Robust Transmission Design for Intelligent Reflecting Surface Aided Secure Communication Systems with Imperfect Cascaded CSI

Sheng Hong, Cunhua Pan, Hong Ren, Kezhi Wang, Kok Keong Chai, *Member, IEEE*, and Arumugam Nallanathan, *Fellow, IEEE*

Abstract

In this paper, we investigate the design of robust and secure transmission in intelligent reflecting surface (IRS) aided wireless communication systems. In particular, a multi-antenna access point (AP) communicates with a single-antenna legitimate receiver in the presence of multiple single-antenna eavesdroppers, where the artificial noise (AN) is transmitted to enhance the security performance. Besides, we assume that the cascaded AP-IRS-user channels are imperfect due to the channel estimation error. To minimize the transmit power, the beamforming vector at the transmitter, the AN covariance matrix, and the IRS phase shifts are jointly optimized subject to the outage rate probability constraints under the statistical cascaded channel state information (CSI) error model. To handle the resulting non-convex optimization problem, we first approximate the outage rate probability constraints by using the Bernstein-type inequality. Then, we develop a suboptimal algorithm based on alternating optimization, the penalty-based and semidefinite relaxation methods. Simulation results reveal that the proposed scheme significantly reduces the transmit power compared to other benchmark schemes.

I. Introduction

The intelligent reflecting surface (IRS) has emerged due to the advancement in metamaterial techniques, and becomes a promising technology in wireless networks [1]. The IRS comprises a large number of passive elements, which can reflect the wireless signal with adjustable phase shifts [2] [3]. The IRS has the capability of reconfiguring the wireless propagation environment between

S. Hong is with Information Engineering School of Nanchang University, Nanchang 330031, China. She is also with the School of Electronic Engineering and Computer Science at Queen Mary University of London, London E1 4NS, U.K. (email: shenghong@ncu.edu.cn). C. Pan, H. Ren, K. K. Chai, and A. Nallanathan are with the School of Electronic Engineering and Computer Science at Queen Mary University of London, London E1 4NS, U.K. (e-mail:{c.pan, h.ren, michael.chai, a.nallanathan}@qmul.ac.uk). K. Wang is with Department of Computer and Information Sciences, Northumbria University, UK. (email: kezhi.wang@northumbria.ac.uk).

access point (AP) and users in a favourable way by properly designing the phase shifts. Thus, the IRS can improve the performance of wireless networks in various aspects. Since the IRS operates in a passive mode by reflecting incident signals [4], it can significantly improve the spectral and energy efficiency of the wireless networks. Besides, it is very appealing that the IRS is low-cost, and can be deployed easily on buildings facades, interior walls, room ceilings, lamp posts and road signs, etc. Therefore, exploiting the IRS device to assist wireless communication systems has received extensive attention. The IRS-aided wireless communication systems in the existing literature include the single-user case [5] [6], the downlink multiuser case [7] [8] [9], multicell scenario [10], wireless power transfer design [11], mobile edge computing [12], and cognitive radio system [13].

Recently, the IRS has emerged as a promising technology to enhance the physical layer security in a wireless communication system. There are various schemes to improve the physical layer security [14], such as the multiple-antenna techniques [15], cooperative relaying [16], artificial noise (AN) [17], and cooperative jamming [18]. By using these schemes, the AP-eavesdroppers links are impaired, and the information leakage to the eavesdroppers is limited. Combining with these schemes, the IRS can further enhance the physical layer security. On one hand, deploying the IRS costs much less than deploying the relay since no active radio frequency chain is required in IRS-aided systems. On the other hand, the IRS can be programmed to configure the radio propagation environment to make the impairment on the eavesdroppers' channels more effective.

In terms of physical layer security enhancement, the IRS-aided secure communication has received considerable attention from academia [19] [20] [21] [22] [23] [24]. In these contributions, the active transmit beamforming and the passive reflecting beamforming of the IRS were jointly designed to improve the achievable secrecy rate. The transmit beamforming can be designed with the CVX solver or a closed-form solution, while the reflecting beamforming is much more difficult to handle due to the unit modulus constraints of the IRS phase shifts. The alternating optimization with minorization maximization (AO-MM) algorithm was proposed in [19]. The semidefinite relax (SDR) technique was utilized to optimize the phase shifts of the IRS in [20] and [21]. By using the bisection search method, a closed form solution for the phase shifts of the IRS was obtained in [22]. In [23], the unit modulus constraints of the phase shifters were approximated by a set of convex constraints. In [24], a novel deep reinforcement learning (DRL)-based secure beamforming approach was proposed to achieve the optimal beamforming. Moreover, the AN is incorporated

with transmit beamforming in IRS-aided wireless secure communications [25] [26] [27]. The phase shift matrix at the IRS as well as the beamforming vectors and AN covariance matrix at the base station (BS) were jointly optimized. It was unveiled that it is beneficial for secrecy enhancement with the aid of AN.

We note that all these existing contributions are based on the ideal assumption of perfect channel state information (CSI) of both the legitimate receiver and the eavesdropper at the transmitters. However, it is difficult to obtain the perfect CSI in IRS-aided wireless communication, because the IRS does not employ any radio frequency (RF) chains. Thus, how to estimate the IRS-related channel of the reflective AP-IRS-user link has attracted a lot of research attention. The contributions of the current channel estimation for the IRS-related channels can be divided into two approaches. The first one is to estimate AP-IRS channel and IRS-user channel separately [28]. As shown in [28], to estimate the two separated IRS-related channels, active transmit radio RF chains are required to be installed at the IRS. The drawback of this approach is that extra hardware and power consumption is required. The second one is to estimate the cascaded AP-IRS-user channel, which is defined as the product of the AP-IRS channel and the IRS-user channel [9] [19] [22] [29]. The benefit of this approach is that the cascaded AP-IRS-user channel can be estimated without installing extra hardware and incurring power cost, and that the estimated cascaded AP-IRS-user channel is sufficient for the transmission beamforming design for the IRS-related links. Thus, more efforts are dedicated to the cascaded channel estimation [30] [31] [32] [33]. The cascaded channel estimation methods were investigated in a single-user multiple-input multiple-output (SU-MIMO) system [30] and a multi-user multiple-input single-output (MU-MISO) system [31], respectively. Then, by exploiting the channel sparsity of the mmWave channels, the compressive sensing methods were adopted to reduce the pilot overhead in [32] and [33] for single-user and multi-user systems, respectively. However, for both approaches, the channel estimation error is inevitable. Therefore, it is imperative to take the channel estimation error into consideration when designing the IRS-aided communication systems.

There are a paucity of contributions investigating the robust transmission design in IRS-aided communications. The work of [34] proposed a worst-case robust design algorithm in an MU-MISO wireless system, where the IRS-user channels were assumed to be imperfect. In addition, a worst-case robust design in IRS-assisted secure communications was investigated in [35], where the IRS-eavesdropper channels were assumed to be imperfect. Both these works only considered

the IRS-user CSI error based on the first IRS-related channel estimation approach. Since it is more practical to consider the cascaded channel uncertainty based on the second IRS-related channel estimation approach, Zhou *et. al* [36] firstly proposed a robust transmission framework with both the bounded and the statistical cascaded CSI errors in an MU-MISO wireless system.

However, to the best of our knowledge, the robust transmission design with cascaded channel error in secure IRS-aided communication systems has not been studied yet. Moreover, the probabilistic CSI error model has not been studied in secure IRS-aided communication systems. To fill this gap, this paper investigates the outage constrained robust secure transmission for IRS-aided secure wireless communication systems, where the statistical model of the cascaded channel error is considered. Specifically, we consider a scenario that a multi-antenna AP serves a single-antenna legitimate user in the presence of several single-antenna eavesdroppers. Moreover, the AN is assumed to be injected to degrade the reception quality of eavesdroppers. In this scenario, we consider an outage-constrained robust design aiming for minimizing the transmit power by considering the imperfect CSI of the eavesdroppers' channels. An outage-constrained power minimization (OC-PM) problem is formulated to jointly optimize the beamforming vector, the AN covariance matrix, and the phase shifts matrix of IRS with the outage constraints of maximum information leakage to eavesdroppers and the constraint of minimum information transmission to legitimate users. Since the optimization variables are highly coupled, an alternating optimization (AO) method is proposed to solve it.

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

1) To the best of our knowledge, this is the first work to study the outage constrained robust secure transmission for IRS-aided wireless communications. In contrast to [34] and [35] that only considered the bounded CSI error model in secure IRS-aided wireless communications, we first consider the statistical CSI error model.
2) To the best of our knowledge, this is the first work to consider the imperfect cascaded channels of AP-IRS-eavesdropper in secure IRS-aided communications, which is more practical compared to the existing literature considering the imperfect IRS-eavesdropper channels.
3) For the outage constrained robust secure transmission design, we formulate an OC-PM problem to optimize the beamforming vector, the AN covariance matrix, and the phase shifts matrix of IRS. To solve it, we first transform the probabilistic constraints into the safe and tractable forms by exploiting Bernstein-type inequality (BTI) [37]. Then, the AO

strategy is utilized to transform the original problems into two subproblems, where the designs of beamforming vector, AN covariance matrix, and phase shifts of IRS are handled by semidefinite relaxation (SDR) based methods. For the nonconvex unit modulus constraints of IRS phase shifts, an equivalent rank-one constraint is incorporated, which is further transformed by the first-order Taylor approximation and added into the objective function as a penalty.

4) Simulation results demonstrate that the robust design of beamforming vector and AN covariance matrix in a secure IRS-aided communication system can reduce the transmit power under the fixed secrecy rate. In comparison to various benchmark methods, the effectiveness of the proposed AO algorithm is verified. Moreover, it is revealed that by deploying the IRS and optimizing the IRS phase shifts, the reliable communication can be guaranteed for the legitimate user, while the information leakage to eavesdropper can be limited.

Notations: Throughout this paper, boldface lower case, boldface upper case and regular letters are used to denote vectors, matrices, and scalars, respectively. $\mathrm{Tr}\left(\mathbf{X}\right)$ and $|\mathbf{X}|$ denote the trace and determinant of $\mathbf{X}$, respectively. $\mathbb{C}^{M\times N}$ denotes the space of $M \times N$ complex matrices. $\mathrm{Re}\{\cdot\}$ denotes the real part of a complex value. $\mathrm{diag}(\cdot)$ is the operator for diagonalization. $\mathbf{1}_M$ presents a column vector with M ones. $\mathcal{CN}(\boldsymbol{\mu}, \mathbf{Z})$ represents a circularly symmetric complex gaussian (CSCG) random vector with mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{Z}$. $(\cdot)^{\mathrm{T}}$, $(\cdot)^{\mathrm{H}}$ and $(\cdot)^*$ denote the transpose, Hermitian and conjugate operators, respectively.

II. System model

In this section, we present the transmission model and CSI error model in a secure IRS-aided communication system as follows.

A. Signal Transmission Model

An IRS-assisted communication system is considered, where an AP equipped with N_t antennas, called Alice, intends to send confidential information to a single-antenna legitimate receiver, called Bob, in the presence of K single-antenna eavesdroppers, called Eves.

