt a lower bound of $\overline{P}_{\mathrm{ea}}^*$ to evaluate the covertness of communications.

Specifically, we use a lower bound denoted as $\hat{\Delta}_{\mathrm{a}}(\gamma)\triangleq e^{\gamma P_{\mathrm{j}}^{\max}\left(\frac{\tilde{\lambda}_{\mathrm{a}}-\lambda_{\mathrm{a}}}{\lambda_{\mathrm{a}}\lambda_{\mathrm{a}}}\right)}$ to replace $\Delta_{\mathrm{a}}(\gamma)$. It is easy to demonstrate that $\Delta_{\mathrm{a}}(\gamma)>\hat{\Delta}_{\mathrm{a}}(\gamma)$ always holds when $\gamma\in(0,+\infty)$, and the relative gap between these two variables gradually goes to zero with the increase of γ.[3] Replacing the $\Delta_{\mathrm{a}}(\gamma)$ in (14) with $\hat{\Delta}_{\mathrm{a}}(\gamma)$, we can get a lower bound of $\overline{P}_{\mathrm{ea}}^*$ given by

$$\begin{aligned}\hat{P}_{\mathrm{ea}}^*&\triangleq\int_0^{+\infty}\left(1-\frac{l_{\mathrm{rw}}\theta_{\mathrm{r}}\varpi_{\mathrm{b}}}{\gamma P_{\mathrm{j}}^{\max}}\left(\hat{\Delta}_{\mathrm{a}}(\gamma)\right)^{\frac{-\varpi_{\mathrm{c}}}{\varpi_{\mathrm{b}}}}\right.\\&\quad\left.\left(e^{\frac{\gamma P_{\mathrm{j}}^{\max}}{l_{\mathrm{rw}}\theta_{\mathrm{r}}(\varpi_{\mathrm{b}}+\varpi_{\mathrm{c}})}}-1\right)\right)\times\frac{1}{\lambda_{\mathrm{rw}}}e^{-\frac{\gamma}{\lambda_{\mathrm{rw}}}}d\gamma\\&=1+\frac{l_{\mathrm{rw}}\theta_{\mathrm{r}}\varpi_{\mathrm{b}}\ln\left(\frac{l_{\mathrm{rw}}\theta_{\mathrm{r}}(\varpi_{\mathrm{b}}+\varpi_{\mathrm{c}})}{l_{\mathrm{rw}}\theta_{\mathrm{r}}(\varpi_{\mathrm{b}}+\varpi_{\mathrm{c}})+P_{\mathrm{j}}^{\max}\lambda_{\mathrm{rw}}}\right)}{P_{\mathrm{j}}^{\max}\lambda_{\mathrm{rw}}}<\overline{P}_{\mathrm{ea}}^*.\end{aligned}\tag{15}$$

Therefore, in the following sections, $\hat{P}_{\mathrm{ea}}^*\ge1-\epsilon$ will be leveraged as a tighter covert constraint to jointly design the active and passive beamforming variables of the system.

C. Analysis on Communication Outage Probability

Note that the instantaneous jamming power P_{j} and the self-interference channel h_{cc} of Carol are unavailable at Alice, thus the randomness introduced by P_{j} and h_{cc} is possible to incur communication outage between Alice and Bob/Carol. When the required communication rate from Alice to Bob (R_{b}) or Carol (R_{c}) exceeds the corresponding channel capacity (C_{b},C_{c}), the communication outage occurs. The communication outage probability at Bob and Carol can be expressed as $\delta_{\mathrm{AB}}=\Pr\left(C_{\mathrm{b}}<R_{\mathrm{b}}\right),\delta_{\mathrm{AC}}=\Pr\left(C_{\mathrm{c}}<R_{\mathrm{c}}\right)$. To guarantee the communication quality between Alice and Bob/Carol, the issue of communication outage should be considered.

It is known that the channel capacity at Bob and Carol can be respectively written as

$$C_{\mathrm{b}}=\log_2\left(1+\frac{\left|\mathbf{h}_{\mathrm{rb}}^H\mathbf{\Theta}_{\mathrm{r}}\mathbf{H}_{\mathrm{AR}}\mathbf{w}_{\mathrm{b}}\right|^2}{\left|\mathbf{h}_{\mathrm{rb}}^H\mathbf{\Theta}_{\mathrm{r}}\mathbf{H}_{\mathrm{AR}}\mathbf{w}_{\mathrm{c}}\right|^2+\left|\mathbf{h}_{\mathrm{rb}}^H\mathbf{\Theta}_{\mathrm{t}}\mathbf{h}_{\mathrm{rc}}^*\right|^2P_{\mathrm{j}}+\sigma_{\mathrm{b}}^2}\right),\tag{16}$$

$$C_{\mathrm{c}}=\log_2\left(1+\frac{\left|\mathbf{h}_{\mathrm{rc}}^H\mathbf{\Theta}_{\mathrm{t}}\mathbf{H}_{\mathrm{AR}}\mathbf{w}_{\mathrm{c}}\right|^2}{\left|\mathbf{h}_{\mathrm{rc}}^H\mathbf{\Theta}_{\mathrm{t}}\mathbf{H}_{\mathrm{AR}}\mathbf{w}_{\mathrm{b}}\right|^2+\left|h_{\mathrm{cc}}\right|^2P_{\mathrm{j}}+\sigma_{\mathrm{c}}^2}\right).\tag{17}$$

Hence, when the required transmission rate between Alice and Bob (or Carol) is selected as R_{b} (or R_{c}), the closed-form expressions of the communication outage probabilities at Bob and Carol can be obtained through Theorem 3.

Theorem 3. *The communication outage probabilities between Alice and Bob/Carol are respectively derived as*

$$\delta_{\mathrm{AB}}=\begin{cases}0, & \Upsilon>P_{\mathrm{j}}^{\max},\\ 1-\frac{\Upsilon}{P_{\mathrm{j}}^{\max}}, & 0\le\Upsilon\le P_{\mathrm{j}}^{\max},\\ 1, & \Upsilon<0,\end{cases}\tag{18}$$

$$\delta_{\mathrm{AC}}=\begin{cases}e^{-\frac{\Gamma}{\phi P_{\mathrm{j}}^{\max}}}+\frac{\Gamma}{\phi P_{\mathrm{j}}^{\max}}\mathrm{Ei}\left(-\frac{\Gamma}{\phi P_{\mathrm{j}}^{\max}}\right), & \Gamma\ge0,\\ 1, & \Gamma<0,\end{cases}\tag{19}$$

where $\Upsilon=\frac{\left|\mathbf{h}_{\mathrm{rb}}^H\mathbf{\Theta}_{\mathrm{r}}\mathbf{H}_{\mathrm{AR}}\mathbf{w}_{\mathrm{b}}\right|^2-\left(2^{R_{\mathrm{b}}}-1\right)\left(\left|\mathbf{h}_{\mathrm{rb}}^H\mathbf{\Theta}_{\mathrm{r}}\mathbf{H}_{\mathrm{AR}}\mathbf{w}_{\mathrm{c}}\right|^2+\sigma_{\mathrm{b}}^2\right)}{\left(2^{R_{\mathrm{b}}}-1\right)\left|\mathbf{h}_{\mathrm{rb}}^H\mathbf{\Theta}_{\mathrm{t}}\mathbf{h}_{\mathrm{rc}}^*\right|^2}$ *and* $\Gamma=\frac{\left|\mathbf{h}_{\mathrm{rc}}^H\mathbf{\Theta}_{\mathrm{t}}\mathbf{H}_{\mathrm{AR}}\mathbf{w}_{\mathrm{c}}\right|^2-\left(2^{R_{\mathrm{c}}}-1\right)\left(\left|\mathbf{h}_{\mathrm{rc}}^H\mathbf{\Theta}_{\mathrm{t}}\mathbf{H}_{\mathrm{AR}}\mathbf{w}_{\mathrm{b}}\right|^2+\sigma_{\mathrm{c}}^2\right)}{\left(2^{R_{\mathrm{c}}}-1\right)}$. *In* (19), $\mathrm{Ei}(x)=-\int_{-x}^{\infty}\frac{e^{-t}}{t}dt$ *is the exponential internal function.*

Proof: The proof is given in Appendix C. ■

The communication outage constraints are then defined as $\delta_{\mathrm{AB}}\le\iota$ and $\delta_{\mathrm{AC}}\le\kappa$ where ι and κ are two communication outage thresholds required by the system performance indicators for Bob and Carol, respectively. In this paper, we try to maximize the covet rate of Bob under the covert constraint $\hat{\mathrm{P}}_{\mathrm{ea}}^*\ge1-\epsilon$ and the communication outage constraints. It is easy to note that δ_{AB} and δ_{AC} are segment functions with uncertain segment points Υ and Γ determined by the active and passive beamforming variables, i.e., $\mathbf{w}_{\mathrm{b}}$, $\mathbf{w}_{\mathrm{c}}$ and $\mathbf{\Theta}_{\mathrm{r}}$, $\mathbf{\Theta}_{\mathrm{t}}$, which will be jointly optimized in the next section. Hence, it is difficult to handle the two communication outage constraints in an optimization problem. In order to facilitate the optimization and analysis of the considered problem, we equivalently re-

[3]The reasonability about the selection of $\hat{\Delta}_{\mathrm{a}}(\gamma)$ (or $\hat{P}_{\mathrm{ea}}^*$) will be further verified based on the simulation results.

express the two communication outage constraints as

$$\Upsilon \geq P_{\mathrm{j}}^{\max}(1-\iota) \Leftrightarrow R_{\mathrm{b}} \leq R_{\mathrm{bb}}, \tag{20}$$

$$\Gamma \geq \sigma^{*} \Leftrightarrow R_{\mathrm{c}} \leq R_{\mathrm{cc}}, \tag{21}$$

where σ^{*} is the solution to the equation of $\delta_{\mathrm{AC}} = \kappa$ which can be numerically solved by the bi-section search method. In addition, R_{bb} and R_{cc} are respectively the upper bounds of R_{b} and R_{c} in order to guarantee the two communication outage constraints which can be expressed as

$$R_{\mathrm{bb}} = \log_2\left(1+\frac{\left|\mathbf{h}_{\mathrm{rb}}^{H}\boldsymbol{\Theta}_{\mathrm{r}}\mathbf{H}_{\mathrm{AR}}\mathbf{w}_{\mathrm{b}}\right|^2}{\left|\mathbf{h}_{\mathrm{rb}}^{H}\boldsymbol{\Theta}_{\mathrm{r}}\mathbf{H}_{\mathrm{AR}}\mathbf{w}_{\mathrm{c}}\right|^2+\left|\mathbf{h}_{\mathrm{rb}}^{H}\boldsymbol{\Theta}_{\mathrm{t}}\mathbf{h}_{\mathrm{rc}}^{*}\right|^2 P_{\mathrm{j}}^{\max}(1-\iota)+\sigma_{\mathrm{b}}^2}\right), \tag{22}$$

$$R_{\mathrm{cc}} = \log_2\left(1+\frac{\left|\mathbf{h}_{\mathrm{rc}}^{H}\boldsymbol{\Theta}_{\mathrm{t}}\mathbf{H}_{\mathrm{AR}}\mathbf{w}_{\mathrm{c}}\right|^2}{\left|\mathbf{h}_{\mathrm{rc}}^{H}\boldsymbol{\Theta}_{\mathrm{t}}\mathbf{H}_{\mathrm{AR}}\mathbf{w}_{\mathrm{b}}\right|^2+\sigma^{*}+\sigma_{\mathrm{c}}^2}\right). \tag{23}$$

Base on the above analysis, we know that the maximum covert rate for Bob under the communication outage constraint is R_{bb}, then we can maximize R_{bb} accordingly to improve the system covert performance. Similarly, the maximum communication rate for Carol under the communication outage constraint is R_{cc}, and we introduce a constraint $R_{\mathrm{cc}} \geq R^{*}$ to guarantee the QoS for the assistant user Carol where R^{*} is a minimum required communication rate for Carol.

