STAR-RIS Enhanced Joint Physical Layer Security and Covert Communications for Multi-antenna mmWave Systems

Han Xiao, *Student Member, IEEE,* Xiaoyan Hu*, *Member, IEEE,*
Ang Li, *Senior Member, IEEE,* Wenjie Wang, *Senior Member, IEEE,*
Zhou Su, *Senior Member, IEEE,* Kai-Kit Wong, *Fellow, IEEE*, Kun Yang, *Fellow, IEEE*

***Abstract*—This paper investigates the utilization of simultaneously transmitting and reflecting RIS (STAR-RIS) in supporting joint physical layer security (PLS) and covert communications (CCs) in a multi-antenna millimeter wave (mmWave) system, where the base station (BS) communicates with both covert and security users while defeating eavesdropping by wardens with the help of a STAR-RIS. Specifically, analytical derivations are performed to obtain the closed-form expression of warden's minimum detection error probability (DEP). Furthermore, the asymptotic result of the minimum DEP and the lower bound of the secure rates are derived, considering the practical assumption that BS only knows the statistical channel state information (CSI) between STAR-RIS and the wardens. Subsequently, an optimization problem is formulated with the aim of maximizing the average sum of the covert rate and the minimum secure rate while ensuring the covert requirement and quality of service (QoS) for legal users by jointly optimizing the active and passive beamformers. Due to the strong coupling among variables, an iterative algorithm based on the alternating strategy and the semi-definite relaxation (SDR) method is proposed to solve the non-convex optimization problem. Simulation results indicate that the performance of the proposed STAR-RIS-assisted scheme greatly surpasses that of the conventional RIS scheme, which validates the superiority of STAR-RIS in simultaneously implementing PLS and CCs.**

***Index Terms*—Covert communications, Physical layer security, STAR-RIS, Multi-antenna, mmWave.**

Manuscript received 15 July, 2023; revised 12 November, 2023; accepted 10 January, 2024. This work is supported in part by the National Natural Science Foundation of China (NSFC) under Grant 62201449, and in part by the Key R&D Projects of Shaanxi Province under Grant 2023-YBGY-040, and 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 Ang Li was supported in part by the Young Elite Scientists Sponsorship Program by CIC under Grant 2021QNRC001 and in part by the Science and Technology Program of Shaanxi Province under Grant 2021KWZ-01. An earlier version of this paper was presented in part at the IEEE 98th Vehicular Technology Conference (VTC2023-fall), Hong Kong, China, October 2023, Doi: 10.1109/VTC2023-Fall60731.2023.10333375. The associate editor coordinating the review of this article and approving it for publication was Prof. An Liu. *(Corresponding author: Xiaoyan Hu.)*

H. Xiao, X. Hu, A. Li, W. Wang, and Z. Su 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, ang.li.2020@xjtu.edu.cn, wjwang@mail.xjtu.edu.cn, zhousu@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

As wireless communication technologies continue to develop rapidly, the security of communications has become a growing concern for both enterprises and individuals. To safeguard users' information from eavesdropping attacks, physical layer security (PLS) has emerged as a promising technique and garnered significant attention in recent years. As a pioneering work, [1] demonstrates that a positive perfect secrecy rate can be achieved at the transceiver if the eavesdropper's channel is a diminished form of the legitimate user's channel. Following this, numerous methods have been proposed with the aim of improving the performance of PLS [2]–[6]. In particular, [2] proposes a transmit antenna selection scheme to enhance the PLS considering the practical case without the knowledge of eavesdroppers' channel state information (CSI). Then [3] examines the potential of active beamforming to improve the security performance of Heterogeneous networks. In [4], the utilization of artificial noise (AN) is shown to be beneficial against eavesdropping. The authors of [5], [6] both explore the uncoordinated cooperative jamming schemes to maximize the secure rate while defeating the eavesdropping by appropriately allocating the jamming power.

However, in some scenarios like secret military operations, the security level provided by PLS may not be sufficient. This is because PLS can only hide the contents of messages but not the existence of communications between authorized users, which may leave security risks that can be exploited by unauthorized users to launch attacks [7]. Recently, covert communication (CC) as a novel security technology has drawn great attention from both military and civilian fields [8]. CC has the ability to fundamentally conceal the presence of communications between users, providing a higher level of security than PLS. Toward this end, Bash *et al.* first demonstrate that $\mathcal{O}(\sqrt{n})$ bits of information can be reliably transmitted with a low probability of detection over additive white Gaussian noise (AWGN) channels [9]. Since then, lots of efforts have been made to improve the covert performance [10]–[14]. In [11], two background noise models with noise uncertainty are proposed based on which the authors investigate the maximum achievable covert rate. Also, the potential covert performance gain brought by the channel uncertainty is explored in [12]. In [13], full-duplex receivers are leveraged to transmit jamming signals so as to degrade the detection capabilities of

wardens and maximize the covert throughput. The strategy of uninformed jamming is implemented to facilitate CCs between legal users through deliberately generating jamming signals under different channel models [14]. Different from the literature above, [15] investigates the advantages of centralized and distributed multi-antenna transmitters in defeating the wardens with random positions. Moreover, [16] exploits the impacts of the number of antennas at the adversary wardens on the covert rate and finds that a slight increase in the number of antennas results in a dramatic decrease in the covert rate.

Although the strategies mentioned above have demonstrated their effectiveness in enhancing the performance of PLS and CCs, it is necessary to acknowledge that their potentials may be highly constrained by the stochastic nature of the wireless propagation environment due to the fact that the proposed schemes will be designed to accommodate different channel conditions. Specifically, in communication systems operating at millimeter-wave (mmWave) frequencies, this constraint will be particularly pronounced due to the susceptibility of mmWave signals to blockages. In order to break through this constraint, reconfigurable intelligent surfaces (RISs) emerged as a promising solution which consists of numerous cost-effective metamaterial elements. Each element equipped at RIS can dynamically modify the electromagnetic characteristics of the incident signals (e.g., phase and amplitude), and reflect or transmit the modified signals to users. With the assistance of this process, RIS can establish a reconfigurable and desirable end-to-end virtual channel. These attractive features of RIS make it popular in both academia and industry, which have been widely investigated in the performance enhancement of wireless applications including PLS and CCs [17]–[20]. In particular, [17] explores the PLS in a RIS-aided multi-antenna communication system with strong eavesdropping channels, and the achievable secure rate is maximized by optimizing the active and passive beamformers. To further enhance the secrecy performance, a double RIS scheme incorporating inter-RIS signal reflections is proposed in [18]. In addition, [19] provides a general summary of the potential applications of RIS in enhancing CCs. The authors in [20] examine the performance gain of CCs facilitated by RIS, which indicates that RIS can enable perfect covertness subject to the instantaneous channel state information (CSI) of wardens being accessible.

It is noteworthy that the traditional RIS in the literature above only reflects incident signals, which requires both transmitters and receivers to situate on the same side of RIS [21]. In order to overcome this limitation, a novel RIS called simultaneously transmitting and reflecting RIS (STAR-RIS) is proposed and developed in [22], [23]. Specifically, compared with the conventional RIS, STAR-RIS can construct a full-space smart radio environment with 360° coverage. This is because STAR-RIS is capable of dividing incident signals into two parts: one that is reflected and one that is transmitted. This unique feature enables users positioned in all directions to enjoy the communications service. Hence, the STAR-RIS has the more enormous potential in wireless communications than traditional RIS, which has sparked significant interest from both academia and industry [22]. However, the research on incorporating STAR-RISs into wireless communication systems is still in its early stages. In terms of the secure/covert communications, only a small number of works investigate the secure/covert performance gain facilitated by STAR-RIS [24]–[27]. In particular, Han *et al.* [24] and Zhang *et al.* [25] investigate the potentials of the STAR-RIS in boosting downlink and uplink PLS, respectively. In [26], [27], the authors initially explore the potentials of STAR-RIS in CCs, indicating that the STAR-RIS-assisted CCs scheme significantly outperforms the conventional RIS-aided scheme.

In practical scenarios, it is highly possible that users have varying security requirements for communications, e.g., some users may require secure information transmissions and some users may need a higher level of covert communications. In this case, [28] first considers a scenario with both PLS and CCs users and analyzes the average sum rate between the secure rate and covert rate under the perfect and imperfect CSI. However, the inherent randomness of the wireless channels results in a limited average rate. To address this problem, we establish a novel system model enabled by the STAR-RIS for joint implementation of PLS and CCs in this paper. Our main contributions are summarized as follows:

- ***STAR-RIS-assisted Joint PLS and CCs Architecture:*** In this paper, we construct a STAR-RIS-assisted joint PLS and CCs architecture for the first time. This architecture allows the legitimate users who are located on two sides of the STAR-RIS and have varying security needs, e.g., PLS and CCs, to be simultaneously served by elaborately designing the passive reflected and transmitted coefficients of the STAR-RIS as well as the active transmit beamforming of the base station (BS).
- ***Closed-form Expressions of PLS/CCs System Indicators:*** For CCs, the optimal detection threshold and minimum detection error probability (DEP) at Warden are derived analytically. Additionally, the large system analytic technique is introduced to further derive the asymptotic analytic result of the minimum DEP, which is leveraged as the covert constraint to jointly optimize the active and passive beamformers. For PLS, we derive the close-form expression of a lower bound for the secure rate considering that only the statistical CSI of the eavesdropper is available at the BS.
- ***Optimization Problem Formulation under Practical Constraints:*** An optimization problem is formulated for the STAR-RIS-aided joint PLS and CCs system to maximize the average sum rate between the minimum secure rate and covert rate, subject to covert and quality of service (QoS) constraints. This is achieved by jointly optimizing the active and passive beamforming variables. In fact, it is challenging to handle this optimization problem due to the strong coupling among variables especially considering the non-convex amplitude constraint introduced by STAR-RIS.
- ***Iterative Algorithm with Guaranteed Convergence and Substantial Performance Gain:*** To solve the formulated optimization problem, we propose an iterative algorithm leveraging the alternating strategy and the semi-definite relaxation (SDR) method. Specifically, the optimization

problem is divided into two subproblems respectively for active and passive beamforming design. The convergence of the proposed algorithm can be guaranteed which is also verified by simulation results. The performance gain of the proposed STAR-RIS-assisted joint PLS and CCs scheme is demonstrated through comparing with a benchmark scheme utilizing the conventional RIS. The numerical results indicate that STAR-RIS offers greater potential in enhancing the performance of joint PLS and CCs as compared to the traditional RIS.

The remainder of this paper is organized as follows. Section II shows the considered STAR-RIS-aided system model for joint PLS and CCs. The minimum DEP and its asymptotic analytic result are analytically derived in Section III. Also, the close-form expression of the lower bound for the secure rate is also given in Section III. The proposed iterative algorithm and the analysis on its convergence and computational complexity are presented in Section IV. Section V gives the numerical simulation results to verify the effectiveness of the proposed scheme. Finally, the conclusion is made in Section VI.

Notation: Operator $\otimes$ denotes the Kronecker 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 $\mathbf{a}$ while $\mathrm{diag}(\mathbf{A})$ denotes a vector whose elements are composed of the diagonal elements of matrix $\mathbf{A}$. $|\cdot|$, $\|\cdot\|_2$ and $\|\cdot\|_{\mathrm{F}}$ indicate the complex modulus, the spectral norm and Frobenius norm, respectively. $\mathbb{C}^{M\times N}$ stands for the set of $M \times N$ complex matrices. $x \sim \mathcal{CN}(a, b)$ is the circularly symmetric complex Gaussian random variable with mean a and variance b. $\mathbf{I}_{M\times 1}$ indicates the vector with $M \times 1$ entries that are 1. $\mathbf{I}_K$ represents the $K \times K$ identity matrix and $\mathbf{e}_k$ is its k-th column.

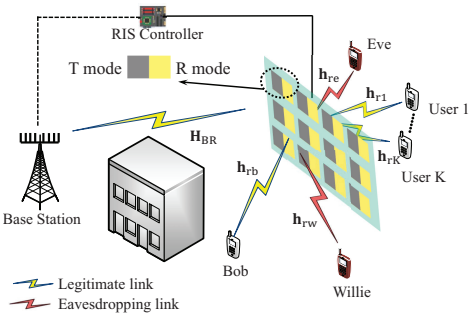


Fig. 1. System model for STAR-RIS-assisted joint PLS and CC.

II. System Model

We consider a STAR-RIS-aided system model for joint PLS and CCs as shown in Fig. 1, which comprises a base station (BS, Alice) with N_{t} antennas, a covert user (Bob), two warden/eavesdropping users (Willie and Eve), K security users with index $k \in \mathcal{K} \triangleq \{1, 2, \cdots, K\}$, and a STAR-RIS with M elements. All users are equipped with a single antenna and operate in half-duplex mode at the mmWave band. A practical scenario is investigated where the direct links between Alice and all users are blocked by obstacles such as buildings. To enhance the communication performance between Alice and legitimate users while impairing the detections by warden/eavesdropping users, an assistant STAR-RIS is deployed near the users. Note that, Bob is chosen as an undercover user, while Willie acts as a warden and attempts to detect the existence of communications between BS and Bob. In other words, Bob and Willie are a pair of legitimate user and eavesdropping user for CCs. Therefore, the CCs technologies will be leveraged to avoid detection by Willie. As an eavesdropper, Eve attempts to decode the information transmitted from Alice to K security users, forming K security user pairs. Hence, the PLS technologies will be utilized to safeguard the information from being decoded by Eve.

In this paper, we assume that the utilized STAR-RIS works at the energy splitting (ES) protocol [23] where all the elements operate in both reflected (R) and transmitted (T) modes simultaneously. Without loss of generality, the covert user Bob and security users are located on the opposite sides of the STAR-RIS as shown in Fig. 1, allowing them to be simultaneously served by the reflected and transmitted signals, respectively. Accordingly, we assume that the warden user Willie is located on the same side with Bob while the eavesdropper Eve is located on the same side with the security users. As a result, the covert user Bob located in the reflected region of the STAR-RIS can enjoy the service via the R mode of the STAR-RIS, while the K security users in the transmitted region can be served by elements operating at T mode.[1]

In this paper, Saleh-Valenzuela channel model [29] is adopted for the mmWave communications. In addition, we assume that the uniform linear array (ULA) of antennas is employed at the BS while the STAR-RIS adopts the uniform planar array (UPA). Hence, the channels between BS and STAR-RIS, and between STAR-RIS and users of {Bob, Willie, Eve, user $k \in \mathcal{K}$} can be modelled as

$$\mathbf{H}_{\mathrm{BR}} = \sqrt{\frac{N_{\mathrm{t}} M \rho_{\mathrm{BR}}}{L}} \sum_{l=1}^{L} \varphi_l^{\mathrm{BR}} \mathbf{a}_{\mathrm{R}}\left(\phi_l^{\mathrm{BR}}, \theta_l^{\mathrm{BR}}\right) \mathbf{a}_{\mathrm{B}}^H\left(\gamma_l^{\mathrm{BR}}\right), \quad (1)$$

$$\mathbf{h}_{\varsigma} = \sqrt{\frac{M \rho_{\varsigma}}{P}} \sum_{p=1}^{P} g_p^{\varsigma} \mathbf{a}_{\mathrm{R}}\left(\phi_p^{\varsigma}, \theta_p^{\varsigma}\right), \ \varsigma \in \{\mathrm{rb}, \mathrm{rw}, \mathrm{re}, \mathrm{r}k\}, \quad (2)$$

where $\mathbf{H}_{\mathrm{BR}} \in \mathbb{C}^{M\times N\mathrm{t}}$ and $\mathbf{h}_{\varsigma} \in \mathbb{C}^{M\times 1}$ with ρ_{BR} and ρ_{ς} being the path loss values related to BS-RIS link and RIS-users links, respectively. L, P denote the total number of paths in $\mathbf{H}_{\mathrm{BR}}$ and $\mathbf{h}_{\varsigma}$, and φ_l^{BR}, $g_p^{\varsigma} \sim \mathcal{CN}(0, 1)$ are the complex gain of the l-th path in $\mathbf{H}_{\mathrm{BR}}$ and p-th path in $\mathbf{h}_{\varsigma}$, respectively. Also, ϕ_l^{BR} and θ_l^{BR} represent the azimuth and elevation angle of arrival (AoA) values at STAR-RIS; γ_l^{BR} indicates the azimuth angle of departure (AoD) associated with BS; ϕ_p^{ς} and θ_p^{ς} denotes the azimuth and elevation AoD values associated with the RIS-users links. In addition, $\mathbf{a}_{\mathrm{R}}(\phi, \theta)$ and $\mathbf{a}_{\mathrm{B}}(\gamma)$ are respectively the beam steering vectors of the ULA and UPA at the BS and STAR-RIS, which are expressed as

- $\mathbf{a}_{\mathrm{B}}(\gamma) = \frac{1}{\sqrt{N_{\mathrm{t}}}}\big[1, \cdots, e^{j\frac{2\pi d}{\lambda}(n_{\mathrm{t}}-1)\sin(\gamma)}, \cdots, e^{j\frac{2\pi d}{\lambda}(N_{\mathrm{t}}-1)\sin(\gamma)}\big]^T$,

[1]We ignore the signals reflected or transmitted more than once by the STAR-RIS in the considered system.