As shown in Fig. 1, the signal transmitted from Alice is reflected by the IRS, which comprises M reflecting elements. The direct links of Alice-Bob and Alice-Eve are blocked by obstacles such as buildings. The Bob and Eves can only receive the signals reflected by the IRS. The received

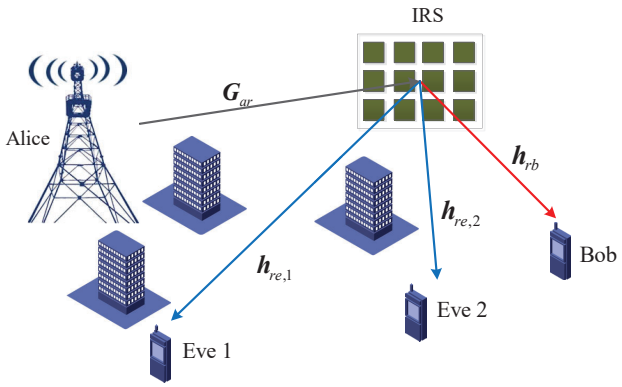


Fig. 1. An IRS-assisted secure communication system.

signals at Bob and Eves are respectively modeled as

$$y_b(t) = \hat{\mathbf{h}}_b^H \mathbf{x}(t) + n_b(t) = (\mathbf{h}_{rb}^H \mathbf{\Phi} \mathbf{G}_{ar})\mathbf{x}(t) + n_b(t), \tag{1}$$

$$\mathrm{y}_{e,k}(t) = \hat{\mathbf{h}}_{e,k}^H \mathbf{x}(t) + \mathrm{n}_{e,k}(t) = (\mathbf{h}_{re,k}^H \mathbf{\Phi} \mathbf{G}_{ar})\mathbf{x}(t) + n_{e,k}(t), k = 1, 2, \cdots, K, \tag{2}$$

where $\hat{\mathbf{h}}_b = \mathbf{G}_{ar}^H \mathbf{\Phi}^H \mathbf{h}_{rb}$, $\hat{\mathbf{h}}_b \in \mathbb{C}^{N_t \times 1}$ is defined as the equivalent channel spanning from Alice to Bob. The channels of the IRS-Bob link and the Alice-IRS link are modeled by $\mathbf{h}_{rb} \in \mathbb{C}^{M \times 1}$ and $\mathbf{G}_{ar} \in \mathbb{C}^{M \times N_t}$, respectively. $\hat{\mathbf{h}}_{e,k} = \mathbf{G}_{ar}^H \mathbf{\Phi}^H \mathbf{h}_{re,k}$, $\hat{\mathbf{h}}_{e,k} \in \mathbb{C}^{N_t \times 1}$ is defined as the equivalent channel spanning from Alice to the kth Eve. The channel of the IRS-Eve link is modeled by $\mathbf{h}_{re,k} \in \mathbb{C}^{M \times 1}$. $n_b(t)$ is the additive white Gaussian noise (AWGN) at Bob with zero mean and noise variance σ_b^2, i.e., $n_b \sim \mathcal{CN}(0, \sigma_b^2)$. $n_{e,k}(t)$ is the AWGN at the kth Eve with zero mean and noise covariance matrix $\sigma_{e,k}^2$, i.e., $n_{e,k} \sim \mathcal{CN}(0, \sigma_{e,k}^2)$. The phase shift coefficients of the IRS are collected in a diagonal matrix defined by $\mathbf{\Phi} = \mathrm{diag}\{\phi_1, \cdots, \phi_m, \cdots, \phi_M\}$, where $\phi_m = e^{j\theta_m}$, and $\theta_m \in [0, 2\pi]$ denotes the phase shift of the m-th reflecting element.

The transmit signal vector is

$$\mathbf{x}(t) = \mathbf{s}(t) + \mathbf{z}(t) = \mathbf{w}s(t) + \mathbf{z}(t), \tag{3}$$

where $s(t)$ is a data stream carrying the confidential information intended for Bob, $\mathbf{z}(t)$ is the noise vector artificially created by Alice to interfere Eves, i.e., the so-called AN, and $\mathbf{w}$ is the beamforming vector. The confidential signal vector $\mathbf{s}(t)$ is assumed to follow a complex Gaussian distribution of $\mathcal{CN}(\mathbf{0}, \mathbf{W})$, where $\mathbf{W} = \mathbf{w}\mathbf{w}^H$ is the transmit covariance matrix to be designed. For the AN, we assume $\mathbf{z}(t) \sim \mathcal{CN}(\mathbf{0}, \mathbf{Z})$, where $\mathbf{Z}$ is the AN covariance matrix to be designed.

The equivalent channel $\hat{\mathbf{h}}_b$ can be rewritten as

$$\hat{\mathbf{h}}_b^H = \boldsymbol{\phi}^T \text{diag}(\mathbf{h}_{rb}^H)\mathbf{G}_{ar} \triangleq \boldsymbol{\phi}^T \mathbf{G}_{cb}, \tag{4}$$

where $\boldsymbol{\phi} = [\phi_1, \cdots, \phi_m, \cdots, \phi_M]^T$ and $\mathbf{G}_{cb} = \text{diag}(\mathbf{h}_{rb}^H)\mathbf{G}_{ar}$. $\mathbf{G}_{cb} \in \mathbb{C}^{M \times N_t}$ is defined as the cascaded channel from Alice to Bob via the IRS. Similarly, the equivalent channel $\hat{\mathbf{h}}_{e,k}$ can be expressed in another way as

$$\hat{\mathbf{h}}_{e,k}^H = \boldsymbol{\phi}^T \text{diag}(\mathbf{h}_{re,k}^H)\mathbf{G}_{ar} \triangleq \boldsymbol{\phi}^T \mathbf{G}_{ce,k}, \tag{5}$$

where $\mathbf{G}_{ce,k} = \text{diag}(\mathbf{h}_{re,k}^H)\mathbf{G}_{ar}$, $k = 1, 2, \cdots, K$. $\mathbf{G}_{ce,k} \in \mathbb{C}^{M \times N_t}$ denotes the cascaded channel from Alice to the kth Eve via the IRS.

B. Statistical CSI Error Model

In IRS-aided communications, the cascaded AP-IRS-user channels at the transmitter is challenging to obtain due to the passive features of the IRS. Hence, we consider the CSI error in the cascaded AP-IRS-user channels. Generally, there are two channel error models: the bounded CSI error model and the statistical CSI error model. In this paper, we focus on the statistical CSI error of the cascaded channel for Eves since the bounded CSI error is more conservative. Furthermore, the statistical CSI error model is more closely related to the channel estimation error, while the bounded CSI error usually models the quantization error.

We assume the cascaded channel $\mathbf{G}_{ce,k}$ from Alice to the kth Eve is imperfect, which can be represented as

$$\mathbf{G}_{ce,k} = \bar{\mathbf{G}}_{ce,k} + \triangle\mathbf{G}_{ce,k}, \mathbf{g}_{ce,k} \triangleq \text{vec}(\triangle\mathbf{G}_{ce,k}) \sim \mathcal{CN}(\mathbf{0}, \boldsymbol{\Sigma}_{e,k}), \tag{6}$$

where the CSI error vector $\mathbf{g}_{ce,k} \triangleq \text{vec}(\triangle\mathbf{G}_{ce,k})$ is assumed to follow the CSCG distribution, and $\boldsymbol{\Sigma}_{e,k} \in \mathbb{C}^{MN_t \times MN_t}$ is the semidefinite error covariance matrix.

III. Problem Formulation

In this section, we first derive the secrecy rate expression and then formulate the outage-constrained power minimization problem.

A. Secrecy Rate

The achievable data rate (bit/s/Hz) of Bob is given by

$$C_b(\mathbf{W},\mathbf{Z},\mathbf{\Phi}) = \log\left(1+\frac{\hat{\mathbf{h}}_b^H\mathbf{W}\hat{\mathbf{h}}_b}{\sigma_b^2+\hat{\mathbf{h}}_b^H\mathbf{Z}\hat{\mathbf{h}}_b}\right) = \log\left(1+\frac{(\boldsymbol{\phi}^T\mathbf{G}_{cb})\mathbf{W}(\boldsymbol{\phi}^T\mathbf{G}_{cb})^H}{\sigma_b^2+(\boldsymbol{\phi}^T\mathbf{G}_{cb})\mathbf{Z}(\boldsymbol{\phi}^T\mathbf{G}_{cb})^H}\right), \tag{7}$$

where the $\log(\cdot)$ denotes the base 2 logarithm, and the cascaded channel $\mathbf{G}_{cb}$ from Alice to Bob is utilized. The achievable data rate (bit/s/Hz) of the kth Eve is given by

$$C_{e,k}(\mathbf{W},\mathbf{Z},\mathbf{\Phi}) = \log\left(1+\frac{\hat{\mathbf{h}}_{e,k}^H\mathbf{W}\hat{\mathbf{h}}_{e,k}}{\sigma_{e,k}^2+\hat{\mathbf{h}}_{e,k}^H\mathbf{Z}\hat{\mathbf{h}}_{e,k}}\right) = \log\left(1+\frac{(\boldsymbol{\phi}^T\mathbf{G}_{ce,k})\mathbf{W}(\boldsymbol{\phi}^T\mathbf{G}_{ce,k})^H}{\sigma_{e,k}^2+(\boldsymbol{\phi}^T\mathbf{G}_{ce,k})\mathbf{Z}(\boldsymbol{\phi}^T\mathbf{G}_{ce,k})^H}\right), \tag{8}$$

where the cascaded channel $\mathbf{G}_{ce,k}$ from Alice to the kth Eve is applied.

Then the achievable secrecy rate is

$$R_s(\mathbf{W},\mathbf{Z},\mathbf{\Phi}) = \min_{k=1,2,\cdots,K}\left\{C_b(\mathbf{W},\mathbf{Z},\mathbf{\Phi}) - C_{e,k}(\mathbf{W},\mathbf{Z},\mathbf{\Phi})\right\}. \tag{9}$$

B. Power Minimization

In this paper, we aim to jointly optimize the transmit beamforming vector $\mathbf{w}$ (or the transmit beamforming matrix $\mathbf{W}=\mathbf{w}\mathbf{w}^H$), the covariance matrix $\mathbf{Z}$ of AN and the IRS phase shifts $\mathbf{\Phi}$ to minimize the transmit power subject to the Bob's data rate requirement and leaked data rate outage for Eves. The OC-PM problem can be formulated as

$$\min_{\mathbf{W},\mathbf{Z},\mathbf{\Phi}} \mathrm{Tr}(\mathbf{W}+\mathbf{Z}) \tag{10a}$$

$$\text{s.t.} \quad C_b(\mathbf{W},\mathbf{Z},\mathbf{\Phi}) \geq \log\gamma, \tag{10b}$$

$$\Pr_{\mathbf{g}_{ce,k}}\left\{C_{e,k}(\mathbf{W},\mathbf{Z},\mathbf{\Phi}) \leq \log\beta\right\} \geq 1-\rho_k, k=1,2,\cdots,K, \tag{10c}$$

$$\mathbf{Z} \succeq 0, \tag{10d}$$

$$\mathbf{W} \succeq 0, \tag{10e}$$

$$\mathrm{rank}(\mathbf{W}) = 1, \tag{10f}$$

$$|\phi_m| = 1, m=1,\cdots,M, \tag{10g}$$

where γ, β, and $\rho_k, \forall k$ are constant values, and (10c) can be seen as the constraints of per-Eve secrecy outage probability.

IV. Problem Solution

Due to the coupling relation of various optimization variables and the complicated probabilistic constraints, the proposed OC-PM problem is nonconvex and challenging to solve. To tackle it, we first transform the probabilistic constraints into the safe and tractable forms. Then, the AO strategy is utilized to decouple the parameters and transform the original problems into two subproblems.