IV. Problem Formulation and Algorithm Design

A. Optimization Problem Formulation

$$\max_{\boldsymbol{\Theta}_{\mathrm{r}},\boldsymbol{\Theta}_{\mathrm{t}},\mathbf{w}_{\mathrm{b}},\mathbf{w}_{\mathrm{c}}} R_{\mathrm{bb}},$$

$$\text{s.t. } \|\mathbf{w}_{\mathrm{b}}\|_2^2+\|\mathbf{w}_{\mathrm{c}}\|_2^2 \leq P_{\max} \tag{24a}$$

$$\frac{l_{\mathrm{rw}}\theta_{\mathrm{r}}\varpi_{\mathrm{b}}\ln\left(\frac{l_{\mathrm{rw}}\theta_{\mathrm{r}}(\varpi_{\mathrm{b}}+\varpi_{\mathrm{c}})}{l_{\mathrm{rw}}\theta_{\mathrm{r}}(\varpi_{\mathrm{b}}+\varpi_{\mathrm{c}})+P_{\mathrm{j}}^{\max}\lambda_{\mathrm{rw}}}\right)}{P_{\mathrm{j}}^{\max}\lambda_{\mathrm{rw}}} \geq -\epsilon, \tag{24b}$$

$$R_{\mathrm{cc}} \geq R^{*}, \tag{24c}$$

$$\beta_{\mathrm{r}}^{n}+\beta_{\mathrm{t}}^{n}=1, \phi_{\mathrm{r}}^{n}, \phi_{\mathrm{t}}^{n} \in [0, 2\pi). \tag{24d}$$

On the basis of the previous discussions in section III, we formulate the optimization problem in this section. Specifically, we will maximize the covert rate between Alice and Bob under the covert communication constraint while ensuring the QoS at Carol with the QoS constraint, by jointly optimizing the active and passive beamforming variables, i.e., $\mathbf{w}_{\mathrm{b}}$, $\mathbf{w}_{\mathrm{c}}$ and $\boldsymbol{\Theta}_{\mathrm{r}}$, $\boldsymbol{\Theta}_{\mathrm{t}}$. Hence, the optimized problem is formulated as (24). Where (24a) is the transmission power constraint for Alice with $P_{\max}$ being the maximum transmitted power; (24b) is an equivalent covert communication constraint of $\hat{P}_{\mathrm{ea}}^{*} \geq 1-\epsilon$; (24c) represents the QoS constraint for Carol; (24d) is the amplitude and phase shift constraints for STAR-RIS. Actually, it is challenging to solve the formulated optimization problem because of the following reasons. Firstly, the active and passive beamforming variables $\mathbf{w}_{\mathrm{b}}$, $\mathbf{w}_{\mathrm{c}}$ and $\boldsymbol{\Theta}_{\mathrm{r}}$, $\boldsymbol{\Theta}_{\mathrm{t}}$ are strongly coupled in the objective function, covert communication constraint (24b) and QoS constraint (24c). In addition, the utilization of STAR-RIS introduces a characteristic amplitude constraint (24d) due to the fact that $\boldsymbol{\Theta}_{\mathrm{r}}$ and $\boldsymbol{\Theta}_{\mathrm{t}}$ depend on each other in terms of element amplitudes.

Hence, the traditional convex optimization algorithms cannot be used directly to solve the non-convex optimized problem (24). To tackle this issue, the alternating strategy is leveraged to design the optimization algorithm. Specifically, we first divide the original problem into three subproblems where two subproblems are focused on designing the active beamformer variables $\mathbf{w}_{\mathrm{b}}$, $\mathbf{w}_{\mathrm{c}}$, while the passive beamformer variables $\boldsymbol{\Theta}_{\mathrm{r}}$, $\boldsymbol{\Theta}_{\mathrm{t}}$ are obtained by solving the last subproblem. After the convergence of the algorithm, we can finally obtain the solution for joint active and passive beamforming.

B. Active Beamforming Design

In this section, we formulate two subproblems based on the original problem (24) which are solved to design the active beamforming variables $\mathbf{w}_{\mathrm{b}}$ and $\mathbf{w}_{\mathrm{c}}$, respectively.

1) Active Beamforming Design for $\mathbf{w}_{\mathrm{b}}$: First, we consider the active beamforming design for $\mathbf{w}_{\mathrm{b}}$ with given $\mathbf{w}_{\mathrm{c}}$ and the passive beamforming variables $\boldsymbol{\Theta}_{\mathrm{r}}$, $\boldsymbol{\Theta}_{\mathrm{t}}$. In this circumstance, the objective function of the original problem turns into maximizing $\left|\mathbf{h}_{\mathrm{rb}}^{H}\boldsymbol{\Theta}_{\mathrm{r}}\mathbf{H}_{\mathrm{AR}}\mathbf{w}_{\mathrm{b}}\right|^2$, and the QoS constraint (24c) is equivalent to $\left|\mathbf{h}_{\mathrm{rc}}^{H}\boldsymbol{\Theta}_{\mathrm{t}}\mathbf{H}_{\mathrm{AR}}\mathbf{w}_{\mathrm{b}}\right|^2 \leq f(R^{*})$ with $f(R^{*}) = \frac{\left|\boldsymbol{\vartheta}_{\mathrm{t}}^{T}\mathbf{H}_{\mathrm{rc}}^{*}\mathbf{H}_{\mathrm{AR}}\mathbf{w}_{\mathrm{c}}\right|^2}{2^{R^{*}}-1}-\sigma^{*}-\sigma_{\mathrm{c}}^2$. Hence, the corresponding subproblem can be formulated as

$$\max_{\mathbf{w}_{\mathrm{b}}} \left|\boldsymbol{\vartheta}_{\mathrm{r}}^{T}\mathbf{H}_{\mathrm{rb}}^{*}\mathbf{H}_{\mathrm{AR}}\mathbf{w}_{\mathrm{b}}\right|^2,$$

$$\text{s.t. } \|\mathbf{w}_{\mathrm{b}}\|_2^2 \leq P_{\max}-\|\mathbf{w}_{\mathrm{c}}\|_2^2, \tag{25a}$$

$$\frac{l_{\mathrm{rw}}\theta_{\mathrm{r}}\varpi_{\mathrm{b}}\ln\left(1+\frac{P_{\mathrm{j}}^{\max}\lambda_{\mathrm{rw}}}{l_{\mathrm{rw}}\theta_{\mathrm{r}}(\varpi_{\mathrm{b}}+\varpi_{\mathrm{c}})}\right)}{P_{\mathrm{j}}^{\max}\lambda_{\mathrm{rw}}} \leq \epsilon, \tag{25b}$$

$$\left|\boldsymbol{\vartheta}_{\mathrm{t}}^{T}\mathbf{H}_{\mathrm{rc}}^{*}\mathbf{H}_{\mathrm{AR}}\mathbf{w}_{\mathrm{b}}\right|^2 \leq f(R^{*}). \tag{25c}$$

Here, it is easy to verify that

$$\mathbf{h}_{\mathrm{rb}}^{H}\boldsymbol{\Theta}_{\mathrm{r}}\mathbf{H}_{\mathrm{AR}} = \boldsymbol{\vartheta}_{\mathrm{r}}^{T}\mathbf{H}_{\mathrm{rb}}^{*}\mathbf{H}_{\mathrm{AR}}, \tag{26}$$

$$\mathbf{h}_{\mathrm{rc}}^{H}\boldsymbol{\Theta}_{\mathrm{t}}\mathbf{H}_{\mathrm{AR}} = \boldsymbol{\vartheta}_{\mathrm{t}}^{T}\mathbf{H}_{\mathrm{rc}}^{*}\mathbf{H}_{\mathrm{AR}}, \tag{27}$$

by vertorizing the diagonal matrixes $\boldsymbol{\Theta}_{\mathrm{r}}$ and $\boldsymbol{\Theta}_{\mathrm{t}}$ as $\boldsymbol{\vartheta}_{\mathrm{r}} = \mathrm{diag}(\boldsymbol{\Theta}_{\mathrm{r}})$ and $\boldsymbol{\vartheta}_{\mathrm{t}} = \mathrm{diag}(\boldsymbol{\Theta}_{\mathrm{t}})$, and diagonalizing the vectors $\mathbf{h}_{\mathrm{rb}}$ and $\mathbf{h}_{\mathrm{rc}}$ as $\mathbf{H}_{\mathrm{rb}} = \mathrm{Diag}(\mathbf{h}_{\mathrm{rb}})$ and $\mathbf{H}_{\mathrm{rc}} = \mathrm{Diag}(\mathbf{h}_{\mathrm{rc}})$. Note that the optimization problem (25) is still non-convex due to the non-concave objective function and the non-convex covert communication constraint (25b) recalling that $\varpi_{\mathrm{b}} = \mathbf{w}_{\mathrm{b}}^{H}\mathbf{w}_{\mathrm{b}}$.

To effectively address the non-convex problem (25), we resort to the SDR method [36]. In particular, let $\mathbf{W}_{\mathrm{b}} = \mathbf{w}_{\mathrm{b}}\mathbf{w}_{\mathrm{b}}^{H}$, then problem (25) can be equivalently transformed as

$$\max_{\mathbf{W}_{\mathrm{b}}} \ \mathrm{Tr}(\mathbf{A}\mathbf{W}_{\mathrm{b}}),$$

$$\text{s.t. } \mathrm{Tr}(\mathbf{W}_{\mathrm{b}}) \leq P_{\max}-\|\mathbf{w}_{\mathrm{c}}\|_2^2, \tag{28a}$$

$$\varpi_{\mathrm{b}}\ln\left(1+\frac{P_{\mathrm{j}}^{\max}\lambda_{\mathrm{rw}}}{l_{\mathrm{rw}}\theta_{\mathrm{r}}(\varpi_{\mathrm{b}}+\varpi_{\mathrm{c}})}\right) \leq \frac{\epsilon P_{\mathrm{j}}^{\max}\lambda_{\mathrm{rw}}}{l_{\mathrm{rw}}\theta_{\mathrm{r}}}, \tag{28b}$$

$$\mathrm{Tr}(\mathbf{B}\mathbf{W}_{\mathrm{b}}) \leq f(R^{*}), \tag{28c}$$

$$\mathbf{W}_{\mathrm{b}} \succeq \mathbf{0}, \tag{28d}$$

$$\mathrm{rank}(\mathbf{W}_{\mathrm{b}}) = 1, \tag{28e}$$

where $\mathbf{A}=(\mathbf{H}_{\rm rb}^{*}\mathbf{H}_{\rm AR})^{H}\boldsymbol{\vartheta}_{\rm r}^{*}\boldsymbol{\vartheta}_{\rm r}^{T}(\mathbf{H}_{\rm rb}^{*}\mathbf{H}_{\rm AR}),\mathbf{B}=(\mathbf{H}_{\rm rc}^{*}\mathbf{H}_{\rm AR})^{H}\boldsymbol{\vartheta}_{\rm t}^{*}\boldsymbol{\vartheta}_{\rm t}^{T}(\mathbf{H}_{\rm rc}^{*}\mathbf{H}_{\rm AR})$. Although (28b) is still a non-convex constraint about $\mathbf{W}_{\rm b}$, we note that the left-side of (28b), denoted as $g(\varpi_{\rm b})$, is a concave function of $\varpi_{\rm b}=\mathrm{Tr}(\mathbf{W}_{\rm b})$. Thus, the first-order Taylor expansion of $g(\varpi_{\rm b})$ can be leveraged to replace it iteratively, which is an upper-bound linear approximation and generate a tighter convex substitute for (28b). Specifically, in the $(m+1)$-th iteration of the overall proposed algorithm ($m=0,1,\cdots$), the first-order Taylor expansion of $g(\varpi_{\rm b})$, with the given point $\varpi_{\rm b}^{(m)}$ obtained in the m-th iteration, can be expressed as

$$\widehat{g}(\varpi_{\rm b},\varpi_{\rm b}^{(m)})=\frac{\partial g}{\partial\varpi_{\rm b}}(\varpi_{\rm b}^{(m)})*(\varpi_{\rm b}-\varpi_{\rm b}^{(m)})+g(\varpi_{\rm b}^{(m)}),\quad(29)$$

where $\frac{\partial g}{\partial\varpi_{\rm b}}(\varpi_{\rm b}^{(m)})$ is given by (30) at the top of the next page.