- $\mathbf{a}_{\mathrm{R}}(\phi,\theta)=\frac{1}{\sqrt{M}}\big[1,\cdots,$
$e^{j\frac{2\pi d}{\lambda}((m_{\mathrm{y}}-1)\sin(\phi)\sin(\theta)+(m_{\mathrm{z}}-1)\cos(\theta))},$
$\cdots,e^{j\frac{2\pi d}{\lambda}((M_{\mathrm{y}}-1)\sin(\phi)\sin(\theta)+(M_{\mathrm{z}}-1)\cos(\theta))}\big]^T$,

where $n_{\mathrm{t}}\in\{1,2,\cdots,N_{\mathrm{t}}\}$, and $m_{\mathrm{y}}\in\{1,2,\cdots,M_{\mathrm{y}}\}$, $m_{\mathrm{z}}\in\{1,2,\cdots,M_{\mathrm{z}}\}$ with M_{y}, M_{z} being the number of elements in horizontal and vertical directions of UPA and $M=M_{\mathrm{y}}M_{\mathrm{z}}$. Note that, we utilize the quasi-static block fading channel assumption across all channels, which states that the channel remains constant within a block of symbols, but can vary independently from one block to another.

It is assumed that the considered STAR-RIS-assisted joint PLS and CCs system operates in time division duplex (TDD) mode, enabling the use of uplink channel estimation techniques to obtain the required CSI with channel reciprocity such as [30]. In addition, we assume that the BS has the knowledge of the instantaneous CSI between STAR-RIS and all legal users, i.e., $\mathbf{H}_{\mathrm{BR}}$, $\mathbf{h}_{\mathrm{rb}}$ and $\mathbf{h}_{\mathrm{r}k}$, while only the statistical CSI between STAR-RIS and Willie/Eve, i.e., $\mathbf{h}_{\mathrm{rw}}$ and $\mathbf{h}_{\mathrm{re}}$, are available at BS. In contrast, Willie knows the instantaneous CSI of $\mathbf{h}_{\mathrm{rw}}$, $\mathbf{h}_{\mathrm{rb}}$ and $\mathbf{h}_{\mathrm{r}k}$, but only the statistical CSI of $\mathbf{H}_{\mathrm{BR}}$ is accessible by Willie, which introduces uncertainty that is beneficial to cover the communications between BS and Bob. The reasonability of these assumptions can be simply verified as follows: (i) BS is able to evaluate the instantaneous CSI of $\mathbf{H}_{\mathrm{BR}}$, $\mathbf{b}_{\mathrm{rb}}$ and $\mathbf{h}_{\mathrm{r}k}$ by the received pilot signals from legitimate users. (ii) Because the signal leakages are nearly impossible to avoid in practical radiometers [31]. Therefore, BS can utilize some advanced detection tools to capture the leaked signals from Willie, which will help to determine the suspected location area and estimate the statistical CSI of $\mathbf{h}_{\mathrm{rw}}$ and $\mathbf{h}_{\mathrm{re}}$. (iii) Willie has the capability to estimate the instantaneous CSI of $\mathbf{h}_{\mathrm{rw}}$, $\mathbf{h}_{\mathrm{rb}}$, and $\mathbf{h}_{\mathrm{r}k}$ based on the received pilot signals from legitimate users. Furthermore, Willie possesses knowledge regarding the suspected positions of both Alice and the STAR-RIS, enabling him to acquire the statistical CSI of $\mathbf{H}_{\mathrm{BR}}$.

III. Analysis on the STAR-RIS-Assisted Joint PLS and CCs System

A. Theoretical Analysis on CCs

In this section, we focus on the theoretical analysis for the CCs of the STAR-RIS assisted system. Specifically, we first discuss Willie's detection strategy for CCs between BS and Bob, and then derive the closed-form expressions of its detection error probability (DEP) and the optimal detection threshold. Specifically, Willie determines the existence of communications between BS and Bob through the received signal sequences in a time slot, denoted as $\{y_{\mathrm{w}}[t]\}_{t=1}^{T}$, where $t\in\mathcal{T}\triangleq\{1,\cdots,T\}$ is the index of each communication channel use with the maximum number of T. It has to face a binary hypothesis for the judgement of CCs, which includes a null hypothesis $\mathcal{H}_0$, denoting that BS only communicates with K security users without CCs to Bob; and an alternative hypothesis $\mathcal{H}_1$, indicating that there exists CCs between BS and Bob. Under these two hypotheses, the received signals at Bob, Willie and the k-th security user can be respectively expressed as

$$y_{\mathrm{b}}[t]=\begin{cases}\sum\limits_{k=1}^{K}\mathbf{h}_{\mathrm{rb}}^{H}\mathbf{\Theta}_{\mathrm{r}}\mathbf{H}_{\mathrm{BR}}\mathbf{w}_k s_k[t]+n_{\mathrm{b}}[t], & \mathcal{H}_0,\\ \mathbf{h}_{\mathrm{rb}}^{H}\mathbf{\Theta}_{\mathrm{r}}\mathbf{H}_{\mathrm{BR}}\mathbf{w}_{\mathrm{b}}s_{\mathrm{b}}[t]+\\ \sum\limits_{k=1}^{K}\mathbf{h}_{\mathrm{rb}}^{H}\mathbf{\Theta}_{\mathrm{r}}\mathbf{H}_{\mathrm{BR}}\mathbf{w}_k s_k[t]+n_{\mathrm{b}}[t], & \mathcal{H}_1,\end{cases}\tag{3}$$

$$y_{\mathrm{w}}[t]=\begin{cases}\sum\limits_{k=1}^{K}\mathbf{h}_{\mathrm{rw}}^{H}\mathbf{\Theta}_{\mathrm{r}}\mathbf{H}_{\mathrm{BR}}\mathbf{w}_k s_k[t]+n_{\mathrm{w}}[t], & \mathcal{H}_0,\\ \mathbf{h}_{\mathrm{rw}}^{H}\mathbf{\Theta}_{\mathrm{r}}\mathbf{H}_{\mathrm{BR}}\mathbf{w}_{\mathrm{b}}s_{\mathrm{b}}[t]+\\ \sum\limits_{k=1}^{K}\mathbf{h}_{\mathrm{rw}}^{H}\mathbf{\Theta}_{\mathrm{r}}\mathbf{H}_{\mathrm{BR}}\mathbf{w}_k s_k[t]+n_{\mathrm{w}}[t], & \mathcal{H}_1,\end{cases}\tag{4}$$

$$y_k[t]=\begin{cases}\sum\limits_{j=1}^{K}\mathbf{h}_{\mathrm{r}k}^{H}\mathbf{\Theta}_{\mathrm{t}}\mathbf{H}_{\mathrm{BR}}\mathbf{w}_j s_j[t]+n_k[t], & \mathcal{H}_0,\\ \mathbf{h}_{\mathrm{r}k}^{H}\mathbf{\Theta}_{\mathrm{t}}\mathbf{H}_{\mathrm{BR}}\mathbf{w}_{\mathrm{b}}s_{\mathrm{b}}[t]+\\ \sum\limits_{j=1}^{K}\mathbf{h}_{\mathrm{r}k}^{H}\mathbf{\Theta}_{\mathrm{t}}\mathbf{H}_{\mathrm{BR}}\mathbf{w}_j s_j[t]+n_k[t], & \mathcal{H}_1,\end{cases}\tag{5}$$

where $\mathbf{\Theta}_{\xi}=\mathrm{Diag}\left\{\sqrt{\beta_{\xi}^{1}}e^{\mathrm{j}\phi_{\xi}^{1}},\ldots,\sqrt{\beta_{\xi}^{M}}e^{\mathrm{j}\phi_{\xi}^{M}}\right\}$ with $\xi\in\{\mathrm{r},\mathrm{t}\}$ indicates the reflected or transmitted coefficient matrix of STAR-RIS, where $\beta_{\xi}^{m}\in[0,1]$, $\phi_{\xi}^{m}\in[0,2\pi)$ and $\beta_{\mathrm{r}}^{m}+\beta_{\mathrm{t}}^{m}=1$ for $\forall m\in\mathcal{M}\triangleq\{1,2,\ldots,M\}$. Also, s_{b}, $s_k\sim\mathcal{CN}(0,1)$ represent the signals transmitted by BS to Bob and the security user k, while $\mathbf{w}_{\mathrm{b}}$, $\mathbf{w}_k$ are the corresponding beamforming vectors. In addition, $n_{\mathrm{b}}\sim\mathcal{CN}(0,\sigma_{\mathrm{b}}^2)$, $n_{\mathrm{w}}\sim\mathcal{CN}(0,\sigma_{\mathrm{w}}^2)$ and $n_k\sim\mathcal{CN}(0,\sigma_k^2)$ for $k\in\mathcal{K}$ denote the AWGN noise at Bob, Willie and the k-th security user.

We assume that Willie leverages a radiometer to detect CCs, where the average power of the received signals, i.e., $\overline{P}_{\mathrm{w}}=\frac{1}{T}\sum_{t=1}^{T}|y_{\mathrm{w}}[t]|^2$, is used to do the statistical test. In line with the existing works (e.g., [7], [26]), it is assumed that Willie utilizes an infinite number of signal samples, i.e., $T\to\infty$, to judge the binary hypotheses. Hence, the received average power can be derived as

$$\begin{aligned}\overline{P}_{\mathrm{w}}&=\lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}|y_{\mathrm{w}}[t]|^2\\&=\begin{cases}\sum\limits_{k=1}^{K}\left|\mathbf{h}_{\mathrm{rw}}^{H}\mathbf{\Theta}_{\mathrm{r}}\mathbf{H}_{\mathrm{BR}}\mathbf{w}_k\right|^2+\sigma_{\mathrm{w}}^2, & \mathcal{H}_0,\\ \left|\mathbf{h}_{\mathrm{rw}}^{H}\mathbf{\Theta}_{\mathrm{r}}\mathbf{H}_{\mathrm{BR}}\mathbf{w}_{\mathrm{b}}\right|^2+\\ \sum\limits_{k=1}^{K}\left|\mathbf{h}_{\mathrm{rw}}^{H}\mathbf{\Theta}_{\mathrm{r}}\mathbf{H}_{\mathrm{BR}}\mathbf{w}_k\right|^2+\sigma_{\mathrm{w}}^2, & \mathcal{H}_1,\end{cases}\end{aligned}\tag{6}$$

where the uncertainty of the noise n_w has been averaged out assuming that Willie is capable to know the noise power σ_{w}^2.

To determine the existence of CCs between BS and Bob, Willie needs to analyze $\overline{P}_{\mathrm{w}}$ under the hypotheses of $\mathcal{H}_0$ and $\mathcal{H}_1$ by leveraging the decision rule $\overline{P}_{\mathrm{w}}\underset{\mathcal{D}_0}{\overset{\mathcal{D}_1}{\gtrless}}\tau_{\mathrm{dt}}$, where $\mathcal{D}_0$ (or $\mathcal{D}_1$) is the decision that Willie favors $\mathcal{H}_0$ (or $\mathcal{H}_1$) and τ_{dt} is the corresponding detection threshold. In this paper, we adopt DEP to characterize Willie's detection ability for CCs between BS and Bob, considering the worst-case scenario where Willie can optimize τ_{dt} to obtain the optimal detection threshold and the minimum DEP. Next, we will analytically derive the minimum DEP based on the false alarm (FA) probability and the miss

detection (MD) probability from Willie's perspective, where FA indicates that Willie makes decision $\mathcal{D}_1$ under hypothesis $\mathcal{H}_0$ with probability $P_{\rm FA} = \Pr(\mathcal{D}_1|\mathcal{H}_0)$ while MD means that Willie makes decision $\mathcal{D}_0$ under hypothesis $\mathcal{H}_1$ with probability $P_{\rm MD} = \Pr(\mathcal{D}_0|\mathcal{H}_1)$. Specifically, $P_{\rm FA}$ and $P_{\rm MD}$ are given by Theorem 1.

Theorem 1. *The FA probability and the MD probability are expressed as*

$$P_{\rm FA} = \begin{cases} 1, & \tau_{\rm dt} \le \sigma_{\rm w}^2, \\ e^{-\frac{\tau_{\rm dt}-\sigma_{\rm w}^2}{\lambda_0}}, & \text{otherwise}, \end{cases} \tag{7}$$

$$P_{\rm MD} = \begin{cases} 0, & \tau_{\rm dt} \le \sigma_{\rm w}^2, \\ 1 - e^{-\frac{\tau_{\rm dt}-\sigma_{\rm w}^2}{\lambda_1}}, & \text{otherwise}, \end{cases} \tag{8}$$

where

- $\lambda_0 = \frac{N_{\rm t} M \rho_{\rm BR}}{L} \sum_{k=1}^{K} \left\| \mathbf{\Phi} \,\mathrm{vec}\left((\mathbf{w}_k \mathbf{h}_{\rm rw}^H \mathbf{\Theta}_{\rm r})^T \right) \right\|_2^2$,
- $\lambda_1 = \frac{N_{\rm t} M \rho_{\rm BR}}{L} \left\| \mathbf{\Phi} \,\mathrm{vec}\left((\mathbf{w}_{\rm b} \mathbf{h}_{\rm rw}^H \mathbf{\Theta}_{\rm r})^T \right) \right\|_2^2 + \lambda_0$,
- $\mathbf{\Phi} = \left[\mathrm{vec}(\mathbf{A}_1), \cdots, \mathrm{vec}(\mathbf{A}_L) \right]^H$,
- $\mathbf{A}_l = \mathbf{a}_{\rm R}\left(\phi_l^{\rm BR}, \theta_l^{\rm BR}\right) \mathbf{a}_{\rm B}^H\left(\gamma_l^{\rm BR}\right)$.

Proof: The proof is given in Appendix A. ∎

According to (7) and (8), we can find that when $\tau_{\rm dt} \le \sigma_{\rm w}^2$, FA will always be performed, but MD can be completely avoided. And with the increase of $\tau_{\rm dt}$ from 0 to ∞, $P_{\rm FA}$ will experience a decrease from 1 to 0, while $P_{\rm MD}$ has an opposite trend. Based on the analytical expression of $P_{\rm FA}$ and $P_{\rm MD}$ in (7) and (8), Willie's DEP can be derived as

$$\begin{aligned} P_{\rm e} =& P_{\rm FA} + P_{\rm MD} \\ =& \begin{cases} 1, & \tau_{\rm dt} \le \sigma_{\rm w}^2, \\ 1 - e^{-\frac{\tau_{\rm dt}-\sigma_{\rm w}^2}{\lambda_1}} + e^{-\frac{\tau_{\rm dt}-\sigma_{\rm w}^2}{\lambda_0}}, & \text{otherwise}. \end{cases} \end{aligned} \tag{9}$$

In this paper, we focus on the uncertain scenario with detection threshold $\tau_{\rm dt} > \sigma_{\rm w}^2$. Next, we will analyze and derive the optimal detection threshold, denoted as $\tau_{\rm dt}^*$, and the minimum DEP $P_{\rm e}^*$. In particular, the first-order partial derivative of $P_{\rm e}$ with respect to (w.r.t.) $\tau_{\rm dt}$ is given by

$$\frac{\partial P_{\rm e}}{\partial \tau_{\rm dt}} = \frac{e^{-\frac{\tau_{\rm dt}-\sigma_{\rm w}^2}{\lambda_1}}}{\lambda_1} - \frac{e^{-\frac{\tau_{\rm dt}-\sigma_{\rm w}^2}{\lambda_0}}}{\lambda_0}. \tag{10}$$

Let $\frac{\partial P_{\rm e}}{\partial \tau_{\rm dt}} = 0$, we can obtain the unique solution of $\tau_{\rm dt} = \frac{\lambda_1 \lambda_0 \ln \frac{\lambda_1}{\lambda_0}}{\lambda_1 - \lambda_0} + \sigma_{\rm w}^2$. It is easy to verify that $\frac{\partial P_{\rm e}}{\partial \tau_{\rm dt}} > 0$ when $\tau_{\rm dt} > \frac{\lambda_1 \lambda_0 \ln \frac{\lambda_1}{\lambda_0}}{\lambda_1 - \lambda_0} + \sigma_{\rm w}^2$, while $\frac{\partial P_{\rm e}}{\partial \tau_{\rm dt}} < 0$ when $\tau_{\rm dt} < \frac{\lambda_1 \lambda_0 \ln \frac{\lambda_1}{\lambda_0}}{\lambda_1 - \lambda_0} + \sigma_{\rm w}^2$. Hence, the optimal detection threshold minimizing $P_{\rm e}$ can be expressed as $\tau_{\rm dt}^* = \frac{\lambda_1 \lambda_0 \ln \frac{\lambda_1}{\lambda_0}}{\lambda_1 - \lambda_0} + \sigma_{\rm w}^2$ and the corresponding minimum DEP is derived as

$$P_{\rm e}^* = 1 - e^{-\frac{\lambda_0 \ln \frac{\lambda_1}{\lambda_0}}{\lambda_1 - \lambda_0}} + e^{-\frac{\lambda_1 \ln \frac{\lambda_1}{\lambda_0}}{\lambda_1 - \lambda_0}}. \tag{11}$$

In order to guarantee the covertness of communications between BS and Bob, $P_{\rm e}^* \ge 1 - \epsilon$ is required where $\epsilon \in (0, 1)$ is a quite small value required by the system performance indicators. Considering that only the statistical CSI of $\mathbf{h}_{\rm rw}$ is available at BS, the average minimum DEP over $\mathbf{h}_{\rm rw}$, i.e., $\overline{P}_{\rm e}^* = \mathbb{E}_{\mathbf{h}_{\rm rw}}(P_{\rm e}^*)$, is utilized to evaluate the covert performance. However, in (11), λ_0 and λ_1 are both random functions of $\mathbf{h}_{\rm rw}$ and are coupled with each other, which makes it challenging to directly calculate $\overline{P}_{\rm e}^*$. To tackle this problem, the large system analytic technique is leveraged to handle the coupling between λ_0 and λ_1, which is widely adopted to analyze the performance limitations of RIS-assisted wireless communication systems (e.g., [26], [32]). Specifically, we assume that a large number of low-cost elements are equipped at STAR-RIS, and thus the asymptotic analytic results of λ_0 and λ_1 can be obtained as in the following Theorem 2.