A. Problem Reformulation

We first reformulate the constraint in (10b). Specifically, it can be simplified by removing the log function as

$$\log\left(1+\frac{\hat{\mathbf{h}}_b^H\mathbf{W}\hat{\mathbf{h}}_b}{\sigma_b^2+\hat{\mathbf{h}}_b^H\mathbf{Z}\hat{\mathbf{h}}_b}\right)\geq\log\gamma \tag{11a}$$

$$\Leftrightarrow\mathrm{Tr}([\mathbf{W}-(\gamma-1)\mathbf{Z}]\hat{\mathbf{h}}_b\hat{\mathbf{h}}_b^H)\geq(\gamma-1)\sigma_b^2, \tag{11b}$$

$$\Leftrightarrow\mathrm{Tr}([\mathbf{W}-(\gamma-1)\mathbf{Z}](\mathbf{G}_{cb}^H\boldsymbol{\phi}^*)(\boldsymbol{\phi}^T\mathbf{G}_{cb}))\geq(\gamma-1)\sigma_b^2, \tag{11c}$$

$$\Leftrightarrow\mathrm{Tr}(\mathbf{G}_{cb}[\mathbf{W}-(\gamma-1)\mathbf{Z}]\mathbf{G}_{cb}^H\mathbf{E})\geq(\gamma-1)\sigma_b^2, \tag{11d}$$

where $\mathbf{E}\triangleq\boldsymbol{\phi}^*\boldsymbol{\phi}^T$. Then, we reformulate the per-Eve outage probability constraint (10c), which is not tractable to handle due to the log function and the probability requirement. We first eliminate the log function, then transform the probability constraint into the deterministic constraint. There are two steps to reformulate the chance constraint (10c) as follows.

Step 1): Firstly, the data rate leakage expression in (10c) can be simplified as

$$\log\left(1+\frac{\hat{\mathbf{h}}_{e,k}^H\mathbf{W}\hat{\mathbf{h}}_{e,k}}{\sigma_{e,k}^2+\hat{\mathbf{h}}_{e,k}^H\mathbf{Z}\hat{\mathbf{h}}_{e,k}}\right)\leq\log\beta \tag{12a}$$

$$\Leftrightarrow\hat{\mathbf{h}}_{e,k}^H[\mathbf{W}-(\beta-1)\mathbf{Z}]\hat{\mathbf{h}}_{e,k}\leq(\beta-1)\sigma_{e,k}^2. \tag{12b}$$

By substituting $\hat{\mathbf{h}}_{e,k}^H=\boldsymbol{\phi}^T(\bar{\mathbf{G}}_{ce,k}+\triangle\mathbf{G}_{ce,k})$ into (12b) and defining $\boldsymbol{\Xi}_e\triangleq(\beta-1)\mathbf{Z}-\mathbf{W}$, we have

$$(12a)\Leftrightarrow[\boldsymbol{\phi}^T(\bar{\mathbf{G}}_{ce,k}+\triangle\mathbf{G}_{ce,k})]\boldsymbol{\Xi}_e[\boldsymbol{\phi}^T(\bar{\mathbf{G}}_{ce,k}+\triangle\mathbf{G}_{ce,k})]^H+(\beta-1)\sigma_{e,k}^2\geq0, \tag{13a}$$

$$\Leftrightarrow\underbrace{\boldsymbol{\phi}^T\triangle\mathbf{G}_{ce,k}\boldsymbol{\Xi}_e\triangle\mathbf{G}_{ce,k}^H\boldsymbol{\phi}^*}_{f_1}+2\mathrm{Re}\{\underbrace{\boldsymbol{\phi}^T\triangle\mathbf{G}_{ce,k}\boldsymbol{\Xi}_e\bar{\mathbf{G}}_{ce,k}^H\boldsymbol{\phi}^*}_{f_2}\}$$
$$+\underbrace{(\boldsymbol{\phi}^T\bar{\mathbf{G}}_{ce,k})\boldsymbol{\Xi}_e(\boldsymbol{\phi}^T\bar{\mathbf{G}}_{ce,k})^H+(\beta-1)\sigma_{e,k}^2}_{c_k}\geq0. \tag{13b}$$

Let us equivalently represent the CSI error $\mathbf{g}_{ce,k} = \text{vec}(\triangle\mathbf{G}_{ce,k})$ as $\mathbf{g}_{ce,k} = \mathbf{\Sigma}_{e,k}^{1/2}\mathbf{v}_{ce,k}$, where $\mathbf{v}_{ce,k} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{MN_t})$, $\mathbf{\Sigma}_{e,k} = \mathbf{\Sigma}_{e,k}^{1/2}\mathbf{\Sigma}_{e,k}^{1/2}$ and $(\mathbf{\Sigma}_{e,k}^{1/2})^H = \mathbf{\Sigma}_{e,k}^{1/2}$. The expression f_1 in (13b) can be reformulated as

$$\begin{aligned} f_1 &= \text{Tr}(\triangle\mathbf{G}_{ce,k}\mathbf{\Xi}_e\triangle\mathbf{G}_{ce,k}^H\boldsymbol{\phi}^*\boldsymbol{\phi}^T) \\ &= \text{Tr}(\triangle\mathbf{G}_{ce,k}^H\mathbf{E}\triangle\mathbf{G}_{ce,k}\mathbf{\Xi}_e) \\ &\overset{(a)}{=} \text{vec}^H(\triangle\mathbf{G}_{ce,k})(\mathbf{\Xi}_e^T \otimes \mathbf{E})\text{vec}(\triangle\mathbf{G}_{ce,k}) \\ &= \mathbf{g}_{ce,k}^H(\mathbf{\Xi}_e^T \otimes \mathbf{E})\mathbf{g}_{ce,k} \\ &= \mathbf{v}_{ce,k}^H\mathbf{\Sigma}_{e,k}^{1/2}(\mathbf{\Xi}_e^T \otimes \mathbf{E})\mathbf{\Sigma}_{e,k}^{1/2}\mathbf{v}_{ce,k} \\ &\triangleq \mathbf{v}_{ce,k}^H\mathbf{A}_k\mathbf{v}_{ce,k}, \end{aligned} \tag{14}$$

where $\mathbf{A}_k \triangleq \mathbf{\Sigma}_{e,k}^{1/2}(\mathbf{\Xi}_e^T \otimes \mathbf{E})\mathbf{\Sigma}_{e,k}^{1/2}$, and (a) is obtained by invoking the identity $\text{Tr}(\mathbf{A}^H\mathbf{BCD}) = \text{vec}^H(\mathbf{A})(\mathbf{D}^T \otimes \mathbf{B})\text{vec}(\mathbf{C})$. The expression f_2 in (13b) can be reformulated as

$$\begin{aligned} f_2 &= \text{Tr}(\triangle\mathbf{G}_{ce,k}\mathbf{\Xi}_e(\bar{\mathbf{G}}_{ce,k}^H\boldsymbol{\phi}^*)\boldsymbol{\phi}^T) \\ &= \text{Tr}(\triangle\mathbf{G}_{ce,k}\mathbf{\Xi}_e(\bar{\mathbf{G}}_{ce,k}^H\mathbf{E})) \\ &\overset{(b)}{=} \text{vec}^H(\mathbf{E}\bar{\mathbf{G}}_{ce,k})(\mathbf{\Xi}_e^T \otimes \mathbf{I}_M)\text{vec}(\triangle\mathbf{G}_{ce,k}) \\ &= \text{vec}^H(\mathbf{E}\bar{\mathbf{G}}_{ce,k})(\mathbf{\Xi}_e^T \otimes \mathbf{I}_M)\mathbf{g}_{ce,k} \\ &= \text{vec}^H(\mathbf{E}\bar{\mathbf{G}}_{ce,k})(\mathbf{\Xi}_e^T \otimes \mathbf{I}_M)\mathbf{\Sigma}_{e,k}^{1/2}\mathbf{v}_{ce,k} \\ &\triangleq \mathbf{u}_{ce,k}^H\mathbf{v}_{ce,k}, \end{aligned} \tag{15}$$

where $\mathbf{u}_{ce,k}^H \triangleq \text{vec}^H(\mathbf{E}\bar{\mathbf{G}}_{ce,k})(\mathbf{\Xi}_e^T \otimes \mathbf{I}_M)\mathbf{\Sigma}_{e,k}^{1/2}$, and (b) is obtained by invoking the identity $\text{Tr}(\mathbf{ABC}^H) = \text{vec}^H(\mathbf{C})(\mathbf{B}^T \otimes \mathbf{I})\text{vec}(\mathbf{A})$. The expression c_k in (13b) can be reformulated as

$$c_k = \text{Tr}[\bar{\mathbf{G}}_{ce,k}\mathbf{\Xi}_e\bar{\mathbf{G}}_{ce,k}^H\mathbf{E}] + (\beta - 1)\sigma_{e,k}^2. \tag{16}$$

By substituting (14), (15) and (16) into (13b), we have

$$\text{(12a)} \Leftrightarrow \mathbf{v}_{ce,k}^H\mathbf{A}_k\mathbf{v}_{ce,k} + 2\text{Re}\{\mathbf{u}_{ce,k}^H\mathbf{v}_{ce,k}\} + c_k \geq 0. \tag{17}$$

The leakage data rate constraint in (10c) becomes

$$\text{(10c)} \Leftrightarrow \text{Pr}_{\mathbf{v}_{ce,k}}\left\{\mathbf{v}_{ce,k}^H\mathbf{A}_k\mathbf{v}_{ce,k} + 2\text{Re}\{\mathbf{u}_{ce,k}^H\mathbf{v}_{ce,k}\} + c_k \geq 0\right\} \geq 1 - \rho_k, k = 1, 2, \cdots, K. \tag{18}$$

The equivalence in (18) implies that the outage probability in (10c) can be characterized by the quadratic inequality with respect to (w.r.t.) the Gaussian random vector $\mathbf{v}_{ce,k}$.

Step 2): the chance constraint is transformed into a deterministic constraint by using the BTI, which provides a safe approximation for (18).

The OC-PM problem in (10) is computationally intractable due to the fact that the rate outage probability constraints (10c) have no simple closed-form expressions. To solve Problem (10), the BTI is utilized to approximate the chance constraints (18) safely, which is given in the following Lemma.

Lemma 1: (Bernstein-Type Inequality[1] [37]) For any $(\mathbf{A}, \mathbf{u}, c) \in \mathbb{H}^n \times \mathbb{C}^n \times \mathbb{R}$,$\mathbf{v} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_n)$ and $\rho \in$ (0,1], the following implication holds:

$$\begin{aligned}
&\Pr_{\mathbf{v}}\left\{\mathbf{v}^H\mathbf{A}\mathbf{v} + 2\mathrm{Re}\{\mathbf{u}^H\mathbf{v}\} + c \geq 0\right\} \geq 1-\rho \\
\Longleftrightarrow &\mathrm{Tr}(\mathbf{A}) - \sqrt{2\ln(1/\rho)}\sqrt{\|\mathbf{A}\|_F^2 + 2\|\mathbf{u}\|^2} + \ln(\rho)\cdot\lambda^+(-\mathbf{A}) + c \geq 0, \\
\Longleftrightarrow &\begin{cases}
\mathrm{Tr}(\mathbf{A}) - \sqrt{-2\ln(\rho)}\cdot x + \ln(\rho)\cdot y + c \geq 0, \\
\left\| \begin{bmatrix} \mathrm{vec}(\mathbf{A}), \\ \sqrt{2}\mathbf{u} \end{bmatrix} \right\|_2 \leq x, \\
y\mathbf{I}_n + \mathbf{A} \succeq \mathbf{0}, y \geq 0,
\end{cases}
\end{aligned}$$

where $\lambda^+(\mathbf{A}) = \max(\lambda_{\max}(\mathbf{A}), 0)$. x and y are slack variables.