Therefore, the problem (28) in the $(m+1)$-th iteration can be reformulated as

$$\begin{aligned}&\max_{\mathbf{W}_{\rm b}}\mathrm{Tr}(\mathbf{A}\mathbf{W}_{\rm b}),\\&\text{s.t. }\widehat{g}(\varpi_{\rm b},\varpi_{\rm b}^{(m)})\le\frac{\epsilon P_{\rm j}^{\max}\lambda_{\rm rw}}{l_{\rm rw}\theta_{\rm r}},&(31a)\\&\quad(28a),(28c),(28d),(28e).&(31b)\end{aligned}$$

Note that the remaining non-convexity of problem (31) is the non-convex rank-one constraint (28e). Here, we employ the SDR techniques to remove this constraint, and then problem (31) can be transformed into a standard convex semidefinite programming (SDP) problem which can be effectively solved by existing convex optimization solvers such as CVX [37]. As stated in [36], [38], there is a relatively high probability that the obtained optimal SDP solution cannot satisfy the rank-one constraint, i.e., $\mathrm{rank}(\mathbf{W}_{\rm b})\neq1$. Thus, additional steps are necessary to construct the rank-one solution from the obtained higher-rank solution, by the commonly used eigenvector approximation or Gaussian randomization methods [36], [39].

2) Active Beamforming Design for $\mathbf{w}_{\rm c}$*:* On the other hand, for given $\mathbf{w}_{\rm b}$, $\boldsymbol{\Theta}_{\rm r}$ and $\boldsymbol{\Theta}_{\rm t}$, the original optimization problem (24) can be simplified as

$$\begin{aligned}&\min_{\mathbf{w}_{\rm c}}\left|\boldsymbol{\vartheta}_{\rm r}^{T}\mathbf{H}_{\rm rb}^{*}\mathbf{H}_{\rm AR}\mathbf{w}_{\rm c}\right|^{2},\\&\text{s.t. }\|\mathbf{w}_{\rm c}\|_2^2\le P_{\max}-\|\mathbf{w}_{\rm b}\|_2^2,&(32a)\\&\quad\varpi_{\rm c}\le C(\epsilon),&(32b)\\&\quad\left|\boldsymbol{\vartheta}_{\rm t}^{T}\mathbf{H}_{\rm rc}^{*}\mathbf{H}_{\rm AR}\mathbf{w}_{\rm c}\right|^{2}\ge\hat{f}(R^{*}),&(32c)\end{aligned}$$

where $C(\epsilon)=\frac{P_{\rm j}^{\max}\lambda_{\rm rw}}{l_{\rm rw}\theta_{\rm r}\left(e^{\frac{\epsilon P_{\rm j}^{\max}\lambda_{\rm rw}}{l_{\rm rw}\theta_{\rm r}\varpi_{\rm b}}}-1\right)}-\varpi_{\rm b}$ and $\hat{f}(R^{*})=\left(2^{R^{*}}-1\right)\left(\left|\boldsymbol{\vartheta}_{\rm t}^{T}\mathbf{H}_{\rm rc}^{*}\mathbf{H}_{\rm AR}\mathbf{w}_{\rm b}\right|^{2}+\sigma^{*}+\sigma_{\rm c}^{2}\right),\varpi_{\rm c}=\mathbf{w}_{\rm c}^{H}\mathbf{w}_{\rm c}$.

$$\begin{aligned}&\min_{\mathbf{W}_{\rm c}}\mathrm{Tr}(\mathbf{A}\mathbf{W}_{\rm c}),\\&\text{s.t. }\mathrm{Tr}(\mathbf{W}_{\rm c})\le P_{\max}-\|\mathbf{w}_{\rm b}\|_2^2,&(33a)\\&\quad\mathrm{Tr}(\mathbf{W}_{\rm c})\le C(\epsilon),&(33b)\\&\quad\mathrm{Tr}(\mathbf{B}\mathbf{W}_{\rm c})\ge\hat{f}(R^{*}),&(33c)\\&\quad\mathrm{rank}(\mathbf{W}_{\rm c})=1,&(33d)\\&\quad\mathbf{W}_{\rm c}\succeq\mathbf{0}.&(33e)\end{aligned}$$

Similarly, we leverage the SDR method to handle the non-convex optimized problem (32). By defining $\mathbf{W}_{\rm c}=\mathbf{w}_{\rm c}\mathbf{w}_{\rm c}^{H}$, the problem can be transformed as (33). The problem (33) can be optimally solved by removing the rank-one constraint (33d), and then the rank-one solution can be constructed through the existing approaches.

C. Joint Passive Beamforming Design for STAR-RIS

After the active beamforming design, we can optimize the passive beamforming variables $\boldsymbol{\vartheta}_{\rm r}$ and $\boldsymbol{\vartheta}_{\rm t}$, with $\mathbf{w}_{\rm b}$ and $\mathbf{w}_{\rm c}$ being fixed as the obtained solutions in the previous subsection. On the basis of the original problem (24), the corresponding optimization problem for joint passive beamforming for STAR-RIS design can be expressed as

$$\begin{aligned}&\max_{\boldsymbol{\vartheta}_{\rm r},\boldsymbol{\vartheta}_{\rm t}}\gamma_{bb}(\boldsymbol{\vartheta}_{\rm r},\boldsymbol{\vartheta}_{\rm t}),\\&\text{s.t. }\frac{l_{\rm rw}\theta_{\rm r}\varpi_{\rm b}\ln\left(1+\frac{P_{\rm j}^{\max}\lambda_{\rm rw}}{l_{\rm rw}\theta_{\rm r}(\varpi_{\rm b}+\varpi_{\rm c})}\right)}{P_{\rm j}^{\max}\lambda_{\rm rw}}\le\epsilon,&(34a)\\&\quad\tilde{f}(\boldsymbol{\vartheta}_{\rm t})\ge0,&(34b)\\&\quad(24d),&(34c)\end{aligned}$$

where

$$\begin{aligned}\gamma_{bb}=&\frac{\left|\mathbf{h}_{\rm rb}^{H}\boldsymbol{\Theta}_{\rm r}\mathbf{H}_{\rm AR}\mathbf{w}_{\rm b}\right|^{2}}{\left|\mathbf{h}_{\rm rb}^{H}\boldsymbol{\Theta}_{\rm r}\mathbf{H}_{\rm AR}\mathbf{w}_{\rm c}\right|^{2}+\left|\mathbf{h}_{\rm rb}^{H}\boldsymbol{\Theta}_{\rm t}\mathbf{h}_{\rm rc}^{*}\right|^{2}P_{\rm j}^{\max}(1-\iota)+\sigma_{\rm b}^{2}}\\=&\frac{\boldsymbol{\vartheta}_{\rm r}^{T}\mathbf{C}\boldsymbol{\vartheta}_{\rm r}^{*}}{\boldsymbol{\vartheta}_{\rm r}^{T}\mathbf{D}\boldsymbol{\vartheta}_{\rm r}^{*}+\boldsymbol{\vartheta}_{\rm t}^{T}\mathbf{E}\boldsymbol{\vartheta}_{\rm t}^{*}P_{\rm j}^{\max}(1-\iota)+\sigma_{\rm b}^{2}},&(35)\\\mathbf{C}=&(\mathbf{H}_{\rm rb}^{*}\mathbf{H}_{\rm AR})\mathbf{w}_{\rm b}\mathbf{w}_{\rm b}^{H}(\mathbf{H}_{\rm rb}^{*}\mathbf{H}_{\rm AR})^{H},&(36)\\\mathbf{D}=&(\mathbf{H}_{\rm rb}^{*}\mathbf{H}_{\rm AR})\mathbf{w}_{\rm c}\mathbf{w}_{\rm c}^{H}(\mathbf{H}_{\rm rb}^{*}\mathbf{H}_{\rm AR})^{H},&(37)\\\mathbf{E}=&(\mathbf{H}_{\rm rb}^{*}\mathbf{H}_{\rm rc}^{*})(\mathbf{H}_{\rm rb}^{*}\mathbf{H}_{\rm rc}^{*})^{H},&(38)\\\tilde{f}(\boldsymbol{\vartheta}_{\rm t})=&\left|\boldsymbol{\vartheta}_{\rm t}\mathbf{H}_{\rm rc}^{*}\mathbf{H}_{\rm AR}\mathbf{w}_{\rm c}\right|^{2}-\\&(2^{R^{*}}-1)\left(\left|\boldsymbol{\vartheta}_{\rm t}^{T}\mathbf{H}_{\rm rc}^{*}\mathbf{H}_{\rm AR}\mathbf{w}_{\rm b}\right|^{2}+\sigma^{*}+\sigma_{\rm c}^{2}\right).&(39)\end{aligned}$$

We can find that the fractional objective function and the constraints (34a), (34b) in problem (34) are non-convex w.r.t. $\boldsymbol{\vartheta}_{\rm r}$ and $\boldsymbol{\vartheta}_{\rm t}$ recalling that $\theta_{\rm r}=\mathrm{diag}(\boldsymbol{\Theta}_{\rm r})^{H}\mathrm{diag}(\boldsymbol{\Theta}_{\rm r})=\boldsymbol{\vartheta}_{\rm r}^{H}\boldsymbol{\vartheta}_{\rm r}$ and $\lambda_{\rm rw}=\|\boldsymbol{\Theta}_{\rm t}\mathbf{h}_{\rm rc}^{*}\|_2^2=\|\boldsymbol{\vartheta}_{\rm t}\circ\mathbf{h}_{\rm rc}^{*}\|_2^2$, which makes this problem difficult to be solved directly.

In order to effectively solve the optimization problem (34), we first adopt the Dinkelbach's algorithm [40, Chapter3.2.1] and the SDR techniques to deal with the objective function. In the $(i+1)$-th iteration of the Dinkelbach's algorithm, the objective function can be transformed as

$$\begin{aligned}\widehat{\gamma}_{\rm bb}^{(i+1)}(\mathbf{Q}_{\rm r},\mathbf{Q}_{\rm t})=&\mathrm{Tr}(\mathbf{C}\mathbf{Q}_{\rm r})-\chi^{(i)}\Big(\mathrm{Tr}(\mathbf{D}\mathbf{Q}_{\rm r})\\&+\mathrm{Tr}(\mathbf{E}\mathbf{Q}_{\rm t})P_{\rm j}^{\max}(1-\iota)+\sigma_{\rm b}^{2}\Big),&(40)\end{aligned}$$

where $\mathbf{Q}_{\rm r}=\boldsymbol{\vartheta}_{\rm r}^{*}\boldsymbol{\vartheta}_{\rm r}^{T}$ and $\mathbf{Q}_{\rm t}=\boldsymbol{\vartheta}_{\rm t}^{*}\boldsymbol{\vartheta}_{\rm t}^{T}$ will be optimized to obtain the reflection and transmission coefficients of STAR-RIS. In addition, the parameter $\chi^{(i)}$ is updated with

$$\chi^{(i)}=\frac{\mathrm{Tr}(\mathbf{C}\mathbf{Q}_{\rm r}^{(i)})}{\left(\mathrm{Tr}(\mathbf{D}\mathbf{Q}_{\rm r}^{(i)})+\mathrm{Tr}(\mathbf{E}\mathbf{Q}_{\rm t}^{(i)})P_{\rm j}^{\max}(1-\iota)+\sigma_{\rm b}^{2}\right)},\quad(41)$$

$$\frac{\partial g}{\partial \varpi_{\rm b}}(\varpi_{\rm b}^{(m)}) = \ln\left(1+\frac{P_{\rm j}^{\max}\lambda_{\rm rw}}{l_{\rm rw}\theta_{\rm r}\left(\varpi_{\rm b}^{(m)}+\varpi_{\rm c}\right)}\right)+\frac{-\varpi_{\rm b}^{(m)}l_{\rm rw}\theta_{\rm r}P_{\rm j}^{\max}\lambda_{\rm rw}}{(\varpi_{\rm b}^{(m)}+\varpi_{\rm c})\left(l_{\rm rw}\theta_{\rm r}(\varpi_{\rm b}^{(m)}+\varpi_{\rm c})+P_{\rm j}^{\max}\lambda_{\rm rw}\right)}, \tag{30}$$

where $\mathbf{Q}_{\rm r}^{(i)}$ and $\mathbf{Q}_{\rm t}^{(i)}$ are the optimized solution through the i-th iteration. In this way, the objective function has been transformed into an affine function w.r.t. $\mathbf{Q}_{\rm r}$ and $\mathbf{Q}_{\rm t}$.