Theorem 2. *Applying the large system analytic technique with $M \to \infty$ on λ_0 and λ_1, the asymptotic analytic results are given by*

$$\begin{aligned} \widehat{\lambda}_0 =& \frac{N_{\rm t} M^2 \rho_{\rm BR} \rho_{\rm rw}}{LP} \sum_{k=1}^{K} \sum_{l=1}^{L} \left(\mathbf{w}_k^H \mathbf{\Psi}_{\rm BR}^l \mathbf{w}_k \right) \\ & \left(\boldsymbol{\vartheta}_{\rm r}^T \mathbf{\Xi}^T \left(\left(\widehat{\mathbf{\Psi}}_{\rm BR}^l \right)^T \otimes \left(\mathbf{\Omega}_{\rm rw}^H \mathbf{\Omega}_{\rm rw} \right) \right) \mathbf{\Xi} \boldsymbol{\vartheta}_{\rm r}^* \right), \end{aligned} \tag{12}$$

$$\begin{aligned} \widehat{\lambda}_1 =& \widehat{\lambda}_0 + \frac{N_{\rm t} M^2 \rho_{\rm BR} \rho_{\rm rw}}{LP} \sum_{l=1}^{L} \left(\mathbf{w}_{\rm b}^H \mathbf{\Psi}_{\rm BR}^l \mathbf{w}_{\rm b} \right) \\ & \left(\boldsymbol{\vartheta}_{\rm r}^T \mathbf{\Xi}^T \left(\left(\widehat{\mathbf{\Psi}}_{\rm BR}^l \right)^T \otimes \left(\mathbf{\Omega}_{\rm rw}^H \mathbf{\Omega}_{\rm rw} \right) \right) \mathbf{\Xi} \boldsymbol{\vartheta}_{\rm r}^* \right), \end{aligned} \tag{13}$$

where

- $\boldsymbol{\vartheta}_{\rm r} = \mathrm{diag}(\mathbf{\Theta}_{\rm r})$, $\mathbf{\Psi}_{\rm BR}^l = \mathbf{a}_{\rm B}\left(\gamma_l^{\rm BR}\right) \mathbf{a}_{\rm B}^H\left(\gamma_l^{\rm BR}\right)$,
- $\widehat{\mathbf{\Psi}}_{\rm BR}^l = \mathbf{a}_{\rm R}\left(\phi_l^{\rm BR}, \theta_l^{\rm BR}\right) \mathbf{a}_{\rm R}^H\left(\phi_l^{\rm BR}, \theta_l^{\rm BR}\right)$,
- $\mathbf{\Omega}_{\rm rw} = \left[\mathbf{a}_{\rm R}\left(\phi_1^{\rm rw}, \theta_1^{\rm rw}\right), \cdots, \mathbf{a}_{\rm R}\left(\phi_P^{\rm rw}, \theta_P^{\rm rw}\right) \right]^H$,
- $\mathbf{\Xi} = \Big[[\mathbf{e}_1, \mathbf{0}_{M\times(M-1)}];\ [\mathbf{0}_{M\times 1}, \mathbf{e}_2, \mathbf{0}_{M\times(M-2)}];\ \cdots; [\mathbf{0}_{M\times(M-1)}, \mathbf{e}_M] \Big]$.

Proof: The proof is given in Appendix B. ∎

Thus, we can further obtain the asymptotic analytic result of the minimum DEP by substituting (12) and (13) into (11) and adopting some algebraic manipulations, which is expressed as

$$P_{\rm ea}^* = 1 - e^{-\frac{\beta \ln \frac{\alpha+\beta}{\beta}}{\alpha}} \left(1 - \frac{\beta}{\alpha + \beta} \right), \tag{14}$$

where

- $\alpha = \frac{N_{\rm t} M^2 \rho_{\rm BR} \rho_{\rm rw}}{LP} \sum_{l=1}^{L} \left(\mathbf{w}_{\rm b}^H \mathbf{\Psi}_{\rm BR}^l \mathbf{w}_{\rm b} \right) \left(\boldsymbol{\vartheta}_{\rm r}^T \mathbf{\Xi}^T \left(\left(\widehat{\mathbf{\Psi}}_{\rm BR}^l \right)^T \otimes \left(\mathbf{\Omega}_{\rm rw}^H \mathbf{\Omega}_{\rm rw} \right) \right) \mathbf{\Xi} \boldsymbol{\vartheta}_{\rm r}^* \right)$,
- $\beta = \frac{N_{\rm t} M^2 \rho_{\rm BR} \rho_{\rm rw}}{LP} \sum_{k=1}^{K} \sum_{l=1}^{L} \left(\mathbf{w}_k^H \mathbf{\Psi}_{\rm BR}^l \mathbf{w}_k \right) \left(\boldsymbol{\vartheta}_{\rm r}^T \mathbf{\Xi}^T \left(\left(\widehat{\mathbf{\Psi}}_{\rm BR}^l \right)^T \otimes \left(\mathbf{\Omega}_{\rm rw}^H \mathbf{\Omega}_{\rm rw} \right) \right) \mathbf{\Xi} \boldsymbol{\vartheta}_{\rm r}^* \right)$.

In the following sections, the convert constraint $P_{\rm ea}^* \ge 1 - \epsilon$ will be utilized to characterize and guarantee the covert performance of the system.

Remark 1. *Based on the* (14)*, several key conclusions can be made. Specifically, the first-order partial derivative of $P_{\rm ea}^*$ w.r.t. α can be expressed as*

$$\frac{\partial P_{\rm ea}^*}{\partial \alpha} = -\frac{\left(\frac{\alpha+\beta}{\beta} \right)^{-\frac{\alpha+\beta}{\alpha}} \ln\left(\frac{\alpha+\beta}{\beta} \right)}{\alpha} < 0, \tag{15}$$

which shows that $P_{\rm ea}^*$ is a monotonically decreasing function regarding α. It is evident that as α increases, Willie's detection ability will improve. Furthermore, by utilizing the L'Hospital's rule, we can determine the limit value of $P_{\rm ea}^*$ as $\alpha \to +\infty$, i.e.,

$\lim_{\alpha\to+\infty} P_{\rm ea}^{*}=0$*, which is consistent with this fact that as the transmit power allocated to Bob increases, the possibility of Willie detecting the communications between Alice and Bob also increases.*

$$\frac{\partial P_{\rm ea}^{*}}{\partial\beta}=\frac{\left(\frac{\alpha+\beta}{\beta}\right)^{-\frac{\beta}{\alpha}}\ln\left(\frac{\alpha+\beta}{\beta}\right)}{\alpha+\beta}>0. \tag{16}$$

Moreover, we can derive the first-order partial derivative of $P_{\rm ea}^{*}$ w.r.t. β, revealing that $P_{\rm ea}^{*}$ exhibits a monotonically increasing behaviour in relation to β, as shown in equation (16)*. Moreover, by applying the L'Hospital's rule, it can be determined that $\lim_{\beta\to+\infty} P_{\rm ea}^{*}=1$, indicating that the security users' signals have the potential to degrade the detection performance at Willie.*

Note that, when hypothesis $\mathcal{H}_1$ is true, the available covert rate at Bob can be expressed as

$$R_{\rm b}^{\rm c}=\log_2\Big(1+\frac{\left|\mathbf{h}_{\rm rb}^{H}\mathbf{\Theta}_{\rm r}\mathbf{H}_{\rm BR}\mathbf{w}_{\rm b}\right|^2}{\sum_{k=1}^{K}\left|\mathbf{h}_{\rm rb}^{H}\mathbf{\Theta}_{\rm r}\mathbf{H}_{\rm BR}\mathbf{w}_{k}\right|^2+\sigma_{\rm b}^2}\Big). \tag{17}$$

B. Theoretical Analysis on PLS

In this section, the theoretical analysis on the PLS of the system is addressed, where we analytically derive the secure rate of all the security users considering that BS only knows the statistical CSI of $\mathbf{h}_{\rm re}$. Specifically, the signals received by the k-th security user are given by equation (5), while the signals received by Eve can be expressed as.

$$y_{\rm e}=\begin{cases}\sum\limits_{k=1}^{K}\mathbf{h}_{\rm re}^{H}\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\mathbf{w}_k s_k+n_{\rm e}, & \mathcal{H}_0,\\ \mathbf{h}_{\rm re}^{H}\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\mathbf{w}_{\rm b}s_{\rm b}+ & \\ \sum\limits_{k=1}^{K}\mathbf{h}_{\rm re}^{H}\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\mathbf{w}_k s_k+n_{\rm e}, & \mathcal{H}_1.\end{cases} \tag{18}$$

Therefore, the secure rate for the k-th user is given as

$$R_{\rm s}^{k}=\begin{cases}\left[\log_2\left(1+\gamma_{k0}\right)-\log_2\left(1+\gamma_{\rm e0}\right)\right]^{+}, & \mathcal{H}_0,\\ \left[\log_2\left(1+\gamma_{k1}\right)-\log_2\left(1+\gamma_{\rm e1}\right)\right]^{+}, & \mathcal{H}_1,\end{cases} \tag{19}$$

where $\gamma_{k1}=\frac{\left|\mathbf{h}_{{\rm r}k}^{H}\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\mathbf{w}_k\right|^2}{\left|\mathbf{h}_{{\rm r}k}^{H}\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\mathbf{w}_{\rm b}\right|^2+\sum_{j\neq k}^{K}\left|\mathbf{h}_{{\rm r}k}^{H}\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\mathbf{w}_j\right|^2+\sigma_k^2}$, $\gamma_{k0}=\frac{\left|\mathbf{h}_{{\rm r}k}^{H}\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\mathbf{w}_k\right|^2}{\sum_{j\neq k}^{K}\left|\mathbf{h}_{{\rm r}k}^{H}\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\mathbf{w}_j\right|^2+\sigma_k^2}$, $\gamma_{e0}=\frac{\left|\mathbf{h}_{\rm re}^{H}\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\mathbf{w}_k\right|^2}{\sum_{j\neq k}^{K}|\mathbf{h}_{\rm re}^{H}\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\mathbf{w}_j|^2+\sigma_{\rm e}^2}$, $\gamma_{e1}=\frac{\left|\mathbf{h}_{\rm re}^{H}\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\mathbf{w}_k\right|^2}{|\mathbf{h}_{\rm re}^{H}\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\mathbf{w}_{\rm b}|^2+\sum_{j\neq k}^{K}|\mathbf{h}_{\rm re}^{H}\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\mathbf{w}_j|^2+\sigma_{\rm e}^2}$, for $k\in\mathcal{K}$. Due to the fact that the BS can only acquire the statistical CSI of $\mathbf{h}_{\rm re}$, the average secure rates over $\mathbf{h}_{\rm re}$ are leveraged. In detail, the average secure rate for the k-th user can be further expressed as

$$\widehat{R}_{\rm s}^{k}=\begin{cases}\left[\log_2\left(1+\gamma_{k0}\right)-\mathbb{E}_{\mathbf{h}_{\rm re}}\left(\log_2\left(1+\gamma_{\rm e0}\right)\right)\right]^{+}, & \mathcal{H}_0,\\ \left[\log_2\left(1+\gamma_{k1}\right)-\mathbb{E}_{\mathbf{h}_{\rm re}}\left(\log_2\left(1+\gamma_{\rm e1}\right)\right)\right]^{+}, & \mathcal{H}_1.\end{cases} \tag{20}$$

It is easy to verify that $\mathbf{h}_{\rm re}\sim\mathcal{CN}\left(\mathbf{0}_{M\times1},\frac{M\rho_{\rm re}}{P}\mathbf{\Omega}_{\rm re}^{H}\mathbf{\Omega}_{\rm re}\right)$, where $\mathbf{\Omega}_{\rm re}=\left[\mathbf{a}_{\rm R}\left(\phi_1^{\rm re},\theta_1^{\rm re}\right),\cdots,\mathbf{a}_{\rm R}\left(\phi_P^{\rm re},\theta_P^{\rm re}\right)\right]^{H}$. Hence, the average eavesdropping rates, denoted as $\mathbb{E}_{\mathbf{h}_{\rm re}}\left(\log_2\left(1+\gamma_{\rm e0}\right)\right)$ and $\mathbb{E}_{\mathbf{h}_{\rm re}}\left(\log_2\left(1+\gamma_{\rm e1}\right)\right)$, can be derived as

$$\begin{aligned}&\mathbb{E}_{\mathbf{h}_{\rm re}}\big(\log_2\left(1+\gamma_{\rm e0}\right)\big)\\&=\mathbb{E}_{\mathbf{h}_{\rm re}}\Big(\log_2\Big(\sum_{k=1}^{K}\left|\mathbf{h}_{\rm re}^{H}\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\mathbf{w}_k\right|^2+\sigma_{\rm e}^2\Big)\\&-\log_2\Big(\sum_{j\neq k}^{K}\left|\mathbf{h}_{\rm re}^{H}\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\mathbf{w}_j\right|^2+\sigma_{\rm e}^2\Big)\Big)\\&=\int_0^{\infty}\log_2\left(x+\sigma_{\rm e}^2\right)\frac{e^{-\frac{x}{\eta_0}}}{\eta_0}dx-\int_0^{\infty}\log_2\left(\widehat{x}+\sigma_{\rm e}^2\right)\frac{e^{-\frac{\widehat{x}}{\widehat{\eta}_0}}}{\widehat{\eta}_0}d\widehat{x}\\&=\frac{e^{\frac{\sigma_{\rm e}^2}{\eta_0}}\Gamma(0,\frac{\sigma_{\rm e}^2}{\eta_0})}{\ln 2}-\frac{e^{\frac{\sigma_{\rm e}^2}{\widehat{\eta}_0}}\Gamma(0,\frac{\sigma_{\rm e}^2}{\widehat{\eta}_0})}{\ln 2}.\end{aligned} \tag{21}$$

Similarly,

$$\mathbb{E}_{\mathbf{h}_{\rm re}}\left(\log_2\left(1+\gamma_{\rm e1}\right)\right)=\frac{e^{\frac{\sigma_{\rm e}^2}{\eta_1}}\Gamma(0,\frac{\sigma_{\rm e}^2}{\eta_1})}{\ln 2}-\frac{e^{\frac{\sigma_{\rm e}^2}{\widehat{\eta}_1}}\Gamma(0,\frac{\sigma_{\rm e}^2}{\widehat{\eta}_1})}{\ln 2}, \tag{24}$$

where $\Gamma(\cdot,\cdot)$ is the upper incomplete Gamma function, and

- $\eta_0=\frac{M\rho_{\rm re}}{P}\sum_{k=1}^{K}\left\|\mathbf{\Omega}_{\rm re}\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\mathbf{w}_k\right\|^2$,
- $\widehat{\eta}_0=\frac{M\rho_{\rm re}}{P}\sum_{j\neq k}^{K}\left\|\mathbf{\Omega}_{\rm re}\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\mathbf{w}_j\right\|^2$,
- $\eta_1=\frac{M\rho_{\rm re}}{P}\left\|\mathbf{\Omega}_{\rm re}\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\mathbf{w}_{\rm b}\right\|^2+\eta_0$,
- $\widehat{\eta}_1=\frac{M\rho_{\rm re}}{P}\left\|\mathbf{\Omega}_{\rm re}\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\mathbf{w}_{\rm b}\right\|^2+\widehat{\eta}_0$.