Please refer to [37] for the detailed proof of Lemma 1.

The strategy behind the BTI for tackling the rate outage constrained problem (10) is to pursue a convex restriction approach, also known as safe tractable approximation in the chance constrained optimization literature [37]. Based on the equivalence in (18), the more tractable upper bounds on the secrecy rate outage probabilities in (10c) can be developed.

By invoking Lemma 1 and introducing the slack variables $\mathbf{x} = [x_1, x_2, \cdots, x_K]^T$ and $\mathbf{y} =$

[1]This method has been widely utilized in the outage constrained robust design for secure communication systems [38] [39] [40].

$[y_1, y_2, \cdots, y_K]^T$, we arrive at the desired safe approximation of OC-PM, which is shown as

$$\min_{\mathbf{W},\mathbf{Z},\boldsymbol{\phi},\mathbf{A},\mathbf{x},\mathbf{y}} \ \mathrm{Tr}(\mathbf{W}+\mathbf{Z}) \tag{19a}$$

$$\text{s.t.} \quad \mathrm{Tr}(\mathbf{A}_k) - \sqrt{-2\ln(\rho_k)} \cdot x_k + \ln(\rho_k) \cdot y_k + c_k \geq 0, \tag{19b}$$

$$\left\| \begin{bmatrix} \mathrm{vec}(\mathbf{A}_k) \\ \sqrt{2}\mathbf{u}_{ce,k} \end{bmatrix} \right\|_2 \leq x_k, \tag{19c}$$

$$y_k \mathbf{I}_{MN_t} + \mathbf{A}_k \succeq \mathbf{0}, y_k \geq 0, \tag{19d}$$

$$\mathrm{Tr}(\mathbf{G}_{cb}[\mathbf{W} - (\gamma-1)\mathbf{Z}]\mathbf{G}_{cb}^H \mathbf{E}) \geq (\gamma-1)\sigma_b^2, \tag{19e}$$

$$(10\text{d}), (10\text{e}), (10\text{f}), (10\text{g}), \tag{19f}$$

where $\mathbf{A} = [\mathbf{A}_1, \mathbf{A}_2, \cdots, \mathbf{A}_K]$. The constraint (19e) is obtained by substituting (11d) into (10b). The constraints (19b), (19c) and (19d) can be further simplified by some mathematical transformations as follows.

The $\mathrm{Tr}(\mathbf{A}_k)$ in (19b) can be rewritten as

$$\mathrm{Tr}(\mathbf{A}_k) = \mathrm{Tr}(\boldsymbol{\Sigma}_{e,k}^{1/2}(\boldsymbol{\Xi}_e^T \otimes \mathbf{E})\boldsymbol{\Sigma}_{e,k}^{1/2}) = \mathrm{Tr}((\boldsymbol{\Xi}_e^T \otimes \mathbf{E})\boldsymbol{\Sigma}_{e,k}). \tag{20}$$

The $\|\mathrm{vec}(\mathbf{A}_k)\|_2^2$ in (19c) can be written as

$$\begin{aligned}
\|\mathrm{vec}(\mathbf{A}_k)\|_2^2 = \|\mathbf{A}_k\|_F^2 &= \mathrm{Tr}[\mathbf{A}_k\mathbf{A}_k^H] = \mathrm{Tr}[\boldsymbol{\Sigma}_{e,k}^{1/2}(\boldsymbol{\Xi}_e^T \otimes \mathbf{E})\boldsymbol{\Sigma}_{e,k}^{1/2}\boldsymbol{\Sigma}_{e,k}^{1/2}(\boldsymbol{\Xi}_e^* \otimes \mathbf{E}^H)\boldsymbol{\Sigma}_{e,k}^{1/2}] \\
&= \mathrm{Tr}[(\boldsymbol{\Xi}_e^T \otimes \mathbf{E})^H \boldsymbol{\Sigma}_{e,k}(\boldsymbol{\Xi}_e^T \otimes \mathbf{E})\boldsymbol{\Sigma}_{e,k}] \\
&\overset{(c)}{=} \mathrm{vec}^H(\boldsymbol{\Xi}_e^T \otimes \mathbf{E})(\boldsymbol{\Sigma}_{e,k}^T \otimes \boldsymbol{\Sigma}_{e,k})\mathrm{vec}(\boldsymbol{\Xi}_e^T \otimes \mathbf{E}) \\
&= \mathrm{vec}^H(\boldsymbol{\Xi}_e^T \otimes \mathbf{E})[(\boldsymbol{\Sigma}_{e,k}^{1/2T}\boldsymbol{\Sigma}_{e,k}^{1/2T}) \otimes (\boldsymbol{\Sigma}_{e,k}^{1/2}\boldsymbol{\Sigma}_{e,k}^{1/2})]\mathrm{vec}(\boldsymbol{\Xi}_e^T \otimes \mathbf{E}) \\
&= \mathrm{vec}^H(\boldsymbol{\Xi}_e^T \otimes \mathbf{E})[(\boldsymbol{\Sigma}_{e,k}^{1/2T} \otimes \boldsymbol{\Sigma}_{e,k}^{1/2})^H(\boldsymbol{\Sigma}_{e,k}^{1/2T} \otimes \boldsymbol{\Sigma}_{e,k}^{1/2})]\mathrm{vec}(\boldsymbol{\Xi}_e^T \otimes \mathbf{E}) \\
&= \left\|(\boldsymbol{\Sigma}_{e,k}^{1/2T} \otimes \boldsymbol{\Sigma}_{e,k}^{1/2})\mathrm{vec}(\boldsymbol{\Xi}_e^T \otimes \mathbf{E})\right\|_2^2,
\end{aligned} \tag{21}$$

where (c) is obtained by invoking the identity $\mathrm{Tr}(\mathbf{A}^H\mathbf{B}\mathbf{C}\mathbf{D}) = \mathrm{vec}^H(\mathbf{A})(\mathbf{D}^T \otimes \mathbf{B})\mathrm{vec}(\mathbf{C})$.

The $\|\mathbf{u}_{ce,k}\|_2^2$ in (19c) can be written as

$$\begin{aligned}
\|\mathbf{u}_{ce,k}\|_2^2 &= \mathbf{u}_{ce,k}^H\mathbf{u}_{ce,k} \\
&= \mathrm{vec}^H(\mathbf{E}\bar{\mathbf{G}}_{ce,k})(\boldsymbol{\Xi}_e^T \otimes \mathbf{I}_M)\boldsymbol{\Sigma}_{e,k}^{1/2}\boldsymbol{\Sigma}_{e,k}^{1/2}(\boldsymbol{\Xi}_e^* \otimes \mathbf{I}_M)\mathrm{vec}(\mathbf{E}\bar{\mathbf{G}}_{ce,k}) \\
&= \left\|\boldsymbol{\Sigma}_{e,k}^{1/2}(\boldsymbol{\Xi}_e^T \otimes \mathbf{I}_M)\mathrm{vec}(\mathbf{E}\bar{\mathbf{G}}_{ce,k})\right\|_2^2.
\end{aligned} \tag{22}$$

The constraint $y_k\mathbf{I}_n + \mathbf{A}_k \succeq \mathbf{0}, y_k \geq 0$ in (19d) can be written as

$$y_k\mathbf{I}_{N_tM} + \mathbf{\Sigma}_{e,k}^{1/2}(\mathbf{\Xi}_e^T \otimes \mathbf{E})\mathbf{\Sigma}_{e,k}^{1/2} \succeq \mathbf{0}, y_k \geq 0. \tag{23}$$

Finally, by substituting (20), (21), (22) and (23) into (19), the OC-PM Problem can be recast as

$$\min_{\mathbf{W},\mathbf{Z},\boldsymbol{\phi},\mathbf{x},\mathbf{y}} \ \mathrm{Tr}(\mathbf{W}+\mathbf{Z}) \tag{24a}$$

$$\text{s.t.} \quad \mathrm{Tr}((\mathbf{\Xi}_e^T \otimes \mathbf{E})\mathbf{\Sigma}_{e,k}) - \sqrt{-2\ln(\rho_k)} \cdot x_k + \ln(\rho_k) \cdot y_k + \mathrm{Tr}[\bar{\mathbf{G}}_{ce,k}\mathbf{\Xi}_e\bar{\mathbf{G}}_{ce,k}^H\mathbf{E}] + (\beta-1)\sigma_{e,k}^2 \geq 0, \tag{24b}$$

$$\left\| \begin{bmatrix} (\mathbf{\Sigma}_{e,k}^{1/2T} \otimes \mathbf{\Sigma}_{e,k}^{1/2})\mathrm{vec}(\mathbf{\Xi}_e^T \otimes \mathbf{E}) \\ \sqrt{2}\mathbf{\Sigma}_{e,k}^{1/2}(\mathbf{\Xi}_e^T \otimes \mathbf{I}_M)\mathrm{vec}(\mathbf{E}\bar{\mathbf{G}}_{ce,k}) \end{bmatrix} \right\|_2 \leq x_k, \tag{24c}$$

$$y_k\mathbf{I}_{N_tM} + \mathbf{\Sigma}_{e,k}^{1/2}(\mathbf{\Xi}_e^T \otimes \mathbf{E})\mathbf{\Sigma}_{e,k}^{1/2} \succeq \mathbf{0}, y_k \geq 0, \tag{24d}$$

$$(19e), (10d), (10e), (10f), (10g). \tag{24e}$$

B. Solving the Beamforming Matrix and AN Covariance Matrix

Problem (24) is not convex due to the coupled $\mathbf{\Xi}_e$ and $\boldsymbol{\phi}$. To solve it, we use the AO method to update $\{\mathbf{W}, \mathbf{Z}, \mathbf{x}, \mathbf{y}\}$ and $\boldsymbol{\phi}$ iteratively. Specifically, when $\boldsymbol{\phi}$ is fixed, Problem (24) becomes convex for $\{\mathbf{W}, \mathbf{Z}, \mathbf{x}, \mathbf{y}\}$ if the rank one constraint in (10f) is relaxed. By using the SDR technique, the problem of optimizing $\{\mathbf{W}, \mathbf{Z}, \mathbf{x}, \mathbf{y}\}$ becomes

$$\min_{\mathbf{W},\mathbf{Z},\mathbf{x},\mathbf{y}} \ \mathrm{Tr}(\mathbf{W}+\mathbf{Z}) \tag{25a}$$

$$\text{s.t.} \quad \mathrm{Tr}((\mathbf{\Xi}_e^T \otimes \mathbf{E})\mathbf{\Sigma}_{e,k}) - \sqrt{-2\ln(\rho_k)} \cdot x_k + \ln(\rho_k) \cdot y_k + \mathrm{Tr}[\bar{\mathbf{G}}_{ce,k}\mathbf{\Xi}_e\bar{\mathbf{G}}_{ce,k}^H\mathbf{E}] + (\beta-1)\sigma_{e,k}^2 \geq 0, \tag{25b}$$

$$\left\| \begin{bmatrix} (\mathbf{\Sigma}_{e,k}^{1/2T} \otimes \mathbf{\Sigma}_{e,k}^{1/2})\mathrm{vec}(\mathbf{\Xi}_e^T \otimes \mathbf{E}) \\ \sqrt{2}\mathbf{\Sigma}_{e,k}^{1/2}(\mathbf{\Xi}_e^T \otimes \mathbf{I}_M)\mathrm{vec}(\mathbf{E}\bar{\mathbf{G}}_{ce,k}) \end{bmatrix} \right\|_2 \leq x_k, \tag{25c}$$

$$y_k\mathbf{I}_{N_tM} + \mathbf{\Sigma}_{e,k}^{1/2}(\mathbf{\Xi}_e^T \otimes \mathbf{E})\mathbf{\Sigma}_{e,k}^{1/2} \succeq \mathbf{0}, y_k \geq 0, \tag{25d}$$

$$(19e), (10d), (10e). \tag{25e}$$

Problem (25) is convex, and can be solved by the CVX toolbox. The SDR for Problem (25) is tight, and the obtained $\mathbf{W}$ always satisfies the rank-one constraint. Thus, the beamforming vector

$\mathbf{w}$ can be obtained from the eigen-decomposition of $\mathbf{W}$. The tightness of the SDR is proved in the following Theorem.