Next, we try to deal with the constraints in problem (34). For the left-side of covert communication constraint (34a), we find that it is a monotonically decreasing function of $\frac{\lambda_{\rm rw}}{\theta_{\rm r}}$. Thus, the constraint (34a) can be equivalently rewritten as $\frac{\lambda_{\rm rw}}{\theta_{\rm r}} \geq \varphi(\epsilon)$, where $\varphi(\epsilon)$ can be numerically obtained by employing the bisection search method. According to the expressions of $\lambda_{\rm rw}$ and $\theta_{\rm r}$, they can be reformulated as $\lambda_{\rm rw} = \|\boldsymbol{\vartheta}_{\rm t}\circ\mathbf{h}_{\rm rc}^*\|_2^2 = \boldsymbol{\beta}_{\rm t}^T(\mathbf{h}_{\rm rc}\circ\mathbf{h}_{\rm rc}^*)$, $\theta_{\rm r} = \boldsymbol{\beta}_{\rm r}^T\mathbf{I}_{N\times 1}$, where $\boldsymbol{\beta}_{\rm r} = \left[\beta_{\rm r}^1,\cdots,\beta_{\rm r}^N\right]^T$, $\boldsymbol{\beta}_{\rm t} = \left[\beta_{\rm t}^1,\cdots,\beta_{\rm t}^N\right]^T$.

Similarly, by using SDR techniques, the equivalent form of $\tilde{f}(\boldsymbol{\vartheta}_{\rm t})$ in constraint (34b) w.r.t. $\mathbf{Q}_{\rm t}$ can be derived as

$$\tilde{f}(\mathbf{Q}_{\rm t}) = {\rm Tr}(\mathbf{F}\mathbf{Q}_{\rm t}) - \left(2^{R^*}-1\right)\left({\rm Tr}(\mathbf{G}\mathbf{Q}_{\rm t})+\sigma^*+\sigma_{\rm c}^2\right), \tag{42}$$

where we have $\mathbf{F} = (\mathbf{H}_{\rm rc}^*\mathbf{H}_{\rm AR})\mathbf{w}_{\rm c}\mathbf{w}_{\rm c}^H(\mathbf{H}_{\rm rc}^*\mathbf{H}_{\rm AR})^H, \mathbf{G} = (\mathbf{H}_{\rm rc}^*\mathbf{H}_{\rm AR})\mathbf{w}_{\rm b}\mathbf{w}_{\rm b}^H(\mathbf{H}_{\rm rc}^*\mathbf{H}_{\rm AR})^H$.

Based on the above analysis, the optimization problem (34) in the $(i+1)$-th iteration can be formulated as

$$\begin{aligned}
\max_{\mathbf{Q}_{\rm r},\mathbf{Q}_{\rm t},\boldsymbol{\beta}_{\rm r},\boldsymbol{\beta}_{\rm t}}\ & \widehat{\gamma}_{\rm bb}^{(i+1)}(\mathbf{Q}_{\rm r},\mathbf{Q}_{\rm t}),\\
{\rm s.t.}\ \ & \frac{\boldsymbol{\beta}_{\rm t}^T(\mathbf{h}_{\rm rc}\circ\mathbf{h}_{\rm rc}^*)}{\boldsymbol{\beta}_{\rm r}^T\mathbf{I}_{N\times 1}} \geq \varphi(\epsilon), && \text{(43a)}\\
& \tilde{f}(\mathbf{Q}_{\rm t}) \geq 0, && \text{(43b)}\\
& \beta_{\rm r}^n+\beta_{\rm t}^n = 1, && \text{(43c)}\\
& {\rm diag}(\mathbf{Q}_{\rm r}) = \boldsymbol{\beta}_{\rm r}, {\rm diag}(\mathbf{Q}_{\rm t}) = \boldsymbol{\beta}_{\rm t}, && \text{(43d)}\\
& {\rm rank}(\mathbf{Q}_{\rm r}) = 1, {\rm rank}(\mathbf{Q}_{\rm t}) = 1, && \text{(43e)}\\
& \mathbf{Q}_{\rm r}\succeq 0, \mathbf{Q}_{\rm t}\succeq 0. && \text{(43f)}
\end{aligned}$$

However, problem (43) is still a non-convex optimization problem due to the two rank-one constraints in (43e). Due to the dependence of $\mathbf{Q}_{\rm r}$ and $\mathbf{Q}_{\rm t}$, it is difficult to re-construct the rank-one solution if we remove the rank-one constraints directly. To handle this issue, we equivalently rewrite the rank-one constraints as [18]

$${\rm rank}(\mathbf{Q}_{\rm r}) = 1 \Leftrightarrow \eta_{\rm r} \triangleq {\rm Tr}(\mathbf{Q}_{\rm r}) - \|\mathbf{Q}_{\rm r}\|_2 = 0, \tag{44}$$

$${\rm rank}(\mathbf{Q}_{\rm t}) = 1 \Leftrightarrow \eta_{\rm t} \triangleq {\rm Tr}(\mathbf{Q}_{\rm t}) - \|\mathbf{Q}_{\rm t}\|_2 = 0, \tag{45}$$

where $\|\mathbf{Q}\|_2$ is the spectral norm which is a convex versus $\mathbf{Q}$. Note that for any positive semidefinite matrix $\mathbf{Q}\succeq 0$, ${\rm Tr}(\mathbf{Q})-\|\mathbf{Q}\|_2 \geq 0$ always holds and the equality is satisfied if and only if ${\rm rank}(\mathbf{Q}) = 1$. Based on the non-negative feature of $\eta_{\rm r}$ and $\eta_{\rm t}$, we add them into the objective function of problem (43) as penalty terms for the rank-one constraints. Hence, the optimization problem can be re-expressed as

$$\begin{aligned}
\max_{\mathbf{Q}_{\rm r},\mathbf{Q}_{\rm t},\boldsymbol{\beta}_{\rm r},\boldsymbol{\beta}_{\rm t}}\ & \widehat{\gamma}_{\rm bb}^{(i+1)}(\mathbf{Q}_{\rm r},\mathbf{Q}_{\rm t}) - \rho_1\eta_{\rm r} - \rho_2\eta_{\rm t},\\
{\rm s.t.}\ \ & \frac{\boldsymbol{\beta}_{\rm t}^T(\mathbf{h}_{\rm rc}\circ\mathbf{h}_{\rm rc}^*)}{\boldsymbol{\beta}_{\rm r}^T\mathbf{I}_{N\times 1}} \geq \varphi(\epsilon), && \text{(46a)}\\
& \text{(43b), (43c), (43d), (43f)}, && \text{(46b)}
\end{aligned}$$

where ρ_1, ρ_2 >0 are the introduced penalty coefficients. Now, the optimization problem (46) is still non-convex because of the non-convexity of the penalty terms $\eta_{\rm r}$ and $\eta_{\rm t}$. By replacing the convex spectral norms in $\eta_{\rm r}$ and $\eta_{\rm t}$ with their linear lower-bound, i.e., first-order Taylor expansions, we can obtain the upper-bound linear approximations for $\eta_{\rm r}$ and $\eta_{\rm t}$ as

$$\begin{aligned}
\eta_{\rm r} &\leq {\rm Tr}(\mathbf{Q}_{\rm r}) - \left(\|\mathbf{Q}_{\rm r}^{(i)}\|_2 + {\rm Tr}\left(\mathbf{q}_{\rm r}^{(i)}(\mathbf{q}_{\rm r}^{(i)})^H(\mathbf{Q}_{\rm r}-\mathbf{Q}_{\rm r}^{(i)})\right)\right)\\
&= \widehat{\eta}_{\rm r}(\mathbf{Q}_{\rm r}), && \text{(47)}\\
\eta_{\rm t} &\leq {\rm Tr}(\mathbf{Q}_{\rm t}) - \left(\|\mathbf{Q}_{\rm t}^{(i)}\|_2 + {\rm Tr}\left(\mathbf{q}_{\rm t}^{(i)}(\mathbf{q}_{\rm t}^{(i)})^H(\mathbf{Q}_{\rm t}-\mathbf{Q}_{\rm t}^{(i)})\right)\right)\\
&= \widehat{\eta}_{\rm t}(\mathbf{Q}_{\rm t}), && \text{(48)}
\end{aligned}$$

where $\mathbf{q}_{\rm r}^{(i)}$ and $\mathbf{q}_{\rm t}^{(i)}$ are the eigenvectors corresponding to the largest eigenvalues of $\mathbf{Q}_{\rm r}^{(i)}$ and $\mathbf{Q}_{\rm t}^{(i)}$. Hence, $\widehat{\gamma}_{\rm bb}^{(i+1)}(\mathbf{Q}_{\rm r},\mathbf{Q}_{\rm t}) - \rho_1\widehat{\eta}_{\rm r}(\mathbf{Q}_{\rm r}) - \rho_2\widehat{\eta}_{\rm t}(\mathbf{Q}_{\rm t})$ is a linear lower-bound of the objective function in problem (46), which will be further utilized as the objective function in problem (46) to obtain the solution of $\mathbf{Q}_{\rm r}$ and $\mathbf{Q}_{\rm t}$ in the $(i+1)$-th iteration, denoted $\mathbf{Q}_{\rm r}^{(i+1)}$ and $\mathbf{Q}_{\rm t}^{(i+1)}$.

In conclusion, to solve the optimization problem (34) for joint passive beamforming design, we propose a two-tier iterative algorithm as summarized in Algorithm 1, where the outer loop is for updating the penalty coefficients ρ_1, ρ_2 through the penalty method [41], and the inner loop is for updating $\mathbf{Q}_{\rm r}$, $\mathbf{Q}_{\rm t}$ through the Dinkelbach's algorithm [40], [42]. Here, $\omega > 1$ is the scaling factor of the penalty coefficient. Also, $v_1 > 0$ denotes the penalty violation and $v_2 > 0$ indicates the gap of the objective functions between two adjacent iterations of the Dinkelbach's algorithm, which are given by

$$v_1 = \max\{\eta_{\rm r},\eta_{\rm t}\}, \tag{49}$$

$$v_2 = \left|\widehat{\gamma}_{\rm bb}^{(i+1)}(\mathbf{Q}_{\rm r}^{(i+1)},\mathbf{Q}_{\rm t}^{(i+1)}) - \widehat{\gamma}_{\rm bb}^{(i)}(\mathbf{Q}_{\rm r}^{(i)},\mathbf{Q}_{\rm t}^{(i)})\right|. \tag{50}$$

D. Proposed Optimization Algorithm & Analysis on Complexity and Convergence

Algorithm 2 concludes the overall processes for solving the original optimization problem (24) for STAR-RIS-assisted covert communications, which is an alternating optimization algorithm to solve three subproblems alternatively as detailed in Section IV. Here, v>0 represents the gap of objective values between two adjacent iterations, and the algorithm converges when v is below a predefined accuracy threshold ε.