The existence of the Gamma functions in average secure rate $\widehat{R}_{\rm s}^{k}$ makes it challenging to be handled for solving the optimization problem in the next section. To tackle this issue, the lower bound of $\widehat{R}_{\rm s}^{k}$ is leveraged to replace $\widehat{R}_{\rm s}^{k}$ as a robust secure rate, which is expressed as

$$\begin{aligned}&\widehat{R}_{\rm s}^{k}>\widehat{R}_{\rm sl}^{k}=\\&\begin{cases}\left[\log_2\left(1+\gamma_{k0}\right)-\mathbb{E}_{\mathbf{h}_{\rm re}}\left(\log_2\left(1+\tilde{\gamma}_{\rm e0}\right)\right)\right]^{+}, & \mathcal{H}_0,\\ \left[\log_2\left(1+\gamma_{k1}\right)-\mathbb{E}_{\mathbf{h}_{\rm re}}\left(\log_2\left(1+\tilde{\gamma}_{\rm e1}\right)\right)\right]^{+}, & \mathcal{H}_1,\end{cases}\end{aligned} \tag{22}$$

where $\tilde{\gamma}_{\rm e1}=\frac{\left|\mathbf{h}_{\rm re}^{H}\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\mathbf{w}_k\right|^2}{|\mathbf{h}_{\rm re}^{H}\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\mathbf{w}_b|^2+\sum_{j\neq k}^{K}|\mathbf{h}_{\rm re}^{H}\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\mathbf{w}_j|^2}>\gamma_{\rm e1}$ and $\tilde{\gamma}_{\rm e0}=\frac{\left|\mathbf{h}_{\rm re}^{H}\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\mathbf{w}_k\right|^2}{\sum_{j\neq k}^{K}|\mathbf{h}_{\rm re}^{H}\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\mathbf{w}_j|^2}>\gamma_{\rm e0}$. We can further derive that $\mathbb{E}_{\mathbf{h}_{\rm re}}\big(\log_2\left(1+\tilde{\gamma}_{\rm e0}\right)\big)=\log_2\left(\frac{\eta_0}{\widehat{\eta}_0}\right)\mathbb{E}_{\mathbf{h}_{\rm re}}\big(\log_2\left(1+\tilde{\gamma}_{\rm e1}\right)\big)=\log_2\left(\frac{\eta_1}{\widehat{\eta}_1}\right)$. Hence, the lower bound robust counterpart of the average secure rate for the k-the security user under two hypotheses are given by

$$\widehat{R}_{\rm sl,0}^{k}=\big[\log_2\left(1+\gamma_{k0}\right)-\log_2\left(\frac{\eta_0}{\widehat{\eta}_0}\right)\big]^{+}, \tag{23}$$

$$\widehat{R}_{\rm sl,1}^{k}=\big[\log_2\left(1+\gamma_{k1}\right)-\log_2\left(\frac{\eta_1}{\widehat{\eta}_1}\right)\big]^{+}, \tag{24}$$

which will be used in the next section for variables optimization and algorithm design.

IV. Problem Formulation and Algorithm Design

A. Optimization Problem Formulation

In this section, we will establish an optimization problem based on the theoretical analysis in Section III. Considering that the existence of the CCs between BS and Bob is under a binary hypothesis, we define a Bernoulli variable b where $b=0$ with the probability of P_0 means that BS only transmits

the secure information, while $b = 1$ with the probability of $P_1 = 1 - P_0$ represents that BS transmits both the covert and secure messages. In this paper, we maximize the average sum rate between the covert rate and the minimum secure rate over b in a time slot while ensuring the covert constraint and the QoS constraints at Bob and K security users by jointly optimizing the active and passive beamforming variables, i.e., $\mathbf{w}_{\rm b}$, $\{\mathbf{w}_k\}_{k=1}^K$, and $\mathbf{\Theta}_{\rm r}$, $\mathbf{\Theta}_{\rm t}$. Specifically, the optimization objective of the average sum rate over the Bernoulli variable b can be expressed as

$$\begin{aligned}&\overline{R}\big(\mathbf{w}_{\rm b}, \{\mathbf{w}_k\}_{k=1}^K, \mathbf{\Theta}_{\rm r}, \mathbf{\Theta}_{\rm t}\big)\\ =&\mathbb{E}_b\Big(bR_{\rm b}^{\rm c} + L(b)\min_k \widehat{R}_{{\rm sl},0}^k + b\min_k \widehat{R}_{{\rm sl},1}^k\Big)\\ =&P_1R_{\rm b}^{\rm c} + P_0\min_k \widehat{R}_{{\rm sl},0}^k + P_1\min_k \widehat{R}_{{\rm sl},1}^k, \end{aligned} \tag{25}$$

where $L(\cdot)$ is the logical operator with $L(0) = 1$, $L(1) = 0$.

Based on the above analysis, the optimization problem is formulated as

$$\max_{\mathbf{w}_{\rm b}, \{\mathbf{w}_k\}_{k=1}^K, \mathbf{\Theta}_{\rm r}, \mathbf{\Theta}_{\rm t}} \overline{R}\big(\mathbf{w}_{\rm b}, \{\mathbf{w}_k\}_{k=1}^K, \mathbf{\Theta}_{\rm r}, \mathbf{\Theta}_{\rm t}\big),$$

$$\text{s.t. } \|\mathbf{w}_{\rm b}\|_2^2 + \sum_{k=1}^K \|\mathbf{w}_k\|_2^2 \le P_{\rm tmax}, \tag{26a}$$

$$e^{-\frac{\beta \ln \frac{\alpha+\beta}{\beta}}{\alpha}}\Big(1 - \frac{\beta}{\alpha+\beta}\Big) \le \epsilon, \tag{26b}$$

$$R_{\rm b}^{\rm c} \ge R_{\rm b}^*, \tag{26c}$$

$$\min_k \widehat{R}_{{\rm sl},0}^k \ge R_{\rm s0}^*,\ k \in \mathcal{K}, \tag{26d}$$

$$\min_k \widehat{R}_{{\rm sl},1}^k \ge R_{\rm s1}^*,\ k \in \mathcal{K}, \tag{26e}$$

$$\beta_{\rm r}^m + \beta_{\rm t}^m = 1, \phi_{\rm r}^m,\ \phi_{\rm t}^m \in [0, 2\pi),\ m \in \mathcal{M}, \tag{26f}$$

where (26a) is the transmit power constraint of the BS with $P_{\rm tmax}$ being the maximum power budget; (26b) denotes the covertness constraint, which is equivalent to $P_{\rm ea}^* \ge 1-\epsilon$; (26c) and (26d), (26e) represent the QoS constraints for covert rate and secure rate with the minimum required covert rate $R_{\rm b}^*$ and secure rate $R_{\rm s0}^*$ and $R_{\rm s1}^*$; (26f) is the amplitude and phase shift constraints for STAR-RIS. In fact, solving this optimization problem is quite challenging due to the strong coupling among variables, i.e., $\mathbf{w}_{\rm b}$, $\{\mathbf{w}_k\}_{k=1}^K$, $\mathbf{\Theta}_{\rm r}$ and $\mathbf{\Theta}_{\rm t}$, in the objective function, covert constraint and QoS constraints. Additionally, the characteristic amplitude constraint introduced by STAR-RIS complicates the problem because $\mathbf{\Theta}_{\rm r}$ and $\mathbf{\Theta}_{\rm t}$ depend on each other in terms of element amplitudes. As a result, it is difficult to directly solve the optimization problem (26) using convex optimization algorithms. To address this challenge, we propose an iterative algorithm that leverages an alternative strategy to effectively solve this optimization problem, which is presented in the next section.

B. Algorithm Design

In this section, we detail the proposed iterative algorithm for solving the originally formulated problem (26). Specifically, this problem is divided into two subproblems which are solved to design the active and passive beamformers, respectively.

1) Joint Active beamforming design for $\mathbf{w}_{\rm b}$ and $\{\mathbf{w}_k\}_{k=1}^K$: We first design the active beamforming variables $\mathbf{w}_{\rm b}$ and $\{\mathbf{w}_k\}_{k=1}^K$ with given the passive beamforming variables, i.e., $\mathbf{\Theta}_{\rm r}$ and $\mathbf{\Theta}_{\rm t}$. In this circumstance, the original optimization problem can be simplified as

$$\max_{\mathbf{w}_{\rm b}, \{\mathbf{w}_k\}_{k=1}^K} \overline{R}\big(\mathbf{w}_{\rm b}, \{\mathbf{w}_k\}_{k=1}^K\big),$$
$$\text{s.t.} \quad (26a) - (26e). \tag{27a}$$

Problem (27) is still a non-convex optimization problem due to and the max-min objective function, the covert constraint and the QoS constraints w.r.t. the active beamforming variables $\mathbf{w}_{\rm b}$ and $\{\mathbf{w}_k\}_{k=1}^K$. To tackle this problem, we first introduce three auxiliary variable $\iota_{\mathbf{w}}$, $\kappa_{\mathbf{w}}$ and $\varpi_{\mathbf{w}}$ to replace $R_{\rm b}^{\rm c}$, $\min_k \widehat{R}_{{\rm sl},1}^k$ and $\min_k \widehat{R}_{{\rm sl},0}^k$ in the objective function and the QoS constraints (26c)-(26e) so that the max-min optimization problem can be transformed as a maximization optimization problem. In addition, it is easy to verify that the left-side of (26b) is a monotonically decreasing function of $\frac{\beta}{\alpha}$, and thus the covert constraint (26b) can be equivalently transformed as $\frac{\beta}{\alpha} \ge \varphi(\epsilon)$, where $\varphi(\epsilon)$ can be obtained by using the numerical methods such bisection search method.

Hence, problem (27) can be equivalently transformed into problem (28) shown at the top of the next page, where we have $\mathbf{D} = \sum_{l=1}^L \mathbf{\Psi}_{\rm BR}^l \boldsymbol{\vartheta}_{\rm r}^T \mathbf{\Xi}^T \Big(\big(\widehat{\mathbf{\Psi}}_{\rm BR}^l\big)^T \otimes \big(\mathbf{\Omega}_{\rm rw}^H \mathbf{\Omega}_{\rm rw}\big)\Big)\mathbf{\Xi}\boldsymbol{\vartheta}_{\rm r}^*$, $\tilde{\mathbf{W}}_{-k} = \{\mathbf{w}_1, \cdots, \mathbf{w}_{k-1}, \mathbf{w}_{k+1}, \cdots, \mathbf{w}_K\}$. The convert constraint can be re-expressed as (28b) based on the definitions of α and β. In fact, (28) is still a non-convex optimization problem because of the non-convexity of the constraints (28b), (28d), (28e) and (28f). To effectively address this problem, we resort to the SDR method [33]. Specifically, we first let $\mathbf{W} = \{\mathbf{w}_{\rm b}, \mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_K\}$ and $\mathbf{W}_{\rm cs} = \mathrm{vec}\,(\mathbf{W})\,\mathrm{vec}\,(\mathbf{W})^H$, then the optimization problem (28) can be equivalently transformed as (29), which is presented in the next page and where

- $\widehat{\mathbf{D}} = \big(\mathbf{E}_{-1}\mathbf{E}_{-1}^T\big) \otimes \mathbf{D},\ \mathbf{D}_1 = \big(\mathbf{e}_1\mathbf{e}_1^T\big) \otimes \mathbf{D}$,
- $\widehat{\mathbf{A}} = \mathbf{I}_{K+1} \otimes \mathbf{A},\ \tilde{\mathbf{A}} = \big(\mathbf{E}_{-1}\mathbf{E}_{-1}^T\big) \otimes \mathbf{A}$,
- $\check{\mathbf{B}}_k = \big(\mathbf{E}_{-1}\mathbf{E}_{-1}^T\big) \otimes \mathbf{B}_k,\ \tilde{\mathbf{B}}_k = \big(\mathbf{E}_{-(k+1)}\mathbf{E}_{-(k+1)}^T\big) \otimes \mathbf{B}_k$,
- $\breve{\mathbf{B}}_k = \big(\mathbf{E}_{-(1,k+1)}\mathbf{E}_{-(1,k+1)}^T\big) \otimes \mathbf{B}_k,\ \widehat{\mathbf{B}}_k = \mathbf{I}_{K+1} \otimes \mathbf{B}_k$,
- $\check{\mathbf{C}}_k = \big(\mathbf{E}_{-1}\mathbf{E}_{-1}^T\big) \otimes \mathbf{C},\ \tilde{\mathbf{C}}_k = \big(\mathbf{E}_{-(k+1)}\mathbf{E}_{-(k+1)}^T\big) \otimes \mathbf{C}$,
- $\breve{\mathbf{C}}_k = \big(\mathbf{E}_{-(1,k+1)}\mathbf{E}_{-(1,k+1)}^T\big) \otimes \mathbf{C},\ \widehat{\mathbf{C}} = \mathbf{I}_{K+1} \otimes \mathbf{C}$,
- $\mathbf{E}_{-(1,k+1)} = \big\{\mathbf{e}_2, \cdots, \mathbf{e}_k, \mathbf{e}_{k+2}, \cdots, \mathbf{e}_{K+1}\big\}$,
- $\mathbf{E}_{-1} = \big\{\mathbf{e}_2, \cdots, \mathbf{e}_k, \cdots, \mathbf{e}_{K+1}\big\}$,
- $\mathbf{E}_{-(k+1)} = \big\{\mathbf{e}_1, \cdots, \mathbf{e}_k, \mathbf{e}_{k+2}, \cdots, \mathbf{e}_{K+1}\big\}$,
- $\mathbf{A} = \big(\mathbf{h}_{\rm rb}^H\mathbf{\Theta}_{\rm r}\mathbf{H}_{\rm BR}\big)^H\big(\mathbf{h}_{\rm rb}^H\mathbf{\Theta}_{\rm r}\mathbf{H}_{\rm BR}\big)$,
- $\mathbf{B}_k = \big(\mathbf{h}_{{\rm r}k}^H\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\big)^H\big(\mathbf{h}_{{\rm r}k}^H\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\big)$,
- $\mathbf{C} = \mathbf{H}_{\rm BR}^H\mathbf{\Theta}_{\rm t}^H\mathbf{\Omega}_{\rm re}^H\mathbf{\Omega}_{\rm re}\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}$.

Note that problem (29) is still a non-convex optimization problem due to the non-convex constraints (29d), (29e), (29f) and the rank-one constraint in (29g). To transform (29) into a solvable convex problem, we first handle the constraints (29d), (29e) and (29f). In particular, we can find that $f(\mathbf{W}_{\rm cs})$, $f_{k,1}(\mathbf{W}_{\rm cs})$ and $f_{k,2}(\mathbf{W}_{\rm cs})$ are all difference of concave (DC) functions, and thus the first-order Taylor expansion can be leveraged on them to obtain their concave lower bounds in

$$\max_{\mathbf{w}_{\rm b},\{\mathbf{w}_k\}_{k=1}^K,\iota_{\mathbf{w}},\kappa_{\mathbf{w}},\varpi_{\mathbf{w}}} P_1\iota_{\mathbf{w}}+P_1\kappa_{\mathbf{w}}+P_0\varpi_{\mathbf{w}},$$

$$\text{s.t. } \|\mathbf{w}_{\rm b}\|_2^2+\sum_{k=1}^K\|\mathbf{w}_k\|_2^2\le P_{\rm tmax}, \tag{28a}$$

$$\sum_{k=1}^K\left(\mathbf{w}_k^H\mathbf{D}\mathbf{w}_k\right)\ge\left(\mathbf{w}_{\rm b}^H\mathbf{D}\mathbf{w}_{\rm b}\right)\varphi(\epsilon), \tag{28b}$$

$$\iota_{\mathbf{w}}\ge R_{\rm b}^*,\varpi_{\mathbf{w}}\ge R_{\rm s0}^*,\kappa_{\mathbf{w}}\ge R_{\rm s1}^*, \tag{28c}$$

$$\log_2\left(1+\frac{\left|\mathbf{h}_{\rm rb}^H\mathbf{\Theta}_{\rm r}\mathbf{H}_{\rm BR}\mathbf{w}_{\rm b}\right|^2}{\sum_{k=1}^K\left|\mathbf{h}_{\rm rb}^H\mathbf{\Theta}_{\rm r}\mathbf{H}_{\rm BR}\mathbf{w}_k\right|^2+\sigma_{\rm b}^2}\right)\ge\iota_{\mathbf{w}}, \tag{28d}$$

$$\log_2\left(1+\frac{\left|\mathbf{h}_{{\rm r}k}^H\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\mathbf{w}_k\right|^2}{\left\|\mathbf{h}_{{\rm r}k}^H\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\tilde{\mathbf{W}}_{-k}\right\|_2^2+\sigma_{{\rm r}k}^2}\right)-\log_2\left(1+\frac{\left\|\mathbf{\Omega}_{\rm re}\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\mathbf{w}_k\right\|_2^2}{\left\|\mathbf{\Omega}_{\rm re}\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\tilde{\mathbf{W}}_{-k}\right\|_{\rm F}^2}\right)\ge\varpi_{\mathbf{w}},\ \forall k, \tag{28e}$$