Theorem 1: Assume that the optimal solution of Problem (25) is $(\mathbf{W}^\star, \mathbf{Z}^\star, \mathbf{x}^\star, \mathbf{y}^\star)$, where $\text{rank}(\mathbf{W}^\star) \geq 1$. Then there always exists another optimal solution of Problem (25), denoted as $(\tilde{\mathbf{W}}^\star, \tilde{\mathbf{Z}}^\star, \tilde{\mathbf{x}}^\star, \tilde{\mathbf{y}}^\star)$, which not only has the same objective value as $(\mathbf{W}^\star, \mathbf{Z}^\star, \mathbf{x}^\star, \mathbf{y}^\star)$, but also satisfies the rank-one constraint, i.e., $\text{rank}(\tilde{\mathbf{W}}^\star) = 1$.

Proof: Please refer to Appendix A. ■

C. Solving the Phase Shifts of IRS

When $\{\mathbf{W}, \mathbf{Z}, \mathbf{x}, \mathbf{y}\}$ are fixed, Problem (24) becomes a feasibility check problem. In order to improve the converged solution in the optimization process, the data rate inequalities for Bob in (11d) and Eve in (13a) are rewritten by introducing slack variables, and recast respectively as

$$\text{Tr}(\mathbf{G}_{cb}[\mathbf{W} - (\gamma-1)\mathbf{Z}]\mathbf{G}_{cb}^H\mathbf{E}) \geq (\gamma-1)\sigma_b^2 + \delta_0, \delta_0 \geq 0, \tag{26a}$$

$$[\boldsymbol{\phi}^T(\bar{\mathbf{G}}_{ce,k} + \triangle\mathbf{G}_{ce,k})]\boldsymbol{\Xi}_e[\boldsymbol{\phi}^T(\bar{\mathbf{G}}_{ce,k} + \triangle\mathbf{G}_{ce,k})]^H + (\beta-1)\sigma_{e,k}^2 - \delta_k \geq 0, \delta_k \geq 0. \tag{26b}$$

Then by fixing $\mathbf{W}$ and $\mathbf{Z}$ obtained in the previous iteration, and using the BTI for the outage of leaked data rate again, the optimization problem for $\boldsymbol{\phi}$ becomes

$$\max_{\boldsymbol{\phi},\boldsymbol{\delta},\mathbf{x},\mathbf{y}} \quad \|\boldsymbol{\delta}\|_1 \tag{27a}$$

$$\text{s.t.} \quad \text{Tr}((\boldsymbol{\Xi}_e^T \otimes \mathbf{E})\boldsymbol{\Sigma}_{e,k}) - \sqrt{-2\ln(\rho_k)} \cdot x_k + \ln(\rho_k) \cdot y_k$$

$$+ \text{Tr}[\bar{\mathbf{G}}_{ce,k}\boldsymbol{\Xi}_e\bar{\mathbf{G}}_{ce,k}^H\mathbf{E}] + (\beta-1)\sigma_{e,k}^2 + \delta_k \geq 0, \tag{27b}$$

$$\left\| \begin{bmatrix} (\boldsymbol{\Sigma}_{e,k}^{1/2T} \otimes \boldsymbol{\Sigma}_{e,k}^{1/2})\text{vec}(\boldsymbol{\Xi}_e^T \otimes \mathbf{E}) \\ \sqrt{2}\boldsymbol{\Sigma}_{e,k}^{1/2}(\boldsymbol{\Xi}_e^T \otimes \mathbf{I}_M)\text{vec}(\mathbf{E}\bar{\mathbf{G}}_{ce,k}) \end{bmatrix} \right\|_2 \leq x_k, \tag{27c}$$

$$y_k\mathbf{I}_{N_tM} + \boldsymbol{\Sigma}_{e,k}^{1/2}(\boldsymbol{\Xi}_e^T \otimes \mathbf{E})\boldsymbol{\Sigma}_{e,k}^{1/2} \succeq \mathbf{0}, y_k \geq 0, \tag{27d}$$

$$\text{Tr}(\mathbf{G}_{cb}[\mathbf{W} - (\gamma-1)\mathbf{Z}]\mathbf{G}_{cb}^H\mathbf{E}) \geq (\gamma-1)\sigma_b^2 + \delta_0, \tag{27e}$$

$$\boldsymbol{\delta} \succeq \mathbf{0}, \tag{27f}$$

$$(10\text{g}), \tag{27g}$$

where $\boldsymbol{\delta} = [\delta_0, \delta_1, \delta_2, \cdots, \delta_K]^T$. Then, the only nonconvex constraint in Problem (27) is the unit-modulus constraint (10g). There is no general approach to solve unit modulus constrained non-convex optimization problems optimally. To deal with it, the semidefinite programming (SDP)

method is utilized. To facilitate SDP, Problem (27) is recast as

$$\max_{\mathbf{E},\boldsymbol{\delta},\mathbf{x},\mathbf{y}} \quad \|\boldsymbol{\delta}\|_1 \tag{28a}$$

$$\text{s.t.} \quad (27b), (27c), (27d), (27e), (27f) \tag{28b}$$

$$\text{Diag}(\mathbf{E}) = \mathbf{1}_M, \tag{28c}$$

$$\mathbf{E} \succeq \mathbf{0}, \tag{28d}$$

$$\text{rank}(\mathbf{E}) = 1, \tag{28e}$$

where the constraints (28c), (28d) and (28e) are equivalent to the constraint (10g), and they are imposed to ensure that the phase shifts vector $\boldsymbol{\phi}$ with unit modulus can be recovered from $\mathbf{E}$. Problem (28) is convex except the nonconvex constraint (28e). Although the SDR method can be used to solve Problem (28) by removing constraint (28e) and solving the resultant SDP problem, the rank of the obtained $\mathbf{E}$ is generally larger than one. To handle this problem, we construct a convex constraint equivalent to the rank one constraint (28e), and the following Lemma is utilized.

Lemma 2: (Lemma 1 in [41]) For any positive semidefinite matrix $\mathbf{A}$, the following inequality holds:

$$|\mathbf{I} + \mathbf{A}| \geq 1 + \text{Tr}(\mathbf{A}), \tag{29}$$

and the equality in (29) holds if and only if $\text{rank}(\mathbf{A}) \leq 1$.

The detailed proof for Lemma 2 can be found in [41]. By invoking Lemma 2, we can construct an equivalent constraint as

$$(28e) \Leftrightarrow |\mathbf{I} + \mathbf{E}| \leq 1 + \text{Tr}(\mathbf{E}), \tag{30a}$$

$$\Leftrightarrow \log\det(\mathbf{I} + \mathbf{E}) \leq \log(1 + M), \tag{30b}$$

where $\text{Tr}(\mathbf{E}) = M$. The constraint (30b) ensures the rank one equality, thus is equivalent to constraint (28e). By utilizing the penalty-based method [42], [43] and putting the constraint (30b) into the objective function (OF) of Problem (28), Problem (28) can be cast as

$$\max_{\mathbf{E},\boldsymbol{\delta},\mathbf{x},\mathbf{y}} \quad \|\boldsymbol{\delta}\|_1 + \frac{1}{\kappa}[\log(1 + M) - \log\det(\mathbf{I} + \mathbf{E})] \tag{31a}$$

$$\text{s.t.} \quad (27b), (27c), (27d), (27e), (27f), (28c), (28d), \tag{31b}$$

where κ is a penalty factor penalizing the violation of constraint $\text{rank}(\mathbf{E}) = 1$. When the penalty factor κ is sufficiently small, the optimal solution to Problem (31) is also the optimal solution for Problem (28). The following proposition states the equivalence.

Proposition 1: Denote the optimal solution of Problem (31) with the penalty factor κ_s by $(\mathbf{E}_s, \boldsymbol{\delta}_s, \mathbf{x}_s, \mathbf{y}_s; \kappa_s)$, and denote the limit solution with sufficiently small κ_s as

$$(\bar{\mathbf{E}}, \bar{\boldsymbol{\delta}}, \bar{\mathbf{x}}, \bar{\mathbf{y}}) = \lim_{\kappa_s \to 0} (\mathbf{E}_s, \boldsymbol{\delta}_s, \mathbf{x}_s, \mathbf{y}_s; \kappa_s). \tag{32}$$

Then the limit solution $(\bar{\mathbf{E}}, \bar{\boldsymbol{\delta}}, \bar{\mathbf{x}}, \bar{\mathbf{y}})$ is an optimal solution for Problem (28).

Proof: Please refer to Appendix B. ∎

Problem (31) is not convex due to its OF. Since the $\log\det(\mathbf{I}+\mathbf{E})$ is a concave function with respect to $\mathbf{E}$, the upper bound of it can be obtained by using the first-order Taylor approximation as

$$\log\det(\mathbf{I}+\mathbf{E}) \leq (\log e)\mathrm{Tr}\{((\mathbf{I}+\mathbf{E}^{(n)})^{-1})^*(\mathbf{E}-\mathbf{E}^{(n)})\} + (\log e)\log\det(\mathbf{I}+\mathbf{E}^{(n)}), \tag{33}$$

where e denotes natural logarithm, and the $\frac{\partial \ln(|\det(\mathbf{X})|)}{\partial \mathbf{X}} = (\mathbf{X}^{-1})^T$ is utilized. By substituting (33) into the OF of Problem (31) and removing the constant terms, Problem (31) can be written as

$$\max_{\mathbf{E},\boldsymbol{\delta},\mathbf{x},\mathbf{y}} \quad \|\boldsymbol{\delta}\|_1 + \frac{1}{\kappa}[-(\log e)\mathrm{Tr}\{((\mathbf{I}+\mathbf{E}^{(n)})^{-1})^*\mathbf{E}\}] \tag{34a}$$

$$\text{s.t.} \quad (27b), (27c), (27d), (27e), (27f), (28c), (28d). \tag{34b}$$

Problem (34) is jointly convex with respect to $\{\mathbf{E}, \boldsymbol{\delta}, \mathbf{x}, \mathbf{y}\}$, hence it can be efficiently solved by standard convex program solvers such as CVX. A rank-one solution $\mathbf{E}^\star$ can be obtained by solving Problem (34) for a sufficiently small value of penalty factor κ, and the optimal $\boldsymbol{\phi}^\star$ can be recovered from the eigen-decomposition of $\mathbf{E}^\star$. The maximum value of Problem (34) serves as the global lower bound for the optimal value of Problem (31).