Algorithm 1: Proposed Iterative Algorithm for Problem (34) on Joint Passive Beamforming Design of STAR-RIS

1: Initialize feasible point ($\mathbf{Q}_{\mathrm{r}}^{(0)}$, $\mathbf{Q}_{\mathrm{t}}^{(0)}$), penalty coefficients ($\rho_1^{(0)}$, $\rho_2^{(0)}$), and calculate v_1; Define accuracy tolerance thresholds ε_1, ε_2; Set iteration index $l = 0$ for outer loop.
2: **While** $v_1 > \varepsilon_1$ or $l = 0$ **do**
3: Initialize $\chi^{(0)}$ and set $i = 0$ for inner loop.
4: **While** $v_2 > \varepsilon_2$ or $i = 0$ **do**
5: Solve the problem (46) with given ($\mathbf{Q}_{\mathrm{r}}^{(i)}$, $\mathbf{Q}_{\mathrm{t}}^{(i)}$). Update the ($\mathbf{Q}_{\mathrm{r}}^{(i+1)}$, $\mathbf{Q}_{\mathrm{t}}^{(i+1)}$) with obtained solution.
6: Calculate v_2, $\chi^{(i+1)}$ based on the obtained solution, and let $i = i + 1$.
7: **end while**
8: Calculate v_1; Update $\rho_1^{(l+1)} = \omega\rho_1^{(l)}$, $\rho_2^{(l+1)} = \omega\rho_2^{(l)}$; Let ($\mathbf{Q}_{\mathrm{r}}^{(0)}$, $\mathbf{Q}_{\mathrm{t}}^{(0)}$)=($\mathbf{Q}_{\mathrm{r}}^{(i)}$, $\mathbf{Q}_{\mathrm{t}}^{(i)}$) and $l = l + 1$.
9: **end while**
10: Calculate the $\mathbf{\Theta}_{\mathrm{r}}$, $\mathbf{\Theta}_{\mathrm{t}}$ through the obtained $\mathbf{Q}_{\mathrm{r}}$, $\mathbf{Q}_{\mathrm{t}}$.

Algorithm 2: Proposed Alternating Algorithm for STAR-RIS-assisted Covert Communications Problem (24)

1: Initialize feasible point ($\mathbf{w}_{\mathrm{b}}^{(0)}$, $\mathbf{w}_{\mathrm{c}}^{(0)}$, $\mathbf{\Theta}_{\mathrm{r}}^{(0)}$, $\mathbf{\Theta}_{\mathrm{t}}^{(0)}$); Define the tolerance accuracy ε; Set iteration index $m = 0$.
2: **While** $v > \varepsilon$ or $m = 0$ **do**
3: Solve the relaxed version of the subproblem (28) with SDR method and use Gaussian randomization method to construct the rank-one solution, then update the $\mathbf{w}_{\mathrm{b}}$.
4: Similarly, solve the relaxed version of the subproblem (33) with SDR method and update the $\mathbf{w}_{\mathrm{c}}$.
5: Solve subproblem (34) with Algorithm 1 and update the $\mathbf{\Theta}_{\mathrm{r}}$ and $\mathbf{\Theta}_{\mathrm{t}}$.
6: Calculate the objective value $R_{\mathrm{bb}}^{(m+1)}$ value and update $v = \left|R_{\mathrm{bb}}^{(m+1)} - R_{\mathrm{bb}}^{(m)}\right|^2$; Let $m = m + 1$.
7: **end while**

Next, we give the analysis on the computational complexity of the proposed Algorithm 2. Specifically, the main complexity comes from solving the standard SDP subproblems[4]. In terms of active beamforming design for Alice, the computational complexity on solving the subproblems (28) and (33) can be respectively calculated as $\mathcal{O}(M^{3.5})$ and $\mathcal{O}(M^{3.5})$. For Algorithm 1 on joint passive beamformeing design of STAR-RIS, the main complexity lies on solving the subproblem (46) which is dominated by $\mathcal{O}(2N^{3.5})$. Besides, we use the bisection search method to find $\varphi(\epsilon)$ with the complexity of $\mathcal{O}\big(\log_2(\frac{s_0}{\epsilon_{\mathrm{t}}})\big)$ where s_0 and ϵ_{t} denote the length of the initial search interval and the accuracy tolerance, respectively. Hence, the computational complexity of Algorithm 1 is $\mathcal{O}\big(I_2 I_3(2N^{3.5}) + \log_2(\frac{s_0}{\epsilon_{\mathrm{t}}})\big)$, where I_2 and I_3 are the number of outer and inner iterations. Based on the above analysis, the computational complexity of the proposed alternating Algorithm 2 for solving the original covert communication problem (24) is dominated by $\mathcal{O}\big(I_1\big(2M^{3.5}+\log_2(\frac{s_0}{\epsilon_{\mathrm{t}}})+I_2 I_3(2N^{3.5})\big)\big)$, where I_1 denotes the total iteration number of the proposed algorithm. The complexity is mainly determined by the number of antennas at Alice (M) and elements at STAR-RIS (N).

It is easy to verify that the convergence of the alternating Algorithm 2. For each iteration of Algorithm 2, we can always find a solution not worse than that of the previous iteration, considering the SDR and the Dinkelbach's algorithm leveraged for solving the subproblems. Hence, the the objective function of problem (24) monotonically non-decreases w.r.t. the iteration index, and the algorithm finally converges subject to the transmit power limitation (24a).

[4]For convex problems, we assume that the Interior point method is adopted and then calculate the computational complexity accordingly [43].

V. Simulation Results

In this section, we show the simulation results to verify the effectiveness of the proposed STAR-RIS-assisted covert communication scheme implemented by optimization Algorithm 2. Specifically, all the simulation results are averaged over 1000 independent channels realizations. The basic simulation parameters are listed in Table I unless specified otherwise.

TABLE I
Parameters Setting

Parameters	Symbol and Value
Distance between RIS and Alice/Bob/Carol	$d_{\mathrm{AR}} = 500$ m, $d_{\mathrm{rb}} = 80$ m, $d_{\mathrm{rc}} = 150$ m
Reference power gain at a distance of one meter (m)	$\rho_0 = -20$dB
Path-loss exponent	$\alpha = 2.6$
Noise power	$\sigma_{\mathrm{b}}^2 = -140$dB, $\sigma_{\mathrm{c}}^2 = -140$dB,
Self-interference cancellation coefficient	$\phi = -160$dB [44]
Accuracy tolerance parameters	$\varepsilon = 10^{-4}$, $\varepsilon_1 = 10^{-8}$, $\varepsilon_2 = 10^{-8}$
Initial penalty coefficients	$\rho_1^{(0)} = 10^{-5}$, $\rho_2^{(0)} = 10^{-5}$
Scaling factor	$\omega = 10$

To highlight the advantage of covert communication aided by STAR-RIS, we consider a baseline scheme which employs two adjacent conventional RISs to replace STAR-RIS where one is the reflection-only RIS and the other one is transmission-only RIS. The number of elements in these two RISs is selected as $N/2$ so as to achieve a fair comparison. We call this baseline scheme as "RIS-aided scheme". In addition, to further validate the effectiveness of the proposed algorithm, a general optimization method called GCMMA is adopted as a comparison algorithm to solve the problem [45], [46], which requires lower computational complexity and is capable of converging to the Karush-Kuhn-Tucker (KKT) solution.

Fig. 2 illustrates the convergence performance of the proposed algorithm in different cases with varying parameter settings. In particular, for different QoS requirements with R^* = 4 and R^* = 5 in two sub-figures, we consider six cases with varying M, N, P_{tmax} and R^*. The achieved results indicate that the covet rates monotonically non-decrease versus the number of iterations, and the proposed algorithm can almost converge within 10 iterations. In addition, we note that less iterations are needed for algorithm convergence if a more relaxed QoS constraint with smaller R^* is required.

In Fig. 3, we show the performance of average covert rate versus the transmitted power $P_{\max}$ at Alice, considering different QoS (R^*) and covert (ϵ) requirements. Specifically, we

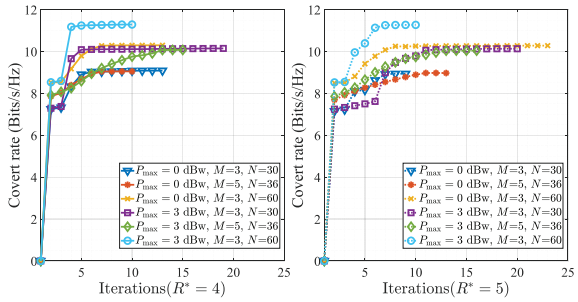


Fig. 2. Covert rate versus iterations with $P_{\mathrm{j}}^{\max}=0$ dBw, $\iota=0.1$, $\kappa=0.1$, $\epsilon=0.1$, and different M, N, $P_{\max}$ and R^*.

can find that the achievable covert rates for all schemes in all scenarios gradually increase with the growth of $P_{\max}$ before it reaching 6 dBw. And then the covert rates approach saturation when $P_{\max}$ further increases due to limitations of system settings. It is obvious that a significant performance improvement can be achieved by the proposed optimization scheme in comparison with the GCMMA method, which clearly validates the effectiveness of the proposed algorithm. Compared with the RIS-aided baseline scheme, we can find that the proposed STAR-RIS-assisted scheme possesses a strong superiority in enhancing the system covet performance, and the advantage may further expanded when with relaxed QoS requirement but limited transmit power budget (smaller R^* and $P_{\max}$). In addition, we can observe that a lower covert rate is achieved if the QoS or the covertness constraint becomes tighter, i.e., from $R^*=4$ to $R^*=5$, or from $\epsilon^*=0.2$ to $\epsilon^*=0.1$, which coincides with our intuition. Compared with the RIS-aided scheme, the performance degradation of proposed STAR-RIS assisted scheme with a moderate $P_{\max}$ is much less serious.

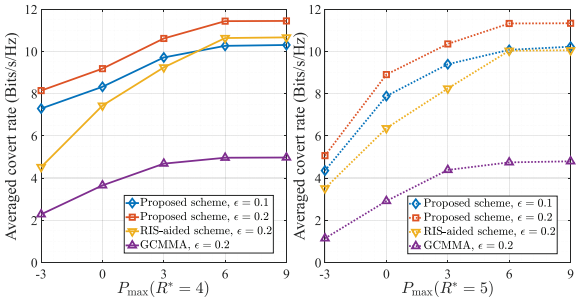


Fig. 3. Average covert rate versus the maximum transmit power $P_{\max}$ of Alice with $P_{\mathrm{j}}^{\max}=0$ dBw, $M=3$, $N=30$, $\iota=0.1$, $\kappa=0.1$, and different ϵ and R^*.

Next in Fig. 4, we investigate the influence of the covert requirement, i.e., ϵ, on the performance of average covert rate, considering different $P_{\max}$ and QoS requirements. In particular, $P_{\max}=6$ dBw is selected to operate the RIS-aided baseline scheme for an evident comparison, and obvious performance improvement can be achieved by the proposed scheme. Even if a lower transmitted power budget, i.e., 3 dBw, is utilized, the proposed scheme can still obtain better performance. This is because the STAR-RIS possesses a more flexible regulation ability compared with the conventional RIS, which can adjust the element phases and amplitudes for both reflection and transmission. It can be seen that the proposed scheme highly outperforms the GCMMA algorithm and the performance gap enlarges with the increase of ϵ, indicating that the proposed scheme can achieve a much better solution than the KKT solution converged by the GAMMA method.