$$\log_2\left(1+\frac{\left|\mathbf{h}_{{\rm r}k}^H\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\mathbf{w}_k\right|^2}{\left|\mathbf{h}_{{\rm r}k}^H\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\mathbf{w}_{\rm b}\right|^2+\left\|\mathbf{h}_{{\rm r}k}^H\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\tilde{\mathbf{W}}_{-k}\right\|_2^2+\sigma_k^2}\right)-$$

$$\log_2\left(1+\frac{\left\|\mathbf{\Omega}_{\rm re}\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\mathbf{w}_k\right\|_2^2}{\left\|\mathbf{\Omega}_{\rm re}\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\mathbf{w}_{\rm b}\right\|_2^2+\left\|\mathbf{\Omega}_{\rm re}\mathbf{\Theta}_{\rm t}\mathbf{H}_{\rm BR}\tilde{\mathbf{W}}_{-k}\right\|_{\rm F}^2}\right)\ge\kappa_{\mathbf{w}},\ \forall k. \tag{28f}$$

$$\max_{\mathbf{W}_{\rm cs},\iota_{\mathbf{w}},\kappa_{\mathbf{w}},\varpi_{\mathbf{w}}} P_1\iota_{\mathbf{w}}+P_1\kappa_{\mathbf{w}}+P_0\varpi_{\mathbf{w}},$$

$$\text{s.t. } {\rm Tr}(\mathbf{W}_{\rm cs})\le P_{\rm tmax}, \tag{29a}$$

$${\rm Tr}(\mathbf{W}_{\rm cs}\widehat{\mathbf{D}})\ge{\rm Tr}(\mathbf{W}_{\rm cs}\mathbf{D}_1)\varphi(\epsilon), \tag{29b}$$

$$\text{(28c)}, \tag{29c}$$

$$f(\mathbf{W}_{\rm cs})=\log_2\left({\rm Tr}(\mathbf{W}_{\rm cs}\widehat{\mathbf{A}})+\sigma_{\rm b}^2\right)-\log_2\left({\rm Tr}(\mathbf{W}_{\rm cs}\tilde{\mathbf{A}})+\sigma_{\rm b}^2\right)\ge\iota_{\mathbf{w}}, \tag{29d}$$

$$f_{k,1}(\mathbf{W}_{\rm cs})=$$

$$\log_2\left({\rm Tr}(\mathbf{W}_{\rm cs}\check{\mathbf{B}}_k)+\sigma_k^2\right)+\log_2\left({\rm Tr}(\mathbf{W}_{\rm cs}\check{\mathbf{C}}_k)\right)-\log_2\left({\rm Tr}(\mathbf{W}_{\rm cs}\check{\mathbf{B}}_k)+\sigma_k^2\right)-\log_2\left({\rm Tr}(\mathbf{W}_{\rm cs}\check{\mathbf{C}})\right)\ge\varpi_{\mathbf{w}},\ \forall k, \tag{29e}$$

$$f_{k,2}(\mathbf{W}_{\rm cs})=$$

$$\log_2\left({\rm Tr}(\mathbf{W}_{\rm cs}\widehat{\mathbf{B}}_k)+\sigma_k^2\right)+\log_2\left({\rm Tr}(\mathbf{W}_{\rm cs}\tilde{\mathbf{C}}_k)\right)-\log_2\left({\rm Tr}(\mathbf{W}_{\rm cs}\tilde{\mathbf{B}}_k)+\sigma_k^2\right)-\log_2\left({\rm Tr}(\mathbf{W}_{\rm cs}\widehat{\mathbf{C}})\right)\ge\kappa_{\mathbf{w}},\ \forall k, \tag{29f}$$

$$\mathbf{W}_{\rm cs}\succeq 0,\ \ {\rm rank}(\mathbf{W}_{\rm cs})=1. \tag{29g}$$

the i-th inner loop iteration of the proposed iterative algorithm (See Algorithm 1 in Section IV-C). These concave lower bounds will be adopted to replace the original expressions in the optimization problem (29) which are derived as

$$\begin{aligned}f(\mathbf{W}_{\rm cs})&\ge\log_2\left({\rm Tr}\left(\mathbf{W}_{\rm cs}\widehat{\mathbf{A}}\right)+\sigma_{\rm b}^2\right)-g_1\left(\mathbf{W}_{\rm cs},\mathbf{W}_{\rm cs}^{(i)}\right)\\&\triangleq\widehat{f}\left(\mathbf{W}_{\rm cs},\mathbf{W}_{\rm cs}^{(i)}\right),\end{aligned} \tag{30}$$

$$\begin{aligned}f_{k,1}(\mathbf{W}_{\rm cs})&\ge\log_2\left({\rm Tr}(\mathbf{W}_{\rm cs}\check{\mathbf{B}}_k)+\sigma_k^2\right)-g_{k,1}\left(\mathbf{W}_{\rm cs},\mathbf{W}_{\rm cs}^{(i)}\right)\\&\quad+\log_2\left({\rm Tr}\left(\mathbf{W}_{\rm cs}\check{\mathbf{C}}_k\right)\right)-g_2\left(\mathbf{W}_{\rm cs},\mathbf{W}_{\rm cs}^{(i)}\right)\\&\triangleq\widehat{f}_{k,1}\left(\mathbf{W}_{\rm cs},\mathbf{W}_{\rm cs}^{(i)}\right),\end{aligned} \tag{31}$$

$$\begin{aligned}f_{k,2}(\mathbf{W}_{\rm cs})&\ge\log_2\left({\rm Tr}\left(\mathbf{W}_{\rm cs}\widehat{\mathbf{B}}_k\right)+\sigma_k^2\right)-g_{k,2}\left(\mathbf{W}_{\rm cs},\mathbf{W}_{\rm cs}^{(i)}\right)\\&\quad+\log_2\left({\rm Tr}\left(\mathbf{W}_{\rm cs}\tilde{\mathbf{C}}_k\right)\right)-g_3\left(\mathbf{W}_{\rm cs},\mathbf{W}_{\rm cs}^{(i)}\right)\\&\triangleq\widehat{f}_{k,2}\left(\mathbf{W}_{\rm cs},\mathbf{W}_{\rm cs}^{(i)}\right),\end{aligned} \tag{32}$$

where the expressions of $g_1\left(\mathbf{W}_{\rm cs},\mathbf{W}_{\rm cs}^{(i)}\right)$, $g_2\left(\mathbf{W}_{\rm cs},\mathbf{W}_{\rm cs}^{(i)}\right)$, $g_3\left(\mathbf{W}_{\rm cs},\mathbf{W}_{\rm cs}^{(i)}\right)$, and $g_{k,1}\left(\mathbf{W}_{\rm cs},\mathbf{W}_{\rm cs}^{(i)}\right)$, $g_{k,2}\left(\mathbf{W}_{\rm cs},\mathbf{W}_{\rm cs}^{(i)}\right)$ for $k\in\mathcal{K}$ are given in (33).

For the rank-one constraint in (29g), we choose to equivalently rewrite it as [26]

$${\rm rank}(\mathbf{W}_{\rm cs})=1\Leftrightarrow{\rm Tr}(\mathbf{W}_{\rm cs})-\|\mathbf{W}_{\rm cs}\|_2=0, \tag{34}$$

where $\|\mathbf{W}_{\rm cs}\|_2$ denotes the spectral norm and is a convex function w.r.t. $\mathbf{W}_{\rm cs}$. It is worth noting that for any positive semidefinite matrix $\mathbf{A}$, $\eta_{\rm cs}(\mathbf{A})\triangleq{\rm Tr}(\mathbf{A})-\|\mathbf{A}\|_2\ge 0$ always holds and the equality is satisfied if and only if ${\rm rank}(\mathbf{A})=1$. Thus, based on the non-negative characteristic of $\eta_{\rm cs}(\mathbf{W}_{\rm cs})$, we add it into the objective function as a penalty term for the rank-one constraint which is subtracted by the objective function. However, the objective function with the penalty term is non-concave and cannot be addressed by convex optimization algorithms directly. To tackle this issue, the spectral norm in

$$g_1\big(\mathbf{W}_{\rm cs},\mathbf{W}_{\rm cs}^{(t)}\big)=\log_2\big(\,{\rm Tr}(\mathbf{W}_{\rm cs}^{(i)}\tilde{\mathbf{A}})+\sigma_{\rm b}^2\big)+\frac{{\rm Tr}(\mathbf{W}_{\rm cs}\tilde{\mathbf{A}})-{\rm Tr}(\mathbf{W}_{\rm cs}^{(i)}\tilde{\mathbf{A}})}{\ln 2\big(\,{\rm Tr}(\mathbf{W}_{\rm cs}^{(i)}\tilde{\mathbf{A}})+\sigma_{\rm b}^2\big)}, \tag{33a}$$

$$g_{k,1}\big(\mathbf{W}_{\rm cs},\mathbf{W}_{\rm cs}^{(i)}\big)=\log_2\big(\,{\rm Tr}(\mathbf{W}_{\rm cs}\breve{\mathbf{B}}_k)+\sigma_k^2\big)+\frac{{\rm Tr}(\mathbf{W}_{\rm cs}\breve{\mathbf{B}}_k)-{\rm Tr}(\mathbf{W}_{\rm cs}^{(i)}\breve{\mathbf{B}}_k)}{\ln 2\big(\,{\rm Tr}(\mathbf{W}_{\rm cs}^{(i)}\breve{\mathbf{B}}_k)+\sigma_k^2\big)}, \tag{33b}$$

$$g_2\big(\mathbf{W}_{\rm cs},\mathbf{W}_{\rm cs}^{(i)}\big)=\frac{{\rm Tr}(\mathbf{W}_{\rm cs}\breve{\mathbf{C}})-{\rm Tr}(\mathbf{W}_{\rm cs}^{(i)}\breve{\mathbf{C}})}{\ln 2\big(\,{\rm Tr}(\mathbf{W}_{\rm cs}^{(i)}\breve{\mathbf{C}})\big)}+\log_2\big(\,{\rm Tr}(\mathbf{W}_{\rm cs}^{(i)}\breve{\mathbf{C}})\big), \tag{33c}$$

$$g_{k,2}\big(\mathbf{W}_{\rm cs},\mathbf{W}_{\rm cs}^{(i)}\big)=\log_2\big(\,{\rm Tr}(\mathbf{W}_{\rm cs}\tilde{\mathbf{B}}_k)+\sigma_k^2\big)+\frac{{\rm Tr}(\mathbf{W}_{\rm cs}\tilde{\mathbf{B}}_k)-{\rm Tr}(\mathbf{W}_{\rm cs}^{(i)}\tilde{\mathbf{B}}_k)}{\ln 2\big(\,{\rm Tr}(\mathbf{W}_{\rm cs}^{(i)}\tilde{\mathbf{B}}_k)+\sigma_k^2\big)}, \tag{33d}$$

$$g_3\big(\mathbf{W}_{\rm cs},\mathbf{W}_{\rm cs}^{(i)}\big)=\frac{{\rm Tr}(\mathbf{W}_{\rm cs}\widehat{\mathbf{C}})-{\rm Tr}(\mathbf{W}_{\rm cs}^{(i)}\widehat{\mathbf{C}})}{\ln 2\big(\,{\rm Tr}(\mathbf{W}_{\rm cs}^{(i)}\widehat{\mathbf{C}})\big)}+\log_2\big(\,{\rm Tr}(\mathbf{W}_{\rm cs}^{(i)}\widehat{\mathbf{C}})\big). \tag{33e}$$

$\eta_{\rm cs}(\mathbf{W}_{\rm cs})$ is replaced by its linear lower bound obtained by its first-order Taylor expansion. Hence, we can obtain the upper bound of $\eta_{\rm cs}(\mathbf{W}_{\rm cs})$, which is expressed as

$$\eta_{\rm cs}(\mathbf{W}_{\rm cs})\le\widehat{\eta}_{\rm cs}(\mathbf{W}_{\rm cs})\triangleq{\rm Tr}(\mathbf{W}_{\rm cs})-\Big(\|\mathbf{W}_{\rm cs}^{(i)}\|_2+{\rm Tr}\big(\mathbf{w}_{\rm cs}^{(i)}\big(\mathbf{w}_{\rm cs}^{(i)}\big)^H\big(\mathbf{W}_{\rm cs}-\mathbf{W}_{\rm cs}^{(i)}\big)\big)\Big), \tag{35}$$

where $\mathbf{w}_{\rm cs}^{(i)}$ represents the eigenvectors corresponding to the largest eigenvalues of $\mathbf{W}_{\rm cs}^{(i)}$ in i-th inner loop iteration. Thus, the objective function with $\widehat{\eta}_{\rm cs}(\mathbf{W}_{\rm cs})$ will be adopted to calculate the $(i+1)$-th solution, denoted as $\mathbf{W}_{\rm cs}^{(i+1)}$. According to the above analysis, the optimization problem (29) can be further transformed as

$$\max_{\mathbf{W}_{\rm cs},\iota_{\mathbf{w}},\kappa_{\mathbf{w}},\varpi_{\mathbf{w}}}\ P_1\iota_{\mathbf{w}}+P_1\kappa_{\mathbf{w}}+P_0\varpi_{\mathbf{w}}-\varrho_{\rm cs}\widehat{\eta}_{\rm cs}(\mathbf{W}_{\rm cs}),$$

$$\text{s.t.}\quad (29a),(29b),(29c), \tag{36a}$$

$$\widehat{f}\big(\mathbf{W}_{\rm cs},\mathbf{W}_{\rm cs}^{(i)}\big)\ge\iota_{\mathbf{w}}, \tag{36b}$$

$$\widehat{f}_{k,1}\big(\mathbf{W}_{\rm cs},\mathbf{W}_{\rm cs}^{(i)}\big)\ge\varpi_{\mathbf{w}},\ \forall k, \tag{36c}$$

$$\widehat{f}_{k,2}\big(\mathbf{W}_{\rm cs},\mathbf{W}_{\rm cs}^{(i)}\big)\ge\kappa_{\mathbf{w}},\ \forall k, \tag{36d}$$

$$\mathbf{W}_{\rm cs}\succeq 0, \tag{36e}$$

where $\varrho_{\rm cs}$ is the penalty coefficient. The optimization problem (36) is a standard convex semidefinite programming (SDP) problem which is able to be effectively solved by the existing convex optimization tools such as CVX [34].