D. Overall algorithm to Solve Problem (24)

The overall AO algorithm proposed is summarized in Algorithm 1. By iteratively solving Problem (25) and Problem (34) optimally in Step 3 and Step 4 in Algorithm 1, the transmit power can be monotonically reduced with guaranteed convergence. We prove the convergence of the proposed AO algorithm as follows. Denote the OF value of Problem (24) and Problem (25) (i.e., the transmit power) with a feasible solution $(\boldsymbol{\phi}, \mathbf{w}, \mathbf{Z})$ as $p(\boldsymbol{\phi}, \mathbf{w}, \mathbf{Z})$. As shown in step 4 of Algorithm 1, if there exists a feasible solution to Problem (34), i.e., $(\boldsymbol{\phi}^{(n+1)}, \mathbf{w}^{(n)}, \mathbf{Z}^{(n)})$ exists, it is also feasible to Problem (25). Then, $(\boldsymbol{\phi}^{(n+1)}, \mathbf{w}^{(n+1)}, \mathbf{Z}^{(n+1)})$ and $(\boldsymbol{\phi}^{(n)}, \mathbf{w}^{(n)}, \mathbf{Z}^{(n)})$ in step 3 are the feasible solutions to Problem (25) in the $(n+1)$th and (n)th iterations, respectively. It then follows that $p(\boldsymbol{\phi}^{(n+1)}, \mathbf{w}^{(n+1)}, \mathbf{Z}^{(n+1)}) \overset{(d)}{\leq} p(\boldsymbol{\phi}^{(n+1)}, \mathbf{w}^{(n)}, \mathbf{Z}^{(n)}) \overset{(e)}{=} p(\boldsymbol{\phi}^{(n)}, \mathbf{w}^{(n)}, \mathbf{Z}^{(n)})$, where (d) holds since for

Algorithm 1 Alternating Optimization Algorithm

1: Parameter Setting: Set the maximum number of iterations $n_{\max}$ and the first iterative number $n = 1$; Give the penalty factor κ and error tolerance ε;

2: Initialize the variables $\mathbf{w}^{(1)}$, $\mathbf{Z}^{(1)}$ and $\boldsymbol{\phi}^{(1)}$ in the feasible region; Compute the OF value of Problem (24) (i.e., the transmit power) as $p(\mathbf{w}^{(1)}, \mathbf{Z}^{(1)})$;

3: Solve Problem (25) to obtain the $\mathbf{w}^{(n)}, \mathbf{Z}^{(n)}$ by fixing $\boldsymbol{\phi}^{(n)}$; Calculate the OF value of Problem (25) as $p(\mathbf{w}^{(n)}, \mathbf{Z}^{(n)})$;

4: Solve Problem (34) to obtain $\mathbf{E}^{(n+1)}$ by fixing $\mathbf{w}^{(n)}, \mathbf{Z}^{(n)}$; then obtain $\boldsymbol{\phi}^{(n+1)}$ from the eigen-decomposition of $\mathbf{E}^{(n+1)}$;

5: If $\left|p(\mathbf{w}^{(n+1)},\mathbf{Z}^{(n+1)})-p(\mathbf{w}^{(n)},\mathbf{Z}^{(n)})\right|/p(\mathbf{w}^{(n)},\mathbf{Z}^{(n)}) < \varepsilon$ or $n \geq n_{\max}$ or Problem (34) becomes infeasible, terminate. Otherwise, update $n \leftarrow n+1$ and jump to Step 3.

given $\boldsymbol{\phi}^{(n+1)}$ in step 3 of Algorithm 1, $\mathbf{w}^{(n+1)}, \mathbf{Z}^{(n+1)}$ is the optimal solution to Problem (25); and (e) holds because the OF of Problem (25) is regardless of $\boldsymbol{\phi}$ and only depends on $\mathbf{w}, \mathbf{Z}$.

Based on the algorithm description, the complexity analysis of the proposed AO algorithm is performed. Since both the resulting convex Problem (25) and Problem (34) involve LMI, second-order cone (SOC) constraints and linear constraints that can be solved by a standard interior point method [44], the computational complexity of the proposed algorithm can be measured in terms of their worst-case runtime, and the general expression for complexity (by ignoring the complexity of the linear constraints) has been given in [36]. For Problem (25), the number of variables is $n_1 = 2N_t^2$; the number of LMIs in (25d) is K with the size of N_tM; the number of LMIs in (10d) and (10e) is 2 with the size of N_t; the number of SOC in (25c) is K with the size of $N_t^2M^2+N_tM$. By utilizing the general expression in [36] and ignoring the linear constraints in (25b) and (19e), the approximate complexity of Problem (25) is

$$\begin{aligned} C_{\mathbf{WZ}} =& \mathcal{O}\{[KN_tM + 2N_t + 2K]^{1/2} n_1 [n_1^2 + n_1(KM^2+2)N_t^2 \\ &+ (KM^3+2)N_t^3 + n_1 K N_t^2 M^2 (N_tM+1)^2]\}, \end{aligned} \tag{35}$$

where $n_1 = 2N_t^2$ and $\mathcal{O}\{\cdot\}$ is the big-O notation. For Problem (34), the number of variables is M^2; the number of LMIs in (27d) is K with the size of N_tM; the number of LMI in (28d) is 1 with the size of M; the number of SOC in (27c) is K with the size of $N_t^2M^2+N_tM$. By utilizing the general expression in [36] and ignoring the linear constraints in (27b), (27e), (27f) and (28c),

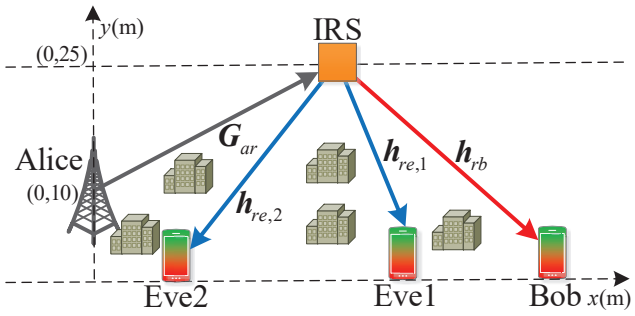


Fig. 2. The simulated secure communication scenario

the approximate complexity of Problem (34) is

$$C_{\mathbf{E}} = \mathcal{O}\{[KN_tM + M + 2K]^{1/2} n_2 [n_2^2 + n_2 M^2 (KN_t^2 + 1) + M^3 (KN_t^3 + 1) + n_2 K N_t^2 M^2 (N_t M + 1)^2]\}, \tag{36}$$

where $n_2 = M^2$.

V. Simulation Results

A. Simulation Setup

One appealing benefit of deploying an IRS in secure wireless systems is to establish favorable communication links for legitimate user that is blocked by obstacles. We consider a scenario where the direct links from the AP to the legitimate user and eavesdroppers are blocked. The simulated system model is shown in Fig. 2. Alice is located at (0,10) m. The IRS is installed at a height of 25 m. Bob and Eves can be randomly distributed on the x axis.

The channel matrix between Alice and the IRS is modeled as $\mathbf{G}_{ar} \in \mathbb{C}^{M \times N_t}$, which is

$$\mathbf{G}_{ar}^T = \sqrt{L_0 d_{ar}^{-\alpha_{ar}}} \left(\sqrt{\frac{\beta_{ar}}{1+\beta_{ar}}} \mathbf{G}_{ar}^{LOS} + \sqrt{\frac{1}{1+\beta_{ar}}} \mathbf{G}_{ar}^{NLOS} \right), \tag{37}$$

where $L_0 = \left(\frac{\lambda_c}{4\pi}\right)^2$ is a constant with λ_c being the wavelength of the center frequency of the information carrier. The distance between the AP and the IRS is denoted by d_{ar}, while α_{ar} denotes the corresponding path loss exponent. The small-scale fading is assumed to be Rician fading with Rician factor β_{ar}. The $\mathbf{G}_{ar}^{LOS}$ denotes the light of sight (LoS) component of the Alice-IRS channel. By assuming that the antennas at Alice and the passive reflecting elements at the IRS are both

TABLE I

SIMULATION PARAMETERS

Carrier center frequency	2.4GHz
Path loss exponents for Alice-IRS channels α_{ar}, IRS-Bob channels α_{rb}, IRS-Eve channels α_{re},	2
Rician factor for Alice-IRS channel, β_{ar}	5
Rician factor for IRS-Bob channel, β_{rb}	5
Rician factor for IRS-Eve channel, β_{re}	5
Noise power at Bob and Eves, σ_b^2, $\sigma_{e,k}^2$	-75dBm
Penalty factor, κ	5×10^{-6}
Convergence tolerance, ε	10^{-3}
Outage probability, ρ_k	0.05

arranged in a uniform linear array (ULA), the $\mathbf{G}_{ar}^{LOS}$ can be modeled as $\mathbf{G}_{ar}^{LOS} = \mathbf{c}\mathbf{d}^H$, where $\mathbf{c} = [c_1, c_2, \cdots, c_{N_t}]^T$ denotes the transmit steering vector of the Alice, and $\mathbf{d} = [d_1, d_2, \cdots, d_M]^T$ denotes the receive steering vector of the IRS. The mth elements of $\mathbf{c}$ and $\mathbf{d}$ are written as

$$c_m = \exp(j2\pi \frac{d_a}{\lambda_c}(m-1)\sin\theta_a), \tag{38a}$$

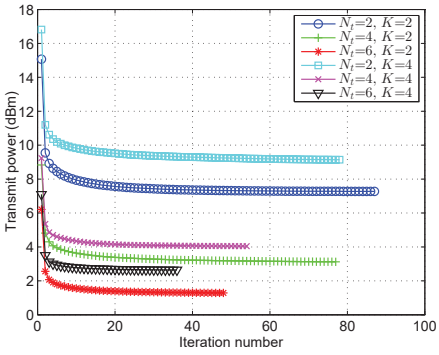
$$d_m = \exp(j2\pi \frac{d_r}{\lambda_c}(m-1)\sin\theta_r), \tag{38b}$$

where d_a and d_r are respectively the element intervals at Alice and the IRS; θ_a represents the elevation angle of the LoS link from Alice; θ_r represents the elevation angle of the LoS link to the IRS. We set $\frac{d_a}{\lambda_c} = \frac{d_r}{\lambda_c} = 0.5$, $\theta_a = \tan^{-1}(\frac{y_r - y_a}{x_r - x_a})$, and $\theta_r = \pi - \theta_a$. $\mathbf{G}_{ar}^{NLOS}$ denotes the non-LoS (NLoS) component of the Alice-IRS channel, which is modeled as Rayleigh fading. The channels of the IRS-Bob link and IRS-Eve link are modeled similarly, where both the LoS component and NLoS component exist simultaneously.

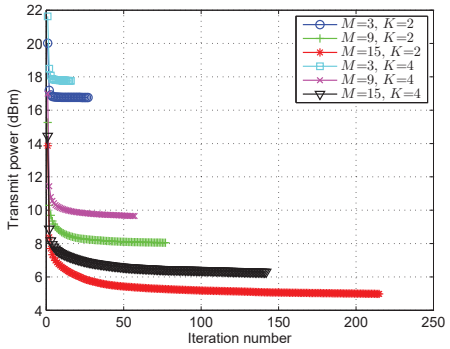
For the statistical cascaded CSI error model, the variance matrix of $\mathbf{g}_{ce,k} = \text{vec}(\triangle \mathbf{G}_{ce,k})$ is defined as $\mathbf{\Sigma}_{e,k} = \varepsilon_{g,k}^2 \mathbf{I}$, where $\varepsilon_{g,k}^2 = \delta_{g,k}^2 \left\| \text{vec}(\bar{\mathbf{G}}_{ce,k}) \right\|_2^2$. $\delta_{g,k}^2 \in [0, 1)$ is the normalized CSI error, which measures the relative amount of CSI uncertainties. The system parameters used in the following simulations are listed in Table I.