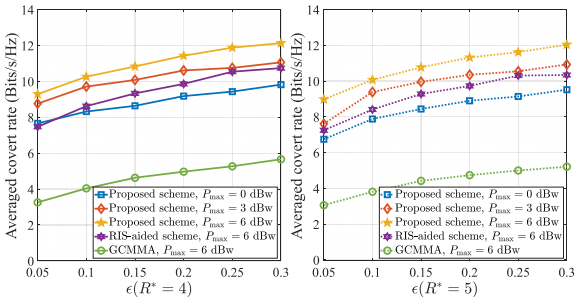


Fig. 4. Average covert rate versus the covert requirement ϵ with $P_{\mathrm{j}}^{\max}=0$ dBw, $M=3$, $N=30$, $\iota=0.1$, $\kappa=0.1$, and different $P_{\max}$ and R^*.

We present the variation curves of average covert rate w.r.t. the number of elements on STAR-RIS (N) in Fig. 5, under different transmit power $P_{\max}$ and QoS constraints R^*. It can be observed that the average covert rates of all the schemes grow with N, since the increased elements can provide higher freedom degree for reconfiguration of propagation environment. However, the increasing rates gradually decrease with the growth of N, and this may due to the limitations of other system settings. Similarly, $P_{\max}=6$ dBw is chosen to implement the two benchmark schemes, i.e., the RIS-aided and the GCMMA schemes. The obtained results further verify the advantages of the proposed STAR-RIS-assisted scheme which can achieve even better performance than the benchmark schemes in the scenario with a much smaller transmit power budget ($P_{\max}=3$ dBw). In addition, the performance enhancement for the proposed scheme is more obvious as the number of elements becomes larger.

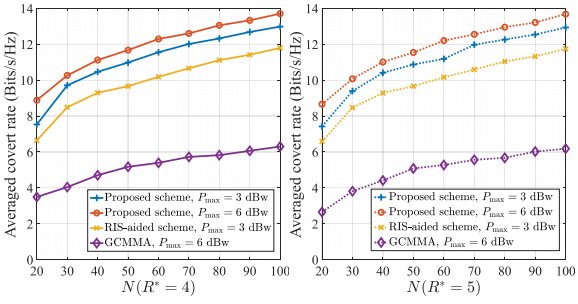


Fig. 5. Average covert rate versus the number of elements on STAR-RIS with $P_{\mathrm{j}}^{\max}=0$ dBw, $M=3$, $\epsilon=0.1$, $\iota=0.1$, $\kappa=0.1$, and different $P_{\max}$ and R^*.

In Fig. 6, we explore the effects of the number of antennas equipped at the transmitter, i.e., M, on the available covert rate, considering different transmitted power $P_{\max}$, number of elements in STAR-RIS (N), and QoS requirements (R^*). In particular, with the growth of M, the average covert rates of all schemes gradually increase, but the increasing rates have downward trends. Besides, we can observe that higher transmitted power or more elements at STAR-RIS contribute

to breaking the performance bottleneck imposed by channel characteristics and the number of antennas at Alice. Similarly, evident performance gaps exist between the proposed scheme and the RIS-aided scheme even for the proposed scheme with a lower $P_{\max}$ or fewer N. Compare with the GCMMA method, the proposed algorithm can still achieve better solutions in both scenarios with $R^* = 4$ and $R^* = 5$.

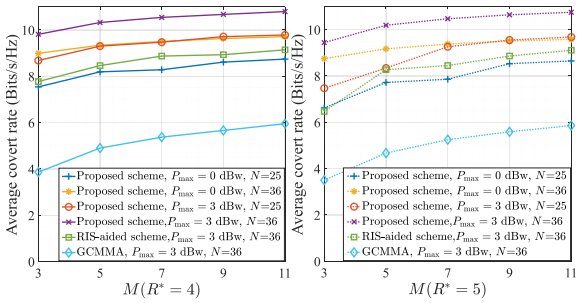


Fig. 6. Averaged covert rate versus the number of antennas at Alice with $P_{\mathrm{j}}^{\max} = 0$ dBw, $\epsilon = 0.1$, $\iota = 0.1$, $\kappa = 0.1$, and different $P_{\max}$, N and R^*.

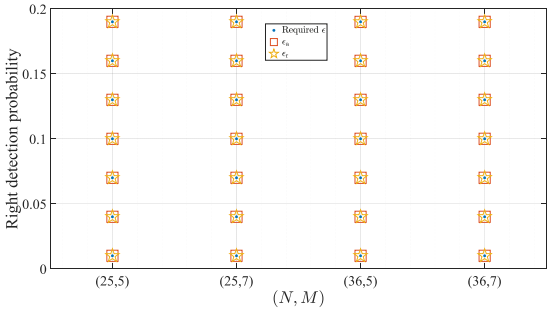


Fig. 7. Verifying the reasonability of choosing the lower bound of the minimum average DEP i.e., $\hat{P}_{\mathrm{ea}}^*$, with $P_{\mathrm{j}}^{\max} = 0$ dBw, $P_{\max} = 3$ dBw, $\iota = 0.1$, $\kappa = 0.1$, and different N and M.

Fig. 7 verifies the reasonability of adopting the lower bound of the minimum average DEP, i.e., $\hat{P}_{\mathrm{ea}}^*$, to replace itself $\overline{P}_{\mathrm{ea}}^*$, under different covertness requirements (ϵ), number of antennas at Alice (M), and elements equipped at RIS (N). In particular, we utilize the solution obtained by choosing the lower bound $\hat{P}_{\mathrm{ea}}^*$ to compute the accurate maximum average correct detection probability at Willie, which denotes ϵ_{r}. Note that $\epsilon_{\mathrm{a}} = 1 - \hat{\mathrm{P}}_{\mathrm{ea}}^*$ represents the upper bound of ϵ_{r}. According to the obtained results, we can find that the lower bound in (15) is tight because ϵ_{r} and ϵ_{a} are almost identical in all considered scenarios with different covert requirements and different communication system configurations.

VI. Conclusions

In this work, we initially investigate the application potentials of STAR-RIS in covert communications. In particular, the closed-form expression of the minimum DEP about the STAR-RIS-aided covert communication system is analytically derived. And then we jointly design the active and passive beamforming at the BS and STAR-RIS, to maximize the covert rate taking into account of the minimum DEP of Willie and the communication outage probability experienced at Bob and Carol. Due to the strong coupling between active and passive beamforming variables and the characteristic amplitude constraint introduced by the STAR-RIS, the proposed optimization problem is a non-convex problem. To effectively solve this covert communication problem, we elaborately design an alternating algorithm based on the SDR method and Dinkelbach's algorithm. Simulation results demonstrate that the STAR-RIS-assisted covert communication scheme highly outperforms the conventional RIS-aided scheme. In addition, the proposed iterative algorithm can effectively solve the formulated problem with guaranteed convergence.

Appendix A
Proof of Theorem 1

To facilitate the analysis, we denote $\varrho_1 \triangleq \mathbf{h}_{\mathrm{rw}}^H \mathbf{\Theta}_{\mathrm{r}} \mathbf{H}_{\mathrm{AR}} \mathbf{w}_{\mathrm{c}}$, and it is easy to demonstrate that ϱ_1 follows the complex Gaussian distribution with mean zero and variance $\lambda = \left\|\mathbf{h}_{\mathrm{rw}}^H \mathbf{\Theta}_{\mathrm{r}}\right\|_2^2 \mathbf{w}_{\mathrm{c}}^H \mathbf{w}_{\mathrm{c}}$. Hence, the PDF of $|\varrho_1|^2$ can be written $f_{|\varrho_1|^2}(x) = \frac{e^{-\frac{x}{\lambda}}}{\lambda}$. In addition, we know that P_{j} follows the uniform distribution, and thus the analytical expression of the FA probability P_{FA} can be derived as (51), which is shown at the bottom of the next page. Similarly, we can derive the MD probability P_{MD} as (8).

Appendix B
Proof of Theorem 2

According to the analytical expression of DEP at Willie, i.e., P_{e} given in (9), we can see that P_{e} is a segment function based on the detection threshold τ_{dt} in three different ranges. Let us discuss the optimal detection τ_{dt}^* threshold in three ranges, respectively.

1) $\tau_{\mathrm{dt}} < \sigma_{\mathrm{w}}^2$*:* It is easy to note that $P_{\mathrm{e}} = 1$ when $\tau_{\mathrm{dt}} < \sigma_{\mathrm{w}}^2$, indicating that Willie is always unable to distinguish the existence of communications between Alice and Bob. Hence, there is no need to optimize τ_{dt} to minimize the DEP when τ_{dt} falls into this range.

2) $\sigma_{\mathrm{w}}^2 \leq \tau_{\mathrm{dt}} < \sigma_{\mathrm{w}}^2 + \gamma P_{\mathrm{j}}^{\max}$*:* The first-order derivative of P_{e} w.r.t. τ_{dt} in this range is given by $\frac{\partial P_{\mathrm{e}}}{\partial \tau_{\mathrm{dt}}} = \frac{e^{-\frac{\tau_{\mathrm{dt}}-\sigma_{\mathrm{w}}^2}{\lambda}} - e^{-\frac{\tau_{\mathrm{dt}}-\sigma_{\mathrm{w}}^2}{\tilde{\lambda}}}}{\gamma P_{\mathrm{j}}^{\max}}$, from which can find that $\frac{\partial P_{\mathrm{e}}}{\partial \tau_{\mathrm{dt}}} < 0$ always holds. Therefore, P_{e} decrease monotonically versus $\tau_{\mathrm{dt}} \in [\sigma_{\mathrm{w}}^2, \sigma_{\mathrm{w}}^2 + \gamma P_{\mathrm{j}}^{\max})$ and the optimal detection threshold is $\tau_{\mathrm{dt}}^* = \sigma_{\mathrm{w}}^2 + \gamma P_{\mathrm{j}}^{\max}$.

3) $\tau_{\mathrm{dt}} \geq \sigma_{\mathrm{w}}^2 + \gamma P_{\mathrm{j}}^{\max}$*:* We can further obtain the first-order derivative of P_{e} w.r.t. τ_{dt} in this range as

$$\frac{\partial P_{\mathrm{e}}}{\partial \tau_{\mathrm{dt}}} = \frac{e^{-\frac{\tau_{\mathrm{dt}}-\sigma_{\mathrm{w}}^2}{\tilde{\lambda}}}\left(e^{\frac{\gamma P_{\mathrm{j}}^{\max}}{\tilde{\lambda}}} - 1\right) + e^{-\frac{\tau_{\mathrm{dt}}-\sigma_{\mathrm{w}}^2}{\lambda}}\left(1 - e^{\frac{\gamma P_{\mathrm{j}}^{\max}}{\lambda}}\right)}{\gamma P_{\mathrm{j}}^{\max}}. \tag{52}$$

Let $\frac{\partial P_{\mathrm{e}}}{\partial \tau_{\mathrm{dt}}} = 0$, we can obtain the unique solution of τ_{dt} in this range, i.e., $\tau_{\mathrm{dt}} = \frac{\tilde{\lambda}\lambda}{\tilde{\lambda}-\lambda} \ln \Delta + \sigma_{\mathrm{w}}^2 \in [\sigma_{\mathrm{w}}^2 + \gamma P_{\mathrm{j}}^{\max}, +\infty)$, where $\Delta = \frac{e^{\frac{\gamma P_{\mathrm{j}}^{\max}}{\lambda}} - 1}{e^{\frac{\gamma P_{\mathrm{j}}^{\max}}{\tilde{\lambda}}} - 1}$. We can prove that P_{e} first decreases and then increases versus τ_{dt} in this range with $\frac{\tilde{\lambda}\lambda}{\tilde{\lambda}-\lambda} \ln \Delta + \sigma_{\mathrm{w}}^2$ as the inflection point. Hence, the optimal detection threshold for minimizing P_{e} is given as $\tau_{\mathrm{dt}}^* = \frac{\tilde{\lambda}\lambda}{\tilde{\lambda}-\lambda} \ln \Delta + \sigma_{\mathrm{w}}^2$.