2) Joint Passive beamforming design for $\mathbf{\Theta}_{\rm r}$ and $\mathbf{\Theta}_{\rm t}$: After obtaining the active beamformers, we then design the passive beamforming variables $\mathbf{\Theta}_{\rm r}$ and $\mathbf{\Theta}_{\rm t}$ with the obtained $\mathbf{w}_{\rm b}$ and $\{\mathbf{w}_k\}_{k=1}^K$. Specifically, based on the original optimization problem (26) and definition of the average sum rate in (25), the optimization problem for joint designing the passive beamforming variables $\mathbf{\Theta}_{\rm r}$ and $\mathbf{\Theta}_{\rm t}$ can be expressed as

$$\max_{\mathbf{\Theta}_{\rm r},\mathbf{\Theta}_{\rm t}}\quad \overline{R}\big(\mathbf{\Theta}_{\rm r},\mathbf{\Theta}_{\rm t}\big),$$

$$\text{s.t.}\quad (26b)-(26f). \tag{37a}$$

Note that problem (37) is a non-convex optimization problem w.r.t. $\mathbf{\Theta}_{\rm r}$ and $\mathbf{\Theta}_{\rm t}$. Similarly, we will adopt the SDR techniques and introduce auxiliary variables $\iota_{\mathbf{\Theta}}$, $\kappa_{\mathbf{\Theta}}$ and $\varpi_{\mathbf{\Theta}}$ to deal with this max-min non-convex optimization problem. The covert constraint $\frac{\beta}{\alpha}\ge\varphi(\epsilon)$ is still utilized to guarantee the covert performance. Let $\mathbf{Q}_{\rm r}=\boldsymbol{\vartheta}_{\rm r}^*\boldsymbol{\vartheta}_{\rm r}^T$, $\mathbf{Q}_{\rm t}=\boldsymbol{\vartheta}_{\rm t}^*\boldsymbol{\vartheta}_{\rm t}^T$ where $\boldsymbol{\vartheta}_{\rm r}={\rm diag}(\mathbf{\Theta}_{\rm r})$, $\boldsymbol{\vartheta}_{\rm t}={\rm diag}(\mathbf{\Theta}_{\rm t})$, and then the optimization problem (37) can be equivalently reformulated as problem (38), where

- $\mathbf{V}=\{\mathbf{Q}_{\rm r},\mathbf{Q}_{\rm t},\boldsymbol{\beta}_{\rm r},\boldsymbol{\beta}_{\rm t},\iota_{\mathbf{\Theta}},\kappa_{\mathbf{\Theta}},\varpi_{\mathbf{\Theta}}\}$ is the defined optimization variable set,
- $\boldsymbol{\beta}_{\rm r}=\{\beta_{\rm r}^1,\cdots,\beta_{\rm r}^M\}$, $\boldsymbol{\beta}_{\rm t}=\{\beta_{\rm t}^1,\cdots,\beta_{\rm t}^M\}$,
- $\mathbf{E}=\sum_{l=1}^L\big(\mathbf{w}_{\rm b}^H\mathbf{\Psi}_{\rm BR}^l\mathbf{w}_{\rm b}\big)\mathbf{\Delta}^l$,
- $\mathbf{F}=\sum_{k=1}^K\sum_{l=1}^L\big(\mathbf{w}_k^H\mathbf{\Psi}_{\rm BR}^l\mathbf{w}_k\big)\mathbf{\Delta}^l$,
- $\mathbf{\Delta}^l=\mathbf{\Xi}^T\big(\big(\widehat{\mathbf{\Psi}}_{\rm BR}^l\big)^T\otimes\big(\mathbf{\Omega}_{\rm rw}^H\mathbf{\Omega}_{\rm rw}\big)\big)\mathbf{\Xi}$,
- $\mathbf{G}=\mathbf{H}_{\rm rb}^*\mathbf{H}_{\rm BR}\mathbf{w}_{\rm b}\mathbf{w}_{\rm b}^H\mathbf{H}_{\rm BR}^H\mathbf{H}_{\rm rb}^T$,
- $\mathbf{O}=\mathbf{H}_{\rm rb}^*\mathbf{H}_{\rm BR}\sum_{k=1}^K\big(\mathbf{w}_k\mathbf{w}_k^H\big)\mathbf{H}_{\rm BR}^H\mathbf{H}_{\rm rb}^T$,
- $\mathbf{P}_k=\mathbf{H}_{{\rm r}k}^*\mathbf{H}_{\rm BR}\sum_{k=1}^K\big(\mathbf{w}_k\mathbf{w}_k^H\big)\mathbf{H}_{\rm BR}^H\mathbf{H}_{{\rm r}k}^T$,
- $\hat{\mathbf{P}}_k=\mathbf{H}_{{\rm r}k}^*\mathbf{H}_{\rm BR}\sum_{j\ne k}^K\big(\mathbf{w}_j\mathbf{w}_j^H\big)\mathbf{H}_{\rm BR}^H\mathbf{H}_{{\rm r}k}^T$,
- $\mathbf{S}=\sum_{k=1}^K\mathbf{\Xi}^T\big(\big(\mathbf{\Omega}_{\rm re}^H\mathbf{\Omega}_{\rm re}\big)^T\otimes\big(\mathbf{H}_{\rm BR}\mathbf{w}_k\mathbf{w}_k^H\mathbf{H}_{\rm BR}^H\big)\big)\mathbf{\Xi}$,
- $\hat{\mathbf{S}}_k=\sum_{j\ne k}^K\mathbf{\Xi}^T\big(\big(\mathbf{\Omega}_{\rm re}^H\mathbf{\Omega}_{\rm re}\big)^T\otimes\big(\mathbf{H}_{\rm BR}\mathbf{w}_j\mathbf{w}_j^H\mathbf{H}_{\rm BR}^H\big)\big)\mathbf{\Xi}$,
- $\mathbf{T}_k=\mathbf{H}_{{\rm r}k}^*\mathbf{H}_{\rm BR}\mathbf{w}_{\rm b}\mathbf{w}_{\rm b}^H\mathbf{H}_{\rm BR}^H\mathbf{H}_{{\rm r}k}^T+\mathbf{P}_k$,
- $\hat{\mathbf{T}}_k=\mathbf{H}_{{\rm r}k}^*\mathbf{H}_{\rm BR}\mathbf{w}_{\rm b}\mathbf{w}_{\rm b}^H\mathbf{H}_{\rm BR}^H\mathbf{H}_{{\rm r}k}^T+\hat{\mathbf{P}}_k$,
- $\mathbf{U}=\mathbf{\Xi}^T\big(\big(\mathbf{\Omega}_{\rm re}^H\mathbf{\Omega}_{\rm re}\big)^T\otimes\big(\mathbf{H}_{\rm BR}\mathbf{w}_{\rm b}\mathbf{w}_{\rm b}^H\mathbf{H}_{\rm BR}^H\big)\big)\mathbf{\Xi}+\mathbf{S}$,
- $\hat{\mathbf{U}}_k=\mathbf{\Xi}^T\big(\big(\mathbf{\Omega}_{\rm re}^H\mathbf{\Omega}_{\rm re}\big)^T\otimes\big(\mathbf{H}_{\rm BR}\mathbf{w}_{\rm b}\mathbf{w}_{\rm b}^H\mathbf{H}_{\rm BR}^H\big)\big)\mathbf{\Xi}+\hat{\mathbf{S}}_k$,
- $\mathbf{H}_{\rm rb}={\rm Diag}(\mathbf{h}_{\rm rb}),\mathbf{H}_{{\rm r}k}={\rm Diag}(\mathbf{h}_{{\rm r}k})$.

To transform (38) into a convex optimization problem, we first need to deal with the non-convex constraints (38c), (38d), (38e) and rank-one constraints (38i). Similarly, the first-order Taylor expansion is adopted to acquire the concave lower bounds of left-sides of constraints (38c), (38d), (38e) in q-th inner loop iteration, denoted as $h\big(\mathbf{Q}_{\rm r},\mathbf{Q}_{\rm r}^{(q)}\big)$, $h_{k,1}\big(\mathbf{Q}_{\rm t},\mathbf{Q}_{\rm t}^{(q)}\big)$ and $h_{k,2}\big(\mathbf{Q}_{\rm t},\mathbf{Q}_{\rm t}^{(q)}\big)$. For the rank-one constraints, we rewrite them as the expressions similar to (35) and add them to the objective function as the penalty terms. Similarly, the linear lower bound of the spectral norm is utilized to replace itself. As a result, the rank-one can be equivalently transformed as

$$\eta_\xi(\mathbf{Q}_\xi)\le\widehat{\eta}_\xi\big(\mathbf{Q}_\xi\big)\triangleq{\rm Tr}(\mathbf{Q}_\xi)-\|\mathbf{Q}_\xi^{(q)}\|_2-{\rm Tr}\big(\mathbf{q}_\xi^{(q)}\big(\mathbf{q}_\xi^{(q)}\big)^H\big(\mathbf{Q}_\xi-\mathbf{Q}_\xi^{(q)}\big)\big),\ \xi\in\{{\rm r},{\rm t}\}, \tag{39}$$

where $\mathbf{q}_{\rm r}^{(q)}$ and $\mathbf{q}_{\rm t}^{(q)}$ are the eigenvectors of the largest eigenvalues of $\mathbf{Q}_{\rm r}^{(q)}$ and $\mathbf{Q}_{\rm t}^{(q)}$ in q-th inner loop iteration. Thus,

$$
\begin{aligned}
\max_{\mathbf{V}}\ & P_1\iota_{\mathbf{\Theta}} + P_1\kappa_{\mathbf{\Theta}} + P_0\varpi_{\mathbf{\Theta}}, \\
\text{s.t. }\ & \mathrm{Tr}(\mathbf{Q}_{\mathrm{r}}\mathbf{F}) \geq \mathrm{Tr}(\mathbf{Q}_{\mathrm{r}}\mathbf{E})\varphi(\epsilon), && \text{(38a)}\\
& \iota_{\mathbf{\Theta}} \geq R_{\mathrm{b}}^{*}, \varpi_{\mathbf{\Theta}} \geq R_{\mathrm{s0}}^{*}, \kappa_{\mathbf{\Theta}} \geq R_{\mathrm{s1}}^{*}, && \text{(38b)}\\
& \log_2\left(\mathrm{Tr}(\mathbf{Q}_{\mathrm{r}}\mathbf{G}) + \mathrm{Tr}(\mathbf{Q}_{\mathrm{r}}\mathbf{O}) + \sigma_{\mathrm{b}}\right) - \log_2\left(\mathrm{Tr}(\mathbf{Q}_{\mathrm{r}}\mathbf{O}) + \sigma_{\mathrm{b}}\right) \geq \iota_{\mathbf{\Theta}}, && \text{(38c)}\\
& \log_2\left(\mathrm{Tr}(\mathbf{Q}_{\mathrm{t}}\mathbf{P}_k) + \sigma_k^2\right) - \log_2\left(\mathrm{Tr}(\mathbf{Q}_{\mathrm{t}}\check{\mathbf{P}}_k) + \sigma_k^2\right) - \log_2\left(\mathrm{Tr}(\mathbf{Q}_{\mathrm{t}}\mathbf{S})\right) + \log_2\left(\mathrm{Tr}\left(\mathbf{Q}_{\mathrm{t}}\check{\mathbf{S}}_k\right)\right) \geq \varpi_{\mathbf{\Theta}},\ \forall k, && \text{(38d)}\\
& \log_2\left(\mathrm{Tr}(\mathbf{Q}_{\mathrm{t}}\mathbf{T}_k) + \sigma_k^2\right) - \log_2\left(\mathrm{Tr}(\mathbf{Q}_{\mathrm{t}}\check{\mathbf{T}}_k) + \sigma_k^2\right) - \log_2\left(\mathrm{Tr}(\mathbf{Q}_{\mathrm{t}}\mathbf{U})\right) + \log_2\left(\mathrm{Tr}(\mathbf{Q}_{\mathrm{t}}\check{\mathbf{U}}_k)\right) \geq \kappa_{\mathbf{\Theta}},\ \forall k, && \text{(38e)}\\
& \mathrm{diag}(\mathbf{Q}_{\mathrm{r}}) = \boldsymbol{\beta}_{\mathrm{r}}, \mathrm{diag}(\mathbf{Q}_{\mathrm{t}}) = \boldsymbol{\beta}_{\mathrm{t}}, && \text{(38f)}\\
& \boldsymbol{\beta}_{\mathrm{r}} + \boldsymbol{\beta}_{\mathrm{t}} = \mathbf{I}_{M\times 1}, && \text{(38g)}\\
& \mathbf{Q}_{\mathrm{r}} \succeq 0, \mathbf{Q}_{\mathrm{t}} \succeq 0, && \text{(38h)}\\
& \mathrm{rank}(\mathbf{Q}_{\mathrm{r}}) = 1, \mathrm{rank}(\mathbf{Q}_{\mathrm{t}}) = 1. && \text{(38i)}
\end{aligned}
$$

optimization problem (38) can be re-expressed as

$$
\begin{aligned}
\max_{\mathbf{V}}\ & P_1\iota_{\mathbf{\Theta}} + P_1\kappa_{\mathbf{\Theta}} + P_0\varpi_{\mathbf{\Theta}} - \varrho_{\mathrm{r}}\widehat{\eta}_{\mathrm{r}}(\mathbf{Q}_{\mathrm{r}}) - \varrho_{\mathrm{t}}\widehat{\eta}_{\mathrm{t}}(\mathbf{Q}_{\mathrm{t}}), \\
\text{s.t. }\ & \text{(38a)}, \text{(38b)}, \text{(38f)} - \text{(38h)}, && \text{(40a)}\\
& h\left(\mathbf{Q}_{\mathrm{r}}, \mathbf{Q}_{\mathrm{r}}^{(q)}\right) \geq \iota_{\mathbf{\Theta}}, && \text{(40b)}\\
& h_{k,1}\left(\mathbf{Q}_{\mathrm{r}}, \mathbf{Q}_{\mathrm{r}}^{(q)}\right) \geq \varpi_{\mathbf{\Theta}},\ \forall k, && \text{(40c)}\\
& h_{k,2}\left(\mathbf{Q}_{\mathrm{r}}, \mathbf{Q}_{\mathrm{r}}^{(q)}\right) \geq \kappa_{\mathbf{\Theta}},\ \forall k, && \text{(40d)}
\end{aligned}
$$

where ϱ_{r} and ϱ_{t} denote the penalty coefficients. Thus, SDP optimization problem (40) can be efficiently solved by CVX.

C. Proposed Optimization Algorithm & Analysis on Complexity and Convergence

Algorithm 1 summarizes the proposed iterative algorithm for solving the optimization problem (26) of the STAR-RIS-assisted joint PLS and CC system. The algorithm alternatively solves two subproblems and converges when the objective function gap $v > 0$ between two consecutive iterations is below a predefined threshold ε. The penalty violations for active and passive beamforming designs are denoted by $\widehat{v} > 0$ and $\tilde{v} > 0$, respectively. Note that, the penalty coefficients ϱ_{cs}, ϱ_{r} and ϱ_{t} are initialized with small values to prevent the penalty terms from dominating the objective function and leading to inefficient solutions. In addition, $\widehat{\xi}$, $\tilde{\xi}_1$ and $\tilde{\xi}_2$ are the scaling factors for penalty coefficients.

In terms of the computational complexity of the proposed algorithm, it is mainly dominated by addressing the two standard SDP subproblems[2]. Specifically, for the joint active beamforming design, the main computed complexity on solving the optimization problem (36) can be calculated as $\mathcal{O}\big(\big((K+1)N_{\mathrm{t}}\big)^{3.5}\big)$. In the aspect of joint design the passive beamformer, the calculated complexity comes from the solving of the optimization problem (40), which is dominated by $\mathcal{O}\big(2M^{3.5}\big)$. In addition, the bisection search method is utilized to find $\varphi(\epsilon)$ to transform the covert constraint (26b) with the computational complexity is $\mathcal{O}\big(\log_2\left(\frac{s}{\varepsilon_{\mathrm{b}}}\right)\big)$, where s and ε_{b} denote the length of the initial search interval and the accuracy tolerance, respectively. Therefore, the overall computational complexity of the proposed iterative algorithm can be calculated as $\mathcal{O}\big(\big(\log_2\left(\frac{s}{\varepsilon_{\mathrm{b}}}\right) + I\big(I_1\big((K+1)N_{\mathrm{t}}\big)^{3.5} + I_2\left(2M^{3.5}\right)\big)\big)\big)$, where I denotes the total iteration number of the proposed algorithm, I_1 and I_2 respectively represent the iteration number of the inner loops for solving two subproblems. Note that the overall computational complexity is highly affected by the number of antennas at BS (N_{t}) and the number of elements equipped at STAR-RIS (M).

Although the alternative strategy is adopted in Algorithm 1, it is easy to verify that the convergence of the proposed iterative algorithm can always be guaranteed. Note that we can always find a solution not worse than that of the previous iteration, and thus the objective function value of the optimization problem (26) is monotonically non-decreasing w.r.t. the iteration. Moreover, the convergence of the proposed algorithm will be further proved by simulation results in Section V.

V. Simulation Results

In this section, the simulation results are presented to validate the effectiveness of the proposed STAR-RIS-aided joint PLS and CCs scheme. In particular, we assume that the mmWave communication system assisted by STAR-RIS operates at 28 GHz with bandwidth 251.1886 MHz. Hence, the noise power can be calculated as $\sigma_{\mathrm{b}}^2 = -90$ dBm and $\sigma_k^2 = -90$ dBm. In addition, we consider that the simulated system has $K = 3$ security users and set the QoS minimum rates as $R_{\mathrm{b}}^{*} = 0.5$, $R_{\mathrm{s0}}^{*} = 0.6$ and $R_{\mathrm{s1}}^{*} = 0.6$. For the large-scale path loss values in (1) and (2), the theoretical free-space distance-dependent path-loss model [36] is leveraged, which is given by $l_{\varpi} = -30 - 22\log d_{\varpi}$ dB, $\varpi \in \{\mathrm{BR}, \mathrm{rb}, \mathrm{r}k\}$ for $k \in \mathcal{K}$. The distances are set as $d_{\mathrm{BR}} = 40$ m, $d_{\mathrm{rb}} = 15$ m and $d_{\mathrm{r}k} = 15$ m. Moreover, the tolerance accuracy ε, $\widehat{\varepsilon}$ and $\tilde{\varepsilon}$ in the proposed iterative algorithm are set as 10^{-4}, 10^{-6} and 10^{-6}, respectively. To highlight the potential of STAR-RIS in jointly implementing the PLS and CCs, and the effectiveness of the proposed optimization algorithm, we compare the obtained results with three benchmark schemes, including: 1) **RIS-aided scheme [23]:** In this baseline scheme, two adjacent conventional RISs with $\frac{M}{2}$ elements where one is

[2]In the case of solving convex problems, it is presumed that the interior point method is employed, and subsequently, the computational complexity is determined [35].

the reflecting-only RIS and the other one is the transmission-only RIS are adopted to replace the STAR-RIS; 2) **Maximum ratio transmission (MRT) scheme [37]:** In this baseline scheme, we perform the MRT to obtain the active beamforming vector $\mathbf{w}_k = \frac{\sqrt{P_k}(\mathbf{h}_{\mathrm{r}k}^H \mathbf{\Theta}_{\mathrm{t}} \mathbf{H}_{\mathrm{BR}})^H}{\left\|\mathbf{h}_{\mathrm{r}k}^H \mathbf{\Theta}_{\mathrm{t}} \mathbf{H}_{\mathrm{BR}}\right\|}$, where P_k is the allocated power for k-th secure user and $\mathbf{w}_{\mathrm{b}}$, $\mathbf{\Theta}_{\mathrm{r}}$ and $\mathbf{\Theta}_{\mathrm{t}}$ are obtained by the proposed scheme; 3) **Zero-forcing (ZF) scheme [38]**: Similarly, the ZF algorithm is utilized to obtain $\mathbf{w}_k$, and then $\mathbf{w}_{\mathrm{b}}$, $\mathbf{\Theta}_{\mathrm{r}}$ and $\mathbf{\Theta}_{\mathrm{t}}$ are solved by the proposed algorithm.