B. Benchmark Schemes

We demonstrate the advantage of the proposed AO algorithm by comparing its performance with the following benchmark schemes:



(a) Convergence of the proposed algorithm with different values of N_t and K for $M = 10$.

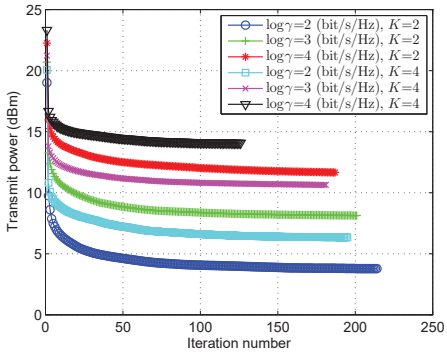


(b) Convergence of the proposed algorithm with different values of M and K for $N_t = 2$.

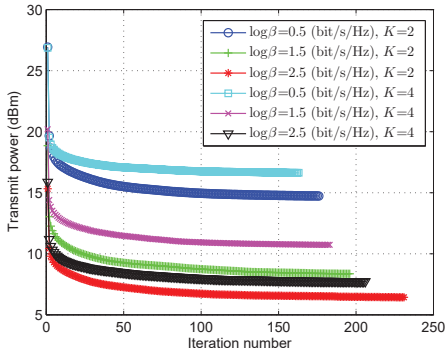
Fig. 3. Convergence of the proposed algorithm with $\log \gamma = 3$ bit/s/Hz and $\log \beta = 1$ bit/s/Hz, $K = 2$, $\delta_{g,k}^2 = 0.0001$, $\forall k$. The location of AP is (0, 10) m, the location of IRS is (100, 25) m, the location of Bob is (180, 0) m, and the locations of Eves are (160, 0) m and (170, 0) m.

- Random-MRT: It performs maximum ratio transmission (MRT) based beamforming design, i.e., $\mathbf{w} = \sqrt{p_w}\frac{\hat{\mathbf{h}}_b}{\|\hat{\mathbf{h}}_b\|} = \sqrt{p_w}\frac{\mathbf{G}_{ar}^H \mathbf{\Phi}^H \mathbf{h}_{rb}}{\|\hat{\mathbf{h}}_b\|}$, where p_w is the power allocated to legitimate user Bob. It applies an isotropic AN [45], i.e., the AN covariance matrix is chosen as $\mathbf{Z} = p_z \mathbf{P}_{\hat{\mathbf{h}}_b}^{\perp}$, where $\mathbf{P}_{\hat{\mathbf{h}}_b}^{\perp} = \mathbf{I}_{N_t} - \hat{\mathbf{h}}_b \hat{\mathbf{h}}_b^H / \left\|\hat{\mathbf{h}}_b\right\|^2$ is the orthogonal complement projector of $\hat{\mathbf{h}}_b$, and p_z is the power invested on AN. Both the beamforming precoder $\mathbf{w}$ and the AN covariance matrix $\mathbf{Z}$ rely on the phase shifts $\mathbf{\Phi}$ of IRS. For Random-MRT schemes, we assume the phase shifts $\mathbf{\Phi}$ of IRS is randomly chosen. The allocated power p_w to beamform and the allocated power p_z to AN are optimized to satisfy the secrecy constraints in Problem (25).
- Optimized-MRT: Based on the MRT beamforming and isotropic AN, the phase shifts $\mathbf{\Phi}$ of IRS is obtained by using the proposed method. The allocated power p_w to beamform and the allocated power p_z to AN are optimized to satisfy the secrecy constraints in Problem (25).
- Random-IRS: It does not optimize the phase shifts of IRS, and the phase shifts of IRS $\mathbf{\Phi}$ is randomly selected. The beamforming matrix $\mathbf{W}$ and AN covariance matrix $\mathbf{Z}$ are optimized by solving Problem (25).

The Random-MRT and Optimized-MRT schemes only exploit the CSI of Bob by assuming that the CSI of Eves is unknown, while the Random-IRS scheme and the proposed algorithm exploit the CSI of both Bob and Eves to realize robust transmission design. For Random-IRS scheme, the phase shifts of IRS $\mathbf{\Phi}$ are not optimized, and only the beamforming matrix $\mathbf{W}$ and AN covariance matrix $\mathbf{Z}$ are optimized robustly, while the proposed algorithm optimizes these variables jointly.



(a) Convergence of the proposed algorithm for different values of $\log\gamma$ and K for $\log\beta = 0.5$ bit/s/Hz.



(b) Convergence of the proposed algorithm for different values of $\log\beta$ and K for $\log\gamma = 4$ bit/s/Hz.

Fig. 4. Convergence of the proposed algorithm with $N_t = 2$, $M = 15$, $K = 2$, $\delta_{g,k}^2 = 0.0001$, $\forall k$. The location of AP is (0, 10) m, the location of IRS is (50, 25) m, the location of Bob is (180, 0) m, and the locations of Eves are (20, 0) m and (10, 0) m.

C. Convergence Performance

The convergence of the proposed method with different numbers of transmit antennas, different numbers of IRS phase shifts and different numbers of Eves is investigated in Fig. 3. As observed from Fig. 3(a), the proposed algorithm is likely to converge with fewer iterations when the number of transmit antennas N_t increases and when the number of Eves K increases. With increased N_t, the dimensions of optimization variables $\mathbf{W}$ and $\mathbf{Z}$ in the CVX problem of each iteration increase, and with increased K, the number of constraints in the two subproblems of the AO algorithm increases. Thus, more computation time is required for each iteration with a larger N_t or a larger K, and the total consumed computation time by the proposed algorithm still increases even with fewer iterations. Fig. 3(b) shows the convergence of the proposed algorithm with different values of M and K. It is seen that the proposed algorithm is likely to converge with more iterations with a larger M, and is likely to converge with fewer iterations with a larger K. Similarly, the dimension of the optimization variable $\mathbf{\Phi}$ in each iteration increases with M, and the number of constraints in each iteration increases with K. Thus the computation time of the proposed algorithm increases with M and K. The convergence of the proposed algorithm for different values of $\log\gamma$ and $\log\beta$ is shown in Fig. 4. It is demonstrated that more iterations are required by decreasing the minimum channel capacity $\log\gamma$ of the legitimate user and increasing the maximum tolerable channel capacity $\log\beta$ of the eavesdroppers. Fewer iterations are required with a larger K. Since

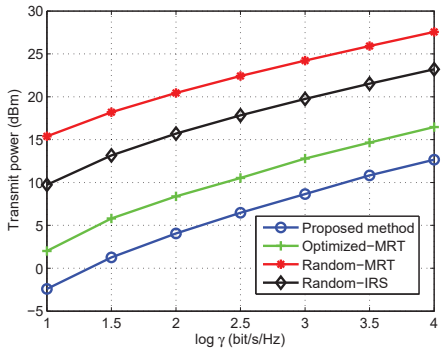


Fig. 5. Transmit power versus the minimum channel capacity $\log\gamma$ of the legitimate user with $N_t = 2$, $M = 15$, $K = 2$, and $\delta_{g,k}^2 = 0.0001$, $\forall k$. The location of AP is (0, 10) m, the location of IRS is (50, 25) m, the location of Bob is (180, 0) m, and the locations of Eves are (20, 0) m and (10, 0) m.

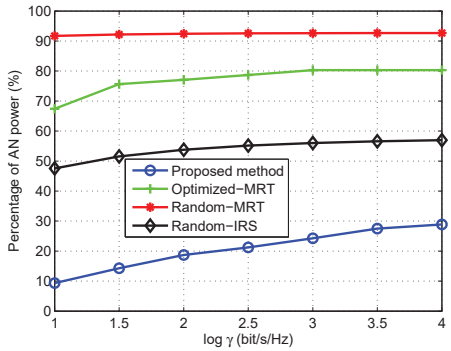


Fig. 6. Percentage of AN power versus the minimum channel capacity $\log\gamma$ of the legitimate user with $N_t = 2$, $M = 15$, $K = 2$, and $\delta_{g,k}^2 = 0.0001$, $\forall k$. The location of AP is (0, 10) m, the location of IRS is (50, 25) m, the location of Bob is (180, 0) m, and the locations of Eves are (20, 0) m and (10, 0) m.

the dimension of the CVX problem in each iteration is not changed with different $\log\gamma$ and $\log\beta$ for a given K, more iterations for a given K imply more computation time.

D. Transmit Power Versus the Minimum Channel Capacity of the Legitimate User

In Fig. 5, we show the transmit power versus the minimum channel capacity $\log\gamma$ of the legitimate user, while limiting the information leakage to the potential eavesdroppers as $\log\beta = 0.5$ bit/s/Hz. From Fig. 5, we find that the transmit power increases monotonically with the minimum channel capacity of the Bob. This is because more transmit power is required to satisfy the increased data rate of Bob. It is also seen that the power budget of the proposed method is the lowest among all these benchmark schemes, which verifies the superiority of the proposed method. By comparing with the Random-IRS and Random-MRT schemes, our proposed algorithm and the Optimized-MRT scheme require much lower transmit power. This confirms that optimizing the phase shifts of the IRS can reconfigure the wireless propagation environment by enhancing the channel quality of the BS-Bob link and degrading the channel condition of the BS-Eve links. When the phase shifts of the IRS have been optimized, our proposed algorithm is better than the optimized-MRT scheme. That's because the MRT scheme roughly transmits the AN isotropically on the orthogonal complement subspace of the Bob's cascaded channel, while the proposed algorithm jointly optimizes the transmit beamforming and AN spatial distribution specifically by making full use of both the Bob's and Eves' cascaded channels.

The percentage of AN power in the total transmit power versus the minimum channel capacity $\log\gamma$ of the legitimate user is shown in Fig. 6. When $\log\gamma$ is small, a small portion of transmit

power is allocated to transmit AN to deteriorate the achievable rates of Eves. When $\log\gamma$ increases, more transmit power should be emitted, thus the data rates of both Bob and Eve increase. To control the data rates of Eves under the threshold $\log\beta$, more AN power should be emitted to interfere the Eves. It is also observed that by optimizing the phase shifts of the IRS, less AN power is required to impair the Eves for the proposed algorithm and Optimized-MRT scheme in comparison to the random schemes. Both the Random-MRT scheme and the Optimized-MRT scheme have to transmit much more AN power than the Random-IRS scheme and the proposed algorithm. The reason is that the AN is transmitted to interfere all the potential Eves distributed in the null subspace of the Bob's cascaded channel for the MRT schemes, while the AN is designed to interfere the specifical Eves according to the exact Eves' locations for the Random-IRS scheme and the proposed algorithm.