Based on the above analysis, the optimal detection threshold $\tau_{\rm dt}^*$ can be finally expressed as

$$\tau_{\rm dt}^* = \begin{cases} \sigma_{\rm w}^2 + \gamma P_{\rm j}^{\max}, & \sigma_{\rm w}^2 \le \tau_{\rm dt} < \sigma_{\rm w}^2 + \gamma P_{\rm j}^{\max}, \\ \frac{\tilde{\lambda}\lambda}{\tilde{\lambda}-\lambda} \ln \Delta + \sigma_{\rm w}^2. & \tau_{\rm dt} \ge \gamma P_{\rm j}^{\max} + \sigma_{\rm w}^2. \end{cases} \tag{53}$$

We can verify that $P_{\rm e}$ in (9) is a continuous segment function at the segment points $\sigma_{\rm w}^2$ and $\sigma_{\rm w}^2+\gamma P_{\rm j}^{\max}$. Therefore, the optimal detection threshold in the overall defined region for minimizing the DEP $P_{\rm e}$ is $\tau_{\rm dt}^* = \frac{\tilde{\lambda}\lambda}{\tilde{\lambda}-\lambda} \ln \Delta + \sigma_{\rm w}^2$.

Appendix C
Proof of Theorem 3

When the required transmission rate between Alice and Bob is chosen as R_b, the communication outage probability at Bob under the randomness of the jamming power $P_{\rm j}$ can be calculated as

$$\begin{aligned} \delta_{\rm AB} &= \Pr\left(C_{\rm b} < R_{\rm b}\right) \\ &= \begin{cases} 1, & \Upsilon < 0, \\ \int_{\Upsilon}^{P_{\rm j}^{\max}} \frac{1}{P_{\rm j}^{\max}} dy, & 0 \le \Upsilon < P_{\rm j}^{\max}, \\ 0, & \Upsilon \ge P_{\rm j}^{\max}, \end{cases} \\ &= \begin{cases} 1, & \Upsilon < 0, \\ 1 - \frac{\Upsilon}{P_{\rm j}^{\max}}, & 0 \le \Upsilon < P_{\rm j}^{\max}, \\ 0, & \Upsilon \ge P_{\rm j}^{\max}, \end{cases} \end{aligned} \tag{54}$$

where

$$\Upsilon = \frac{\left|\mathbf{h}_{\rm rb}^H \mathbf{\Theta}_{\rm r} \mathbf{H}_{\rm AR} \mathbf{w}_{\rm b}\right|^2}{\left(2^{R_{\rm b}}-1\right)\left|\mathbf{h}_{\rm rb}^H \mathbf{\Theta}_{\rm t} \mathbf{h}_{\rm rc}^*\right|^2} - \frac{\left(\left|\mathbf{h}_{\rm rb}^H \mathbf{\Theta}_{\rm r} \mathbf{H}_{\rm AR} \mathbf{w}_{\rm c}\right|^2 + \sigma_{\rm b}^2\right)}{\left|\mathbf{h}_{\rm rb}^H \mathbf{\Theta}_{\rm t} \mathbf{h}_{\rm rc}^*\right|^2}. \tag{55}$$

Similarly, when the required transmission rate between Alice and Carol is chosen as R_c, the communication outage probability at Carol under the randomness of the jamming power $P_{\rm j}$ and the self-interference channel $h_{\rm cc}$ is derived as

$$\begin{aligned} \delta_{\rm AC} &= \Pr\left(C_{\rm c} < R_{\rm c}\right) \\ &= \begin{cases} \int_0^{P_{\rm j}^{\max}} \int_{\frac{\Gamma}{\phi y}}^{+\infty} e^{-x} \frac{1}{P_{\rm j}^{\max}} dx dy, & \Gamma \ge 0, \\ 1, & \Gamma < 0, \end{cases} \\ &= \begin{cases} e^{-\frac{\Gamma}{\phi P_{\rm j}^{\max}}} + \frac{\Gamma}{\phi P_{\rm j}^{\max}} \mathrm{Ei}\left(-\frac{\Gamma}{\phi P_{\rm j}^{\max}}\right), & \Gamma \ge 0, \\ 1, & \Gamma < 0, \end{cases} \end{aligned} \tag{56}$$

where

$$\Gamma = \frac{\left|\mathbf{h}_{\rm rc}^H \mathbf{\Theta}_{\rm t} \mathbf{H}_{\rm AR} \mathbf{w}_{\rm c}\right|^2}{\left(2^{R_{\rm c}}-1\right)} - \left(\left|\mathbf{h}_{\rm rc}^H \mathbf{\Theta}_{\rm t} \mathbf{H}_{\rm AR} \mathbf{w}_{\rm b}\right|^2 + \sigma_{\rm c}^2\right). \tag{57}$$

References

[1] A. Chorti, A. N. Barreto, S. Köpsell, M. Zoli, M. Chafii, P. Sehier, G. Fettweis, and H. V. Poor, "Context-aware security for 6G wireless: The role of physical layer security," *IEEE Commun. Standards Mag.*, vol. 6, no. 1, pp. 102–108, 2022.

[2] M. Cui, G. Zhang, and R. Zhang, "Secure wireless communication via intelligent reflecting surface," *IEEE Wireless Commun. Lett.*, vol. 8, no. 5, pp. 1410–1414, 2019.

[3] X. Hu, P. Mu, B. Wang, and Z. Li, "On the secrecy rate maximization with uncoordinated cooperative jamming by single-antenna helpers," *IEEE Trans. Veh. Technol.*, vol. 66, no. 5, pp. 4457–4462, 2017.

[4] 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, 2021.

[5] X. Chen, W. Sun, C. Xing, N. Zhao, Y. Chen, F. R. Yu, and A. Nallanathan, "Multi-antenna covert communication via full-duplex jamming against a warden with uncertain locations," *IEEE Trans. Wireless Commun.*, vol. 20, no. 8, pp. 5467–5480, 2021.

[6] S. Yan, X. Zhou, J. Hu, and S. V. Hanly, "Low probability of detection communication: Opportunities and challenges," *IEEE Wireless Commun.*, vol. 26, no. 5, pp. 19–25, 2019.

[7] B. A. Bash, D. Goeckel, and D. Towsley, "Limits of reliable communication with low probability of detection on AWGN channels," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1921–1930, 2013.

[8] D. Goeckel, B. Bash, S. Guha, and D. Towsley, "Covert communications when the warden does not know the background noise power," *IEEE Commun. Lett.*, vol. 20, no. 2, pp. 236–239, 2015.

[9] J. Wang, W. Tang, Q. Zhu, X. Li, H. Rao, and S. Li, "Covert communication with the help of relay and channel uncertainty," *IEEE Wireless Commun. Lett.*, vol. 8, no. 1, pp. 317–320, 2018.

[10] J. Hu, S. Yan, X. Zhou, F. Shu, and J. Li, "Covert wireless communications with channel inversion power control in rayleigh fading," *IEEE Trans. Veh. Technol.*, vol. 68, no. 12, pp. 12 135–12 149, 2019.

[11] L. Tao, W. Yang, S. Yan, D. Wu, X. Guan, and D. Chen, "Covert communication in downlink NOMA systems with random transmit power," *IEEE Wireless Commun. Lett.*, vol. 9, no. 11, pp. 2000–2004, 2020.

[12] K. Li, P. A. Kelly, and D. Goeckel, "Optimal power adaptation in covert communication with an uninformed jammer," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 3463–3473, 2020.

[13] K. Shahzad, X. Zhou, and S. Yan, "Covert communication in fading channels under channel uncertainty," in *Proc. IEEE Veh. Technol. Conf (VTC Spring)*, 2017, pp. 1–5.

[14] R. Ma, W. Yang, L. Tao, X. Lu, Z. Xiang, and J. Liu, "Covert communications with randomly distributed wardens in the finite blocklength regime," *IEEE Trans. Veh. Technol.*, vol. 71, no. 1, pp. 533–544, 2021.

[15] T.-X. Zheng, H.-M. Wang, D. W. K. Ng, and J. Yuan, "Multi-antenna covert communications in random wireless networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 3, pp. 1974–1987, 2019.

[16] K. Shahzad, X. Zhou and S. Yan, "Covert wireless communication in presence of a multi-antenna adversary and delay constraints," *IEEE Trans. Veh. Technol.*, vol. 68, no. 12, pp. 12 432–12 436, 2019.

$$P_{\rm FA} = \begin{cases} 1, & \tau_{\rm dt} < \sigma_{\rm w}^2, \\ \int_0^{\infty} \frac{e^{-\frac{x}{\lambda}}}{\lambda} dx - \int_0^{\tau_{\rm dt}-\sigma_{\rm w}^2} \int_0^{\tau_{\rm dt}-\sigma_{\rm w}^2-x} \frac{e^{-\frac{x}{\lambda}}}{\lambda} \frac{1}{\gamma P_{\rm j}^{\max}} dy dx, & \sigma_{\rm w}^2 \le \tau_{\rm dt} < \sigma_{\rm w}^2 + \gamma P_{\rm j}^{\max}, \\ \int_0^{\gamma P_{\rm j}^{\max}} \int_{\tau_{\rm dt}-\sigma_{\rm w}^2-y}^{\infty} \frac{e^{-\frac{x}{\lambda}}}{\lambda} \frac{1}{\gamma P_{\rm j}^{\max}} dx dy, & \tau_{\rm dt} \ge \gamma P_{\rm j}^{\max} + \sigma_{\rm w}^2, \end{cases}$$
$$= \begin{cases} 1, & \tau_{\rm dt} < \sigma_{\rm w}^2, \\ 1 - \frac{\left(\tau_{\rm dt}-\sigma_{\rm w}^2\right) + \lambda e^{-\frac{\tau_{\rm dt}-\sigma_{\rm w}^2}{\lambda}} - \lambda}{\gamma P_{\rm j}^{\max}}, & \sigma_{\rm w}^2 \le \tau_{\rm dt} < \sigma_{\rm w}^2 + \gamma P_{\rm j}^{\max}, \\ \frac{e^{-\frac{\tau_{\rm dt}-\sigma_{\rm w}^2}{\lambda}} \left(e^{\frac{\gamma P_{\rm j}^{\max}}{\lambda}} - 1\right) \lambda}{\gamma P_{\rm j}^{\max}}, & \tau_{\rm dt} \ge \gamma P_{\rm j}^{\max} + \sigma_{\rm w}^2, \end{cases} \tag{51}$$

[17] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, 2019.

[18] X. Hu, C. Masouros, and K.-K. Wong, "Reconfigurable intelligent surface aided mobile edge computing: From optimization-based to location-only learning-based solutions," *IEEE Trans. Commun.*, vol. 69, no. 6, pp. 3709–3725, 2021.

[19] X. Lu, E. Hossain, T. Shafique, S. Feng, H. Jiang, and D. Niyato, "Intelligent reflecting surface enabled covert communications in wireless networks," *IEEE Netw.*, vol. 34, no. 5, pp. 148–155, 2020.

[20] X. Zhou, S. Yan, Q. Wu, F. Shu, and D. W. K. Ng, "Intelligent reflecting surface (IRS)-aided covert wireless communications with delay constraint," *IEEE Trans. Wireless Commun.*, vol. 21, no. 1, pp. 532–547, 2021.

[21] X. Chen, T.-X. Zheng, L. Dong, M. Lin, and J. Yuan, "Enhancing MIMO covert communications via intelligent reflecting surface," *IEEE Wireless Commun. Lett.*, vol. 11, no. 1, pp. 33–37, 2021.