Algorithm 1: Proposed Iterative Algorithm for STAR-RIS-assisted joint PLS and CCs Problem (26)

1: Initialize feasible point $\left(\mathbf{w}_{\mathrm{b}}^{(0,0)}, \mathbf{w}_k^{(0,0)}, \mathbf{\Theta}_{\mathrm{r}}^{(0,0)}, \mathbf{\Theta}_{\mathrm{t}}^{(0,0)}\right)$; Define the tolerance accuracy thresholds ε, $\widehat{\varepsilon}$ and $\tilde{\varepsilon}$; Set the outer iteration index t = 0.
2: **While** $v > \varepsilon$ or $t = 0$ **do**
3: Set inner iteration index i = 0; Initialize $\varrho_{\mathrm{cs}}^{(0)}$.
4: **While** $\widehat{v} > \widehat{\varepsilon}$ or $i = 0$ **do**
5: Solve the optimization problem (36) with the given $\left(\mathbf{w}_{\mathrm{b}}^{(t,i)}, \mathbf{w}_k^{(t,i)}, \mathbf{\Theta}_{\mathrm{r}}^{(t,0)}, \mathbf{\Theta}_{\mathrm{t}}^{(t,0)}\right)$ and update $\left(\mathbf{w}_{\mathrm{b}}^{(t,i+1)}, \mathbf{w}_k^{(t,i+1)}\right)$ with the obtained solutions.
6: Calculate $\widehat{v} = \eta_{\mathrm{cs}}$ based on the acquired solutions; Update penalty coefficients $\varrho_{\mathrm{cs}} = \widehat{\xi}\varrho_{\mathrm{cs}}$; Let $i = i+1$.
7: **end while**
8: Update $\left(\mathbf{w}_{\mathrm{b}}^{(t,0)}, \mathbf{w}_k^{(t,0)}\right)$ with the $\left(\mathbf{w}_{\mathrm{b}}^{(t,i)}, \mathbf{w}_k^{(t,i)}\right)$.
9: Set inner iteration index q = 0; Initialize $\varrho_{\mathrm{r}}^{(0)}$ and $\varrho_{\mathrm{t}}^{(0)}$.
10: **While** $\tilde{v} > \tilde{\varepsilon}$ or $q = 0$ **do**
11: Solve the optimization problem (40) with the given $\left(\mathbf{w}_{\mathrm{b}}^{(t,0)}, \mathbf{w}_k^{(t,0)}, \mathbf{\Theta}_{\mathrm{r}}^{(t,q)}, \mathbf{\Theta}_{\mathrm{t}}^{(t,q)}\right)$; Update the $\left(\mathbf{\Theta}_{\mathrm{r}}^{(t,q+1)}, \mathbf{\Theta}_{\mathrm{t}}^{(t,q+1)}\right)$ with obtained solutions.
12: Calculate $\tilde{v} = \max\{\eta_{\mathrm{r}}, \eta_{\mathrm{t}}\}$ based on the acquired solution; Update the penalty coefficients $\varrho_{\mathrm{r}}^{(q+1)} = \tilde{\xi}_1 \varrho_{\mathrm{r}}^{(q)}$, $\varrho_{\mathrm{t}}^{(q+1)} = \tilde{\xi}_2 \varrho_{\mathrm{t}}^{(q)}$; Let $q = q + 1$.
13: **end while**
14: Update $\left(\mathbf{w}_{\mathrm{b}}^{(t+1,0)}, \mathbf{w}_k^{(t+1,0)}, \mathbf{\Theta}_{\mathrm{r}}^{(t+1,0)}, \mathbf{\Theta}_{\mathrm{t}}^{(t+1,0)}\right)$ with $\left(\mathbf{w}_{\mathrm{b}}^{(t,0)}, \mathbf{w}_k^{(t,0)}, \mathbf{\Theta}_{\mathrm{r}}^{(t,q)}, \mathbf{\Theta}_{\mathrm{t}}^{(t,q)}\right)$
15: Calculate the objective value $\overline{R}^{(t+1)}$ and update $v = \left|\overline{R}^{(t+1)} - \overline{R}^{(t)}\right|$ based the obtained solutions; Let $t = t + 1$.
16: **end while**

The convergence curves of the proposed iterative algorithm are depicted in Fig. 2, taking into account of various maximum transmit power at the BS, as well as the number of elements and antennas equipped at the STAR-RIS and the BS. Specifically, we conducts evaluation of convergence using eight diverse cases for the proposed method. The presented results indicate that the obtained sum rates exhibit a monotonically non-decreasing behavior versus the number of iterations. In addition, the proposed algorithm consistently achieves rapid convergence to a stable value within a few iterations. Hence, the efficiency of the proposed algorithm can be validated.

Fig. 3 presents the variation curves of the average sum rates versus the maximal transmit power P_{tmax} with different covert

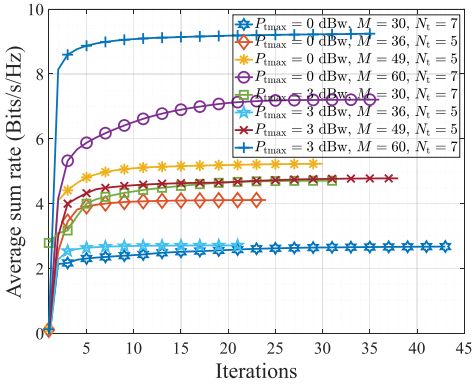


Fig. 2. Average sum rate versus the iterations with $\epsilon = 0.1$ and $P_1 = 0.5$, and different P_{tmax}, M and N_{t}.

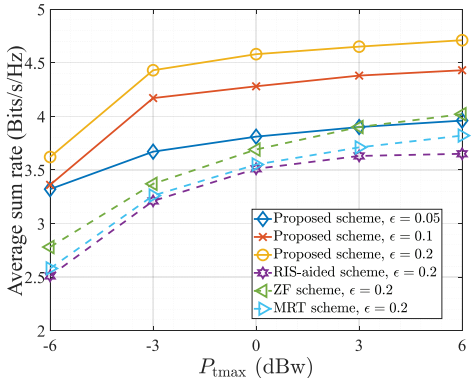


Fig. 3. Average sum rate versus the maximum transmit power P_{tmax} at BS with $M = 30$, $N_{\mathrm{t}} = 7$, $P_1 = 0.5$, and different covert requirements ϵ.

requirements ϵ, in comparison with the baselines respectively utilizing the traditional RIS, ZF algorithm and MRT algorithm. It can be observed that the average sum rates gradually increase w.r.t. P_{tmax} in all cases, indicating that there exists a positive correlation between the average sum rates and P_{tmax}. However, the speeds of increase diminish with the growth of the maximum transmit power. Additionally, a relaxed covert requirement contributes to breaking through the performance bottleneck constrained by other system indicators. It is obvious that the proposed scheme exhibits significant performance benefits in jointly implementing the PLS and CCs in comparison to the baseline schemes. Even if the proposed scheme is operated at a tighter covert requirement (i.e., $\epsilon = 0.05$), it can still achieve better performance.

Next, we investigate the influence of the covert requirements ϵ on the average sum rate considering different P_{tmax}, as shown in Fig. 4. According to the given results, we can find that the average sum rates increase progressively versus ϵ in all scenarios due to the fact that the covert requirement becomes more relaxed. To acquire an apparent comparison, $P_{\mathrm{tmax}} = 3$ dBw is selected to operate the baseline schemes. Despite this, the achieved performance gain of the baselines falls significantly short to that of the proposed scheme, even if the proposed scheme is operated at a much lower maximum transmit power of $P_{\mathrm{tmax}} = -3$ dBw. Furthermore, both the ZF scheme and MRT scheme exhibit superior performance gains compared to the RIS-aided scheme when operating under

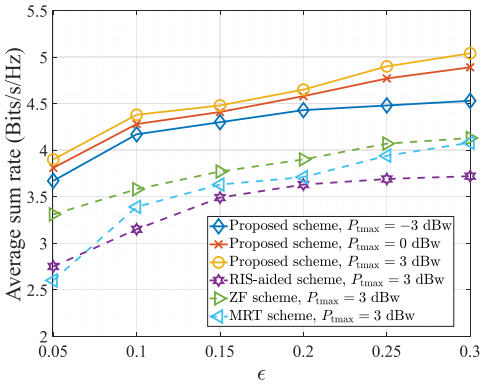


Fig. 4. Average sum rate versus the covert requirement ϵ with $M = 30$, $N_{\text{t}} = 7$, $P_1 = 0.5$, and different P_{tmax}.

identical conditions. The results demonstrate that the STAR-RIS-aided scheme offers a significant benefit in improving system performance as compared to the conventional RIS, and the proposed iterative algorithm proves to be successful in addressing the proposed optimization problem.

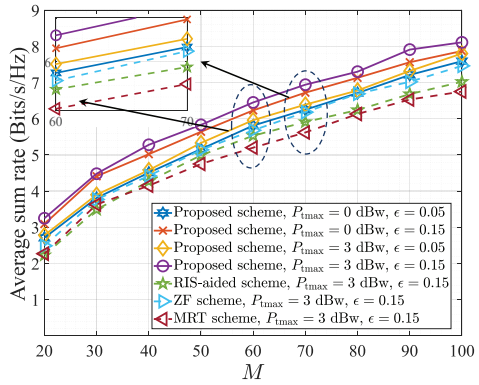


Fig. 5. Average sum rate versus the number of elements equipped at STAR-RIS with $N_{\text{t}} = 7$, $P_1 = 0.5$, and different maximum transmit power P_{tmax} and covert requirements ϵ.

In Fig. 5, the performance trends of the average sum rate w.r.t. the number of elements at STAR-RIS (M) are presented, taking into account of various P_{tmax} and covert requirements ϵ. In particular, it is discernible that the average sum rates exhibit ascending trends with the increased M, which is due to the fact that more elements can provide a higher degree of freedom to augment performance gains. Besides, we also find that the relaxed covert demands may offer more potential to break through the performance limitation imposed by the system settings than the incremental maximal transmit power by respectively comparing the simulation results with the same P_{tmax} and ϵ. Similarly, the most relaxed condition (i.e., $P_{\text{tmax}} = 3$ dBw, $\epsilon = 0.15$) is adopted to implement the baseline schemes (RIS-aided, ZF and MRT), however, the acquired performance is still worse than the proposed scheme under the strictest condition (i.e., $P_{\text{tmax}} = 0$ dBw, $\epsilon = 0.05$). Likewise, the ZF scheme consistently outperforms the RIS-aided scheme. However, as the number of elements equipped at STAR-RIS increases, the performance of the MRT scheme gradually falls below that of the RIS-aided scheme. The reason for this is that the principle of the MRT scheme is to boost the desired signal but ignore interference among users. Nevertheless, as M increases, the relationship among users in the channel will also grow, resulting in an elevated presence of interference among users.

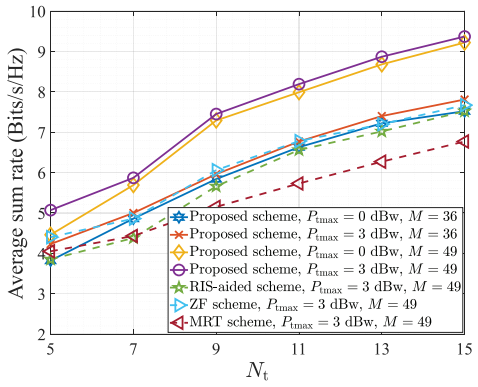


Fig. 6. Average sum rate versus the number of antennas at BS with $\epsilon = 0.1$, $P_1 = 0.5$, and different maximal transmit power P_{tmax} and M.

We explore the impact of the number of antennas installed at the BS (N_{t}) on the system performance in Fig. 6 with different P_{tmax} and M. Specifically, a comparable performance trend can still be noted, wherein the average sum rate gradually raises as N_{t} is augmented. In addition, we can observe that increasing the number of elements at the STAR-RIS from $M = 36$ to $M = 49$ can achieve much more performance improvement than enlarging P_{tmax} from 0 dBw to 3 dBw. Although we choose $P_{\text{tmax}} = 3$ dBw and $M = 49$ to operate all baseline schemes, the presented performance gain is still far below the proposed scheme under the same condition (i.e., $P_{\text{tmax}} = 3$ dBw, $M = 49$), which further indicates the superiority of STAR-RIS in ensuring the performance of joint PLS and CCs, and the effectiveness of the proposed algorithm. Additionally, the three baseline schemes yielded consistent performance trends as observed during the investigation of performance gains with varying M.

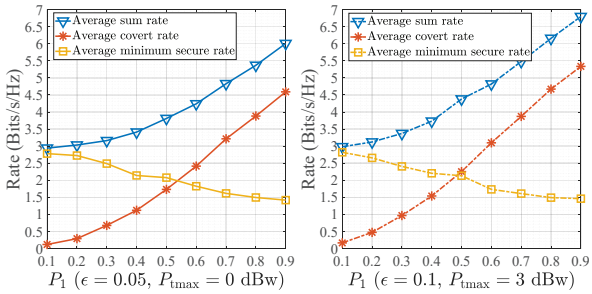


Fig. 7. Average sum rate, average covert rate and average minimum secure rate versus the covert transmission probability P_1 with $M = 30$, $N_{\text{t}} = 7$, and different P_{tmax} and ϵ.

Lastly, we investigate the influence of the probability for CCs to Bob, i.e., P_1, on the average sum rate, the average covert rate ($P_1 R_{\text{b}}^{\text{c}}$), and the average minimum secure rate ($P_0 \min_k \widehat{R}_{\text{sl},0}^k + P_1 \min_k \widehat{R}_{\text{sl},1}^k$), in a time slot, as presented in Fig. 7. According to the simulated results, it can be concluded that an increase in the value of P_1 from 0.1 to 0.9 leads to an

upward trend in both the average sum rate and the average covert rate. Conversely, the average minimum secure rate displays a downward trend, suggesting that the improvement in the average sum rate versus P_1 is mainly attributed to the average covert rate. The rationale behind this phenomenon is that the energy splitting protocol is adopted at the STAR-RIS. As a result, more amplitude energy will be allocated to the reflected coefficients while less to the transmitted coefficients as the CCs probability P_1 increases, leading to an increase of the covert rate while a decrease of the secure rate.

VI. Conclusion

In this paper, we initially investigate the STAR-RIS enhanced joint PLS and CCs for mmWave systems. In particular, the analytical derivations of the minimum DEP and the lower bound of the secure rates are obtained by considering the practical assumptions, where only the statistical CSI between STAR-RIS and the wardens is accessible at the BS. An optimization problem is constructed that focuses on maximizing the average sum rate between the covert rate and the minimum secure rate, while also ensuring the covert constraint and QoS constraints. In order to effectively solve this non-convex optimization problem with strong coupling variables, an alternative algorithm based on the SDR method is proposed. Numerical results demonstrate the performance gains of the proposed STAR-RIS-assisted scheme in comparison with the benchmark scheme adopting the traditional RIS, which further indicates that the STAR-RIS exhibits more benefits in the implementation of the joint PLS and CCs.