[22] C. Wang, Z. Li, J. Shi, and D. W. K. Ng, "Intelligent reflecting surface-assisted multi-antenna covert communications: Joint active and passive beamforming optimization," *IEEE Trans. Commun.*, vol. 69, no. 6, pp. 3984–4000, 2021.

[23] Y. Liu, X. Mu, J. Xu, R. Schober, Y. Hao, H. V. Poor, and L. Hanzo, "STAR: Simultaneous transmission and reflection for 360° coverage by intelligent surfaces," *IEEE Wireless Commun.*, vol. 28, no. 6, pp. 102–109, 2021.

[24] X. Mu, Y. Liu, L. Guo, J. Lin, and R. Schober, "Simultaneously transmitting and reflecting (STAR) RIS aided wireless communications," *IEEE Trans. Wireless Commun.*, vol. 21, no. 5, pp. 3083–3098, 2022.

[25] Y. Han, N. Li, Y. Liu, T. Zhang, and X. Tao, "Artificial noise aided secure NOMA communications in STAR-RIS networks," *IEEE Wireless Commun. Lett.*, 2022.

[26] Z. Zhang, J. Chen, Y. Liu, Q. Wu, B. He, and L. Yang, "On the secrecy design of STAR-RIS assisted uplink NOMA networks," *IEEE Trans. Wireless Commun.*, vol. 21, no. 12, pp. 11 207–11 221, 2022.

[27] C. Wu, C. You, Y. Liu, X. Gu, and Y. Cai, "Channel estimation for STAR-RIS-aided wireless communication," *IEEE Commun. Lett.*, vol. 26, no. 3, pp. 652–656, 2021.

[28] D. Kim, H. Lee, and D. Hong, "A survey of in-band full-duplex transmission: From the perspective of PHY and MAC layers," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2017–2046, 2015.

[29] A. Chaman, J. Wang, J. Sun, H. Hassanieh, and R. Roy Choudhury, "Ghostbuster: Detecting the presence of hidden eavesdroppers," in *Proc. 24th Annu. Int. Conf. Mobile Comput. Netw. MobiCom*, 2018, pp. 337–351.

[30] A. Chaman, "Detecting the presence of hidden wireless eavesdroppers," M.S. thesis, Dept. Elect. Comput. Eng, Univ. Illinois Urbana-Champaign, Champaign, IL, USA, Nov. 2018.

[31] X. Zhou, S. Yan, J. Hu, J. Sun, J. Li, and F. Shu, "Joint optimization of a UAV's trajectory and transmit power for covert communications," *IEEE Trans. Signal Process.*, vol. 67, no. 16, pp. 4276–4290, 2019.

[32] J. Evans and D. N. C. Tse, "Large system performance of linear multiuser receivers in multipath fading channels," *IEEE Trans. Inf. Theory.*, vol. 46, no. 6, pp. 2059–2078, 2000.

[33] Y. Wu, R. Schober, D. W. K. Ng, C. Xiao, and G. Caire, "Secure massive MIMO transmission with an active eavesdropper," *IEEE Trans. Inf. Theory.*, vol. 62, no. 7, pp. 3880–3900, 2016.

[34] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, 2020.

[35] Z. Ding, R. Schober, and H. V. Poor, "On the impact of phase shifting designs on IRS-NOMA," *IEEE Wireless Commun. Lett.*, vol. 9, no. 10, pp. 1596–1600, 2020.

[36] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, 2010.

[37] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," 2014.

[38] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5394–5409, 2019.

[39] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network: Joint active and passive beamforming design," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*. IEEE, 2018, pp. 1–6.

[40] A. Zappone, E. Jorswieck *et al.*, "Energy efficiency in wireless networks via fractional programming theory," *Found. Trends Commun. Inf. Theory.*, vol. 11, no. 3-4, pp. 185–396, 2015.

[41] J. Nocedal and S. J. Wright, *Numerical optimization*. New York, NY, USA: Springer, 2006.

[42] X. Hu, L. Wang, K. Wong, M. Tao, Y. Zhang, and Z. Zheng, "Edge and central cloud computing: A perfect pairing for high energy efficiency and low-latency," *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 1070–1083, 2020.

[43] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex optimization*. Cambridge, U.K.:Cambridge Univ. Press, 2004.

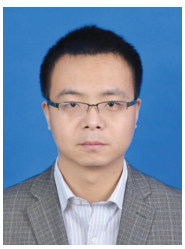
[44] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," in *Proc. ACM SIGCOMM*, 2013, pp. 375–386.

[45] K. Svanberg, "MMA and GCMMA-two methods for nonlinear optimization," *Tech. Rep. Optim. Theory.*, vol. 1, 2007.

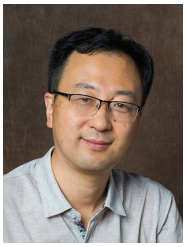
[46] K. Svanberg, "A class of globally convergent optimization methods based on conservative convex separable approximations," *SIAM J. Optim.*, vol. 12, no. 2, pp. 555–573, 2002.

Han Xiao received the M.Eng. degree in Vehicle Engineering from Dalian University of Technology, Dalian, China, in 2021. He is currently pursuing a Ph.D. degree at the School of Information and Communications Engineering, Xi'an Jiaotong University, Xi'an, China. His research interests include physical layer security, covert communications, mobile edge computing and reconfigurable intelligent surface.

Xiaoyan Hu (Member, IEEE) received the Ph.D. degree in Electronic and Electrical Engineering from University College London (UCL), London, U.K., in 2020. From 2019 to 2021, she was a Research Fellow with the Department of Electronic and Electrical Engineering, UCL, U.K. She is currently an Associate Professor with the School of Information and Communications Engineering, Xi'an Jiaotong University, Xi'an, China. Her research interests are in the areas of 5G&6G wireless communications, including topics such as edge computing, reconfigurable intelligent surface, UAV communications, ISAC, secure&covert communications, and learning-based communications. She has served as a Guest Editor for ELECTRONICS on Physical Layer Security and for China Communications Blue Ocean Forum on MAC and Networks. She has also been honored as an Exemplary Reviewer for IEEE COMMUNICATIONS LETTERS. Since 2020, she has been serving as the Assistant to the Editor-in-Chief of IEEE WIRELESS COMMUNICATIONS LETTERS.



Pengcheng Mu (Member, IEEE) received the B.S. degree in Information Engineering and the M.S. degree in Communication and Information system from Xi'an Jiaotong University, Xi'an, China, in 2003 and 2006, respectively, the engineering diploma from the Ecole Centrale Paris, Paris, France, in 2006, and the Ph.D. degree in Signal and Image Processing from INSA-Rennes, Rennes, France, in 2009. He is currently an Associate Professor with the School of Information and Communications Engineering, Xi'an Jiaotong University. His research interests include wireless communications and digital signal processing.



Wenjie Wang (Member, IEEE) received the B.S., M.S., and Ph.D. degrees in Information and Communications Engineering from Xi'an Jiaotong University, Xi'an, China, in 1993, 1998, and 2001, respectively. He was a Visiting Scholar with the Department of Electrical and Computer Engineering, University of Delaware, Newark, DE, USA, from 2009 to 2010. Now, he is a Professor and the Dean of the School of Information and Communications Engineering, Xi'an Jiaotong University. His research interests include ad-hoc networks, smart antennas, wireless communication, signal processing, artificial intelligence, and data analysis.

Tong-Xing Zheng (Member, IEEE) received the B.S. degree in information engineering and Ph.D. degree in information and communications engineering from Xi'an Jiaotong University, Xi'an, China, in 2010 and 2016, respectively. From 2017 to 2018, he was a Visiting Scholar with the School of Electrical Engineering and Telecommunications, University of New South Wales, Sydney, Australia. He is currently an Associate Professor with Xi'an Jiaotong University. His current research interests include 5G&6G wireless networks and key technologies, physical layer security, and covert communications. He has co-authored the book *Physical Layer Security in Random Cellular Networks* (Springer, 2016), one book chapter, and has authored or co-authored over 80 papers in telecommunications journals and conference proceedings. He was a recipient of the Excellent Doctoral Dissertation Award of Shaanxi Province in 2019, the First Prize of Science and Technology Award in Higher Institution of Shaanxi Province in 2019, and the Best Paper Award at the 2023 International Conference on Ubiquitous Communication (Ucom 2023). He was honored as an Exemplary Reviewer of IEEE TRANSACTIONS ON COMMUNICATIONS in 2017, 2018, and 2021, respectively. He was a Leading Guest Editor of FRONTIERS IN COMMUNICATIONS AND NETWORKS for the Special Issue on Covert Communications for Next-Generation Wireless Networks in 2021 and a Guest Editor of WIRELESS COMMUNICATIONS AND MOBILE COMPUTING for the Special Issue on Physical Layer Security for Internet of Things in 2018. He served as a Session Chair for the IEEE International Conference on Communications (ICC) 2022. He is currently serving as an Associate Editor of IET ELECTRONIC LETTERS and a Review Editor of FRONTIERS IN COMMUNICATIONS AND NETWORKS.

Kai-Kit Wong (Fellow, IEEE) received the BEng, the MPhil, and the PhD degrees, all in Electrical and Electronic Engineering, from the Hong Kong University of Science and Technology, Hong Kong, in 1996, 1998, and 2001, respectively. After graduation, he took up academic and research positions at the University of Hong Kong, Lucent Technologies, Bell-Labs, Holmdel, the Smart Antennas Research Group of Stanford University, and the University of Hull, UK. He is currently the Chair of Wireless Communications with the Department of Electronic and Electrical Engineering, University College London, UK.

His current research centers around 5G and beyond mobile communications, including topics such as massive MIMO, full-duplex communications, millimetre-wave communications, edge caching and fog networking, physical layer security, wireless power transfer and mobile computing, V2X communications, fluid antenna communications systems, and of course cognitive radios. He is a co-recipient of the 2013 IEEE Signal Processing Letters Best Paper Award and the 2000 IEEE VTS Japan Chapter Award at the IEEE Vehicular Technology Conference in Japan in 2000, and a few other international best paper awards.

He is Fellow of IEEE and IET and is also on the editorial board of several international journals. He has served as Senior Editor for IEEE COMMUNICATIONS LETTERS since 2012 and for IEEE WIRELESS COMMUNICATIONS LETTERS since 2016. He had also previously served as Associate Editor for IEEE SIGNAL PROCESSING LETTERS from 2009 to 2012 and Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2005 to 2011. He was also Guest Editor for IEEE JSAC SI on virtual MIMO in 2013 and on physical layer security for 5G in 2018. He has been the Editor-in-Chief of the IEEE WIRELESS COMMUNICATIONS LETTERS, since 2020.

Kun Yang (Fellow, IEEE) received his PhD from the Department of Electronic & Electrical Engineering of University College London (UCL), UK. He is currently a Chair Professor in the School of Computer Science & Electronic Engineering, University of Essex, UK, leading the Network Convergence Laboratory (NCL). His main research interests include wireless networks and communications, future Internet and edge computing. In particular he is interested in energy aspects of future communication systems such as 6G, promoting energy self-sustainability via both energy efficiency (green communications and networks) and energy harvesting (wireless charging). He has managed research projects funded by UK EPSRC, EU FP7/H2020, and industries. He has published 400+ papers and filed 30 patents. He serves on the editorial boards of a number of IEEE journals (e.g., IEEE TNSE, TVT, WCL). He is a Deputy Editor-in-Chief of IET Smart Cities Journal. He has been a Judge of GSMA GLOMO Award at World Mobile CongressC Barcelona since 2019. He was a Distinguished Lecturer of IEEE ComSoc (2020-2021). He is a Member of Academia Europaea (MAE), a Fellow of IEEE, a Fellow of IET and a Distinguished Member of ACM.