Appendix A
Proof of Theorem 1

In this section, we derive the FA probability and MD probability in detail. In particular, we first give the expression of P_{FA} based on the definition of FA probability, given as

$$\begin{aligned} P_{\mathrm{FA}} &= \Pr\left(\overline{P}_{\mathrm{w}} > \tau_{dt} | \mathcal{H}_0\right) \\ &= \Pr\left(\sum_{k=1}^{K} \left|\mathbf{h}_{\mathrm{rw}}^H \boldsymbol{\Theta}_{\mathrm{r}} \mathbf{H}_{\mathrm{BR}} \mathbf{w}_k\right|^2 + \sigma_{\mathrm{w}}^2 > \tau_{\mathrm{dt}}\right). \end{aligned} \tag{41}$$

Considering only the statistical CSI of $\mathbf{H}_{\mathrm{BR}}$ is available at Willie, thus letting $\chi_k = \left|\mathbf{h}_{\mathrm{rw}}^H \boldsymbol{\Theta}_{\mathrm{r}} \mathbf{H}_{\mathrm{BR}} \mathbf{w}_k\right|^2$, we have

$$\begin{aligned} \chi_k &= \left|\mathrm{Tr}(\mathbf{h}_{\mathrm{rw}}^H \boldsymbol{\Theta}_{\mathrm{r}} \mathbf{H}_{\mathrm{BR}} \mathbf{w}_k)\right|^2 = \left|\mathrm{Tr}(\mathbf{H}_{\mathrm{BR}} \mathbf{w}_k \mathbf{h}_{\mathrm{rw}}^H \boldsymbol{\Theta}_{\mathrm{r}})\right|^2 \\ &\overset{(a)}{=} \left|\mathrm{vec}(\mathbf{H}_{\mathrm{BR}})^T \mathrm{vec}\left((\mathbf{w}_k \mathbf{h}_{\mathrm{rw}}^H \boldsymbol{\Theta}_{\mathrm{r}})^T\right)\right|^2, \end{aligned} \tag{42}$$

where (a) is due to [37, (eq.1.11.12)]. Besides, it is easy to verify that $\mathrm{vec}(\mathbf{H}_{\mathrm{BR}}) \sim \mathcal{CN}(\mathbf{0}, \frac{N_{\mathrm{t}} M \rho_{\mathrm{BR}}}{L} \boldsymbol{\Phi}^H \boldsymbol{\Phi})$. Hence, we can observe that χ_k is an exponential random variable whose probability density function (PDF) is $f_{\chi_k}(x) = \frac{e^{\frac{-x}{\lambda_k}}}{\lambda_k}$, where $\lambda_k = \frac{N_{\mathrm{t}} M \rho_{\mathrm{BR}}}{L} \left\|\boldsymbol{\Phi}\, \mathrm{vec}\left((\mathbf{w}_k \mathbf{h}_{\mathrm{rw}}^H \boldsymbol{\Theta}_{\mathrm{r}})^T\right)\right\|_2^2$.

According to the above analysis, the analytic expressions of P_{FA} and P_{MD} can be derived as [28]

$$P_{\mathrm{FA}} = \begin{cases} 1, \\ \int_{\tau_{\mathrm{dt}} - \sigma_{\mathrm{w}}^2}^{+\infty} \frac{e^{-\frac{x}{\lambda_0}}}{\lambda_0} dx, \end{cases} = \begin{cases} 1, & \tau_{\mathrm{dt}} \le \sigma_{\mathrm{w}}^2, \\ e^{-\frac{\tau_{\mathrm{dt}} - \sigma_{\mathrm{w}}^2}{\lambda_0}}, & \text{otherwise}, \end{cases}$$

$$\begin{aligned} P_{\mathrm{MD}} &= \Pr\Big(\left|\mathbf{h}_{\mathrm{rw}}^H \boldsymbol{\Theta}_{\mathrm{r}} \mathbf{H}_{\mathrm{BR}} \mathbf{w}_{\mathrm{b}}\right|^2 + \sum_{k=1}^{K} \left|\mathbf{h}_{\mathrm{rw}}^H \boldsymbol{\Theta}_{\mathrm{r}} \mathbf{H}_{\mathrm{BR}} \mathbf{w}_k\right|^2 \\ &\quad + \sigma_{\mathrm{w}}^2 < \tau_{\mathrm{dt}} \Big) \\ &= \begin{cases} 0, \\ \int_0^{\tau_{\mathrm{dt}} - \sigma_{\mathrm{w}}^2} \frac{e^{-\frac{x}{\lambda_1}}}{\lambda_1} dx, \end{cases} = \begin{cases} 0, & \tau_{\mathrm{dt}} \le \sigma_{\mathrm{w}}^2, \\ 1 - e^{-\frac{\tau_{\mathrm{dt}} - \sigma_{\mathrm{w}}^2}{\lambda_1}}, & \text{otherwise}. \end{cases} \end{aligned}$$

Appendix B
Proof of Theorem 2

The asymptotic results of λ_0 and λ_1 leveraging the large system analytic technique are derived in this section. Specifically, we first equivalently rewrite $\psi = \left\|\boldsymbol{\Phi}\, \mathrm{vec}\left((\mathbf{w}_k \mathbf{h}_{\mathrm{rw}}^H \boldsymbol{\Theta}_{\mathrm{r}})^T\right)\right\|_2^2$ as

$$\begin{aligned} \psi &= \mathrm{vec}\left((\mathbf{w}_k \mathbf{h}_{\mathrm{rw}}^H \boldsymbol{\Theta}_{\mathrm{r}})^T\right)^H \boldsymbol{\Phi}^H \boldsymbol{\Phi}\, \mathrm{vec}\left((\mathbf{w}_k \mathbf{h}_{\mathrm{rw}}^H \boldsymbol{\Theta}_{\mathrm{r}})^T\right) \\ &= \left\|\boldsymbol{\Phi}\, \mathrm{vec}\left((\mathbf{w}_k \mathbf{h}_{\mathrm{rw}}^H \boldsymbol{\Theta}_{\mathrm{r}})^T\right)\right\|_2^2 \overset{(a)}{=} \left\|\boldsymbol{\Phi}\left(\mathbf{w}_k \otimes \boldsymbol{\Theta}_{\mathrm{r}}\right) \mathbf{h}_{\mathrm{rw}}^*\right\|_2^2 \\ &= \frac{M \rho_{\mathrm{rw}}}{P} \left\|\boldsymbol{\Phi}\left(\mathbf{w}_k \otimes \boldsymbol{\Theta}_{\mathrm{r}}\right) \boldsymbol{\Omega}_{\mathrm{rw}}^H \mathbf{g}^*\right\|_2^2, \end{aligned} \tag{43}$$

where $\boldsymbol{\Omega}_{\mathrm{rw}} = \left[\mathbf{a}_{\mathrm{R}}\left(\phi_1^{\mathrm{rw}}, \theta_1^{\mathrm{rw}}\right), \cdots, \mathbf{a}_{\mathrm{R}}\left(\phi_P^{\mathrm{rw}}, \theta_P^{\mathrm{rw}}\right)\right]^H$, $\mathbf{g} = [g_1^{\mathrm{rw}}, \cdots, g_P^{\mathrm{rw}}]^T$, and (a) is based on the result from [37, (eq. 1.11.18)]. Then, by applying the large system analytical method in [38] on (43), we can further obtain (44) given below

$$\begin{aligned} &\lim_{M\to\infty} \frac{\left\|\boldsymbol{\Phi}\big(\mathbf{w}_k \otimes \boldsymbol{\Theta}_{\mathrm{r}}\big) \boldsymbol{\Omega}_{\mathrm{rw}}^H \mathbf{g}^*\right\|_2^2}{M} \\ &= \lim_{M\to\infty} \frac{\mathrm{Tr}\left(\mathbf{g}^T \boldsymbol{\Omega}_{\mathrm{rw}} \left(\mathbf{w}_k \otimes \boldsymbol{\Theta}_{\mathrm{r}}\right)^H \boldsymbol{\Phi}^H \boldsymbol{\Phi} (\mathbf{w}_k \otimes \boldsymbol{\Theta}_{\mathrm{r}}) \boldsymbol{\Omega}_{\mathrm{rw}}^H \mathbf{g}^*\right)}{M} \\ &\overset{(a)}{\to} \frac{\mathrm{Tr}\left(\boldsymbol{\Omega}_{\mathrm{rw}}^H \boldsymbol{\Omega}_{\mathrm{rw}} \big(\mathbf{w}_k \otimes \boldsymbol{\Theta}_{\mathrm{r}}\big)^H \boldsymbol{\Phi}^H \boldsymbol{\Phi} (\mathbf{w}_k \otimes \boldsymbol{\Theta}_{\mathrm{r}})\right)}{M} \\ &\overset{(b)}{=} \frac{\mathrm{Tr}\left(\sum_{l=1}^{L} \left(\mathbf{w}_k^H \boldsymbol{\Psi}_{\mathrm{BR}}^l \mathbf{w}_k\right) \otimes \left(\boldsymbol{\Omega}_{\mathrm{rw}}^H \boldsymbol{\Omega}_{\mathrm{rw}} \boldsymbol{\Theta}_{\mathrm{r}}^H \widehat{\boldsymbol{\Psi}}_{\mathrm{BR}}^{l} \boldsymbol{\Theta}_{\mathrm{r}}\right)\right)}{M} \\ &\overset{(c)}{=} \frac{\sum_{l=1}^{L} \left(\mathbf{w}_k^H \boldsymbol{\Psi}_{\mathrm{BR}}^l \mathbf{w}_k\right) \mathrm{Tr}\left(\boldsymbol{\Omega}_{\mathrm{rw}}^H \boldsymbol{\Omega}_{\mathrm{rw}} \boldsymbol{\Theta}_{\mathrm{r}}^H \widehat{\boldsymbol{\Psi}}_{\mathrm{BR}}^{l} \boldsymbol{\Theta}_{\mathrm{r}}\right)}{M} \\ &= \frac{\sum_{l=1}^{L} \left(\mathbf{w}_k^H \boldsymbol{\Psi}_{\mathrm{BR}}^l \mathbf{w}_k\right) \left(\boldsymbol{\vartheta}_{\mathrm{r}}^T \boldsymbol{\Xi}^T \left(\left(\widehat{\boldsymbol{\Psi}}_{\mathrm{BR}}^{l}\right)^T \otimes \left(\boldsymbol{\Omega}_{\mathrm{rw}}^H \boldsymbol{\Omega}_{\mathrm{rw}}\right)\right) \boldsymbol{\Xi} \boldsymbol{\vartheta}_{\mathrm{r}}^*\right)}{M}, \end{aligned} \tag{44}$$

where the convergence (a) is from the corollary in [38, Corollary 1]; steps (b) and (c) are because of [37, (eq. 1.10.15)] and [37, (eq. 1.10.11)], respectively. On the basis of the derived result in (44), the asymptotic results of λ_0 and λ_1 can be obtained as in (12) and (13).

References

[1] A. D. Wyner, "The wire-tap channel," *The Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, 1975.

[2] N. Yang, P. L. Yeoh, M. Elkashlan, R. Schober, and I. B. Collings, "Transmit antenna selection for security enhancement in MIMO wiretap channels," *IEEE Trans. Commun.*, vol. 61, no. 1, pp. 144–154, 2012.

[3] T. Lv, H. Gao, and S. Yang, "Secrecy transmit beamforming for heterogeneous networks," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 6, pp. 1154–1170, 2015.

[4] N. Zhao, Y. Cao, F. R. Yu, Y. Chen, M. Jin, and V. C. Leung, "Artificial noise assisted secure interference networks with wireless power transfer," *IEEE Trans. Veh. Technol.*, vol. 67, no. 2, pp. 1087–1098, 2017.

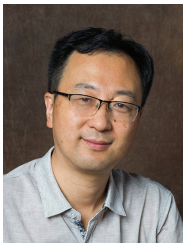
[5] X. Hu, P. Mu, B. Wang, and Z. Li, "On the secrecy rate maximization with uncoordinated cooperative jamming by single-antenna helpers," *IEEE Trans. Veh. Technol.*, vol. 66, no. 5, pp. 4457–4462, 2016.

[6] T.-X. Zheng, X. Chen, C. Wang, K.-K. Wong, and J. Yuan, "Physical layer security in large-scale random multiple access wireless sensor networks: a stochastic geometry approach," *IEEE Trans. Commun.*, vol. 70, no. 6, pp. 4038–4051, 2022.
[7] 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.
[8] 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.
[9] 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.
[10] 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.
[11] B. He, S. Yan, X. Zhou, and V. K. Lau, "On covert communication with noise uncertainty," *IEEE Commun. Lett.*, vol. 21, no. 4, pp. 941–944, 2017.
[12] 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.
[13] 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.
[14] 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.
[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.
[17] 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.
[18] L. Dong, H.-M. Wang, J. Bai, and H. Xiao, "Double intelligent reflecting surface for secure transmission with inter-surface signal reflection," *IEEE Trans. Veh. Technol.*, vol. 70, no. 3, pp. 2912–2916, 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. 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.
[22] 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.
[23] X. Mu, Y. Liu, L. Guo, J. Lin, and R. Schober, "Simultaneously tansmitting and reflecting (STAR) RIS aided wireless communications," *IEEE Trans. Wireless Commun.*, vol. 21, no. 5, pp. 3083–3098, 2022.
[24] 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.*, vol. 11, no. 6, pp. 1191–1195, 2022.
[25] 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.
[26] H. Xiao, X. Hu, P. Mu, W. Wang, T.-X. Zheng, K.-K. Wong, and K. Yang, "Simultaneously Transmitting and Reflecting RIS (STAR-RIS) Assisted Multi-Antenna Covert Communication: Analysis and Optimization," *IEEE Trans. Wireless Commun.*, pp. 1–1, early access, Doi: 10.1109/TWC.2023.3 331 706, 2023.
[27] H. Xiao, X. Hu, T.-X. Zheng, and K.-K. Wong, "Star-ris assisted covert communications in noma systems," *IEEE Trans. Veh. Technol.*, pp. 1–6, early access, Doi: 10.1109/TVT.2023.3 329 873, 2023.
[28] M. Forouzesh, P. Azmi, A. Kuhestani, and P. L. Yeoh, "Joint information-theoretic secrecy and covert communication in the presence of an untrusted user and warden," *IEEE Internet Things J.*, vol. 8, no. 9, pp. 7170–7181, 2020.
[29] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1164–1179, 2014.
[30] H. Guo and V. K. Lau, "Uplink cascaded channel estimation for intelligent reflecting surface assisted multiuser miso systems," *IEEE Trans. Signal Process.*, vol. 70, pp. 3964–3977, 2022.
[31] 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.
[32] 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.
[33] 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.
[34] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," 2014.
[35] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex optimization*. Cambridge Univ. Press, 2004.
[36] C. Feng, W. Shen, J. An, and L. Hanzo, "Joint hybrid and passive RIS-assisted beamforming for mmwave MIMO systems relying on dynamically configured subarrays," *IEEE Internet Things J.*, vol. 9, no. 15, pp. 13 913–13 926, 2022.
[37] X.-D. Zhang, *Matrix analysis and applications*. Cambridge Univ. Press, 2017.
[38] J. Evans and D. N. C. Tse, "Large system performance of linear multiuser receivers in multipath fading channels," *IEEE Transa. Inf. Theory*, vol. 46, no. 6, pp. 2059–2078, 2000.

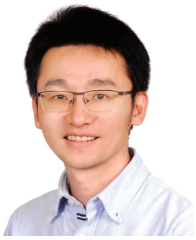
Han Xiao received the M.Eng. degree in Vehicle Engineering from Dalian University of Technology, Dalian, China, in 2021. He is currently pursuing a Ph.D. degree at the School of Information and Communications Engineering, Xi'an Jiaotong University, Xi'an, China. His research interests include physical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

Ang Li (Senior Member, IEEE) received his Ph.D. degree in the Communications and Information Systems research group, Department of Electrical and Electronic Engineering, University College London in April 2018. He was a postdoctoral research associate in the School of Electrical and Information Engineering, The University of Sydney from May 2018 to February 2020. He joined Xi'an Jiaotong University in March 2020 and is now a Professor in the School of Information and Communications Engineering, Faculty of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an, China. His main research interests lie in the physical-layer techniques in wireless communications, including MIMO/massive MIMO, interference exploitation, symbol-level precoding, and reconfigurable MIMO, etc. He currently serves as the Associate Editor for IEEE COMMUNICATIONS LETTERS, IEEE OPEN JOURNAL OF SIGNAL PROCESSING, and EURASIP JOURNAL ON WIRELESS COMMUNICATIONS AND NETWORKING. He is the recipient of the 2021 IEEE Signal Processing Society Young Author Best Paper Award. He has been an Exemplary Reviewer for IEEE COMMUNICATIONS LETTERS, IEEE TRANSACTIONS ON COMMUNICATIONS, and IEEE WIRELESS COMMUNICATIONS LETTERS. He has served as the Co-Chair of the IEEE ICASSP 2020 Special Session on 'Hardware-Efficient Large-Scale Antenna Arrays: The Stage for Symbol-Level Precoding', and has organized a Tutorial in IEEE ICC 2021 on 'Interference Exploitation through Symbol Level Precoding: Energy Efficient Transmission for 6G and Beyond'.



Wenjie Wang (Senior 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.



Zhou Su (Senior Member, IEEE) has published technical papers, including top journals and top conferences, such as IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING, IEEE TRANSACTIONS ON MOBILE COMPUTING, IEEE/ACM TRANSACTION ON NETWORKING, and INFOCOM. His research interests include multimedia communication, wireless communication, and network traffic. He received the Best Paper Award of International Conference IEEE ICC2020, IEEE BigdataSE2019, and IEEE CyberSciTech2017. He is also an Associate Editor of IEEE INTERNET OF THINGS JOURNAL, IEEE OPEN JOURNAL OF THE COMPUTER SOCIETY, and *IET Communications*.

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 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 Congress 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.