E. Transmit Power Versus the Maximum Tolerable Channel Capacity of the Eavesdroppers

Fig. 7 depicts the transmit power versus the maximum tolerable channel capacity $\log\beta$ of the Eves by assuming that the minimum channel capacity of the legitimate user is $\log\gamma$=4 bit/s/Hz. As shown in Fig. 7, when $\log\beta$ becomes large, more information leakage is allowed for the Eves, and less AN power is required. Thus the transmit power decreases versus the increased maximum tolerable channel capacity $\log\beta$ of Eves. When $\log\beta$ becomes larger, the consumed transmit power of the Random-MRT scheme approaches that of the Random-IRS scheme, while the consumed transmit power of the Optimized-MRT scheme approaches that of the proposed algorithm. This reveals that, when the allowable information leakage to Eves is relatively large, the transmit power of the MRT schemes by exploiting only the Bob's CSI is comparable with that of the Random-IRS and proposed schemes by exploiting both the Bob's and the Eves' CSI, and the importance of exploiting the Eves' CSI is weakened. However, the proposed method still consumes the lowest transmit power among all these schemes.

The percentage of AN power in the total transmit power versus the maximum tolerable channel capacity $\log\beta$ of the Eves is depicted in Fig. 8. When $\log\beta$ becomes large, the percentages of AN power for all methods reduce. When $\log\beta$ is small, a large portion of the transmit power is allocated to transmit AN to deteriorate the achievable rates of the potential eavesdroppers. As $\log\beta$ increases, the constraints on Eves' data rates are relaxed, and more transmit power is allocated for Bob. When $\log\beta$ is relatively small, the joint optimization of beamforming and AN covariance seems more effective to reduce the transmit power than the optimization of the phase shifts of the IRS, thus the AN percentage of the Random-IRS scheme is lower than the Optimized-MRT

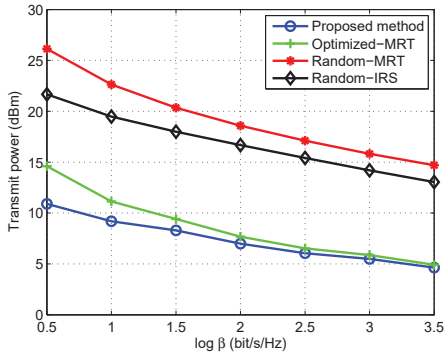


Fig. 7. Transmit power versus the maximum tolerable channel capacity $\log \beta$ of the eavesdropper with $N_t = 2$, $M = 15$, $K = 2$ and $\delta_{g,k}^2 = 0.0001$, $\forall k$. The location of AP is (0, 10) m, the location of IRS is (50, 25) m, the location of Bob is (180, 0) m, and the locations of Eves are (20, 0) m and (10, 0) m.

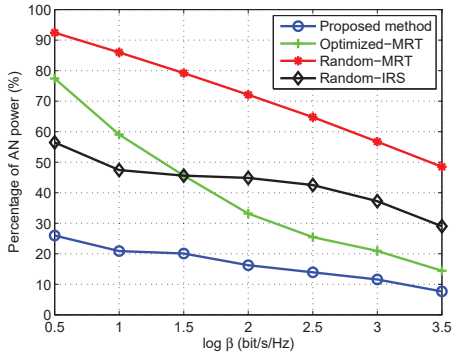


Fig. 8. Percentage of AN power versus the maximum channel capacity $\log \beta$ of the eavesdropper with $N_t = 2$, $M = 15$, $K = 2$ and $\delta_{g,k}^2 = 0.0001$, $\forall k$. The location of AP is (0, 10) m, the location of IRS is (50, 25) m, the location of Bob is (180, 0) m, and the locations of Eves are (20, 0) m and (10, 0) m.

scheme. When $\log \beta$ is relatively large, the optimization of the phase shifts of the IRS becomes more effective to reduce the transmit power than the joint optimization of beamforming and AN covariance, thus the AN percentage of the Optimized-MRT scheme is lower than the Random-IRS scheme. However, the AN percentage of the proposed algorithm remains to be the lowest.

F. Transmit Power Versus the Number of IRS Elements

We further examine the minimum transmit power consumption versus different numbers of IRS elements in Fig. 9. The maximum tolerable channel capacity of the Eves is $\log \beta = 1$ bit/s/Hz. The required minimum data rate for the Bob is set as $\log \gamma = 3$ bit/s/Hz. The transmit power is reduced when the IRS element number M increases. To enhance the security, the phase shifts of the IRS can be optimized to help the data transmission for Bob while degrade the data transmission for Eves. A large value of M will increase the beamforming gain of the IRS, thus less transmit power is required. The transmit power of the Optimized-MRT scheme and the proposed algorithm decreases more quickly than that of the Random-MRT scheme and Random-IRS scheme, which shows that optimizing the phase shifts of IRS can reduce the transmit power effectively. It is also observed that when M is large enough (e.g., $M = 21$), the transmit power of all these four schemes increases slightly. That's because a larger CSI error will be introduced with a very large M, which requires more transmit power to compensate for the channel errors, thus outweighs the benefit of increasing M for reducing the transmit power. Compared with the other benchmark schemes, the proposed algorithm requires the lowest transmit power, which validates the superiority of the

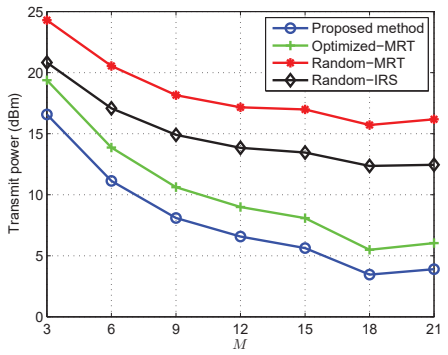


Fig. 9. Transmit power versus the IRS element number M with $N_t = 2$, $K = 2$, $\delta_{g,k}^2 = 0.0001$, $\forall k$, $\log \beta = 1$ bit/s/Hz and $\log \gamma = 3$ bit/s/Hz. The location of AP is (0, 10) m, the location of IRS is (100, 25) m, the location of Bob is (180, 0) m, and the locations of Eves are (160, 0) m and (170, 0) m.

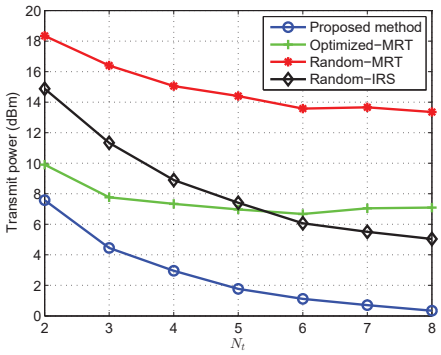


Fig. 10. Transmit power versus transmit antenna number N_t with $M = 10$, $K = 2$, $\delta_{g,k}^2 = 0.0001$, $\forall k$, $\log \beta = 1$ bit/s/Hz and $\log \gamma = 3$ bit/s/Hz. The location of AP is (0, 10) m, the location of IRS is (100, 25) m, the location of Bob is (180, 0) m, and the locations of Eves are (160, 0) m and (170, 0) m.

proposed algorithm.

G. Transmit Power Versus the Number of Transmit Antennas

We further examine the transmit power consumption versus different number of transmit antennas N_t in Fig. 10. The maximum tolerable channel capacity of the Eves is $\log \beta = 1$ bit/s/Hz. The required minimum data rate for the Bob is set as $\log \gamma = 3$ bit/s/Hz. Fig. 10 shows that the transmit power of all schemes reduces when the number of transmit antennas increases. A large value of N_t helps improve the channels for both the Bob and Eves. Then, the robust beamforming design for Bob and AN design for Eves are easier to achieve with a larger N_t. Thus, the required transmit power reduces with increased N_t. Compared with the Random-MRT scheme and Optimized-MRT scheme, the required transmit power of the Random-IRS scheme and proposed algorithm decreases more quickly with N_t. The transmit power of the Random-MRT scheme and Optimized-MRT scheme reduces slowly (even increases slightly) with increased N_t, which means that optimizing the phase shifts of the IRS is necessary to compensate the power increase for larger CSI errors with a larger N_t. The required transmit power of Random-IRS scheme is further reduced by the proposed algorithm with optimized phase shifts of the IRS.

H. Transmit Power Versus CSI Uncertainty

Fig. 11 shows the transmit power versus the normalized CSI error variance δ^2, where the normalized CSI errors of different eavesdroppers are assumed to be identical, i.e., $\delta_{g,k} = \delta$, $\forall k$. As

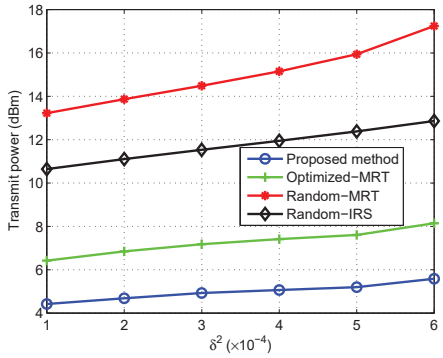


Fig. 11. Transmit power versus normalized CSI error variance δ^2 with $N_t = 2$, $M = 15$, $K = 2$, $\log \beta = 1$ bit/s/Hz and $\log \gamma = 3$ bit/s/Hz. The location of AP is (0, 10) m, the location of IRS is (100, 25) m, the location of Bob is (180, 0) m, and the locations of Eves are (160, 0) m and (170, 0) m.

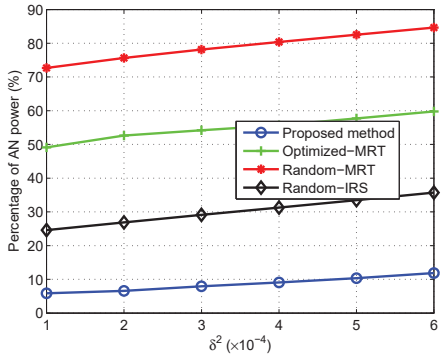


Fig. 12. Percentage of AN power versus normalized CSI error variance δ^2 with $N_t = 2$, $M = 15$, $K = 2$, $\log \beta = 1$ bit/s/Hz and $\log \gamma = 3$ bit/s/Hz. The location of AP is (0, 10) m, the location of IRS is (100, 25) m, the location of Bob is (180, 0) m, and the locations of Eves are (160, 0) m and (170, 0) m.

can be observed, for the proposed scheme and the other benchmark schemes, the transmit power increases as the quality of the CSI degrades. When the CSI errors of Eves' channels increase, more AN power is required to interfere Eves for information leakage limitation, thus the allocated power for beamforming must also increase to guarantee Bob's data requirement. This leads to the increase of total transmit power with deteriorating channel. In comparison to other three benchmark schemes, the required transmit power of the proposed method has been significantly reduced. Both the Random MRT and Optimized MRT schemes need much more transmit power to impair the Eves, which shows that the utilization of Eves' CSI can help reduce the transmit power. The Random IRS scheme also needs much more transmit power than the proposed method, which indicates the effectiveness of the optimization on the IRS phase shifts in the robust transmission design. It can also be found from Fig. 11 that the transmit power of the Random MRT and Optimized MRT schemes increases more quickly than the proposed method, which shows that they are more sensitive to the CSI errors, and that the proposed method is more robust with respect to the CSI error.

Fig. 12 gives the percentage of AN power in the total transmit power versus the normalized CSI error variance δ^2. It is shown in Fig. 12 that the percentage of AN power increases with CSI estimation error. When the quality of the CSI degrades, more AN power is required to interfere the eavesdroppers. In comparison to other three benchmark schemes, the percentage of AN power in the total transmit power for the proposed method is reduced to be less than 10%. Without the Eves's CSI, more AN power is required to impair the Eves. On the contrary, by exploiting the

Eves' CSI robustly, the AN interference becomes more effective, thus less percentage of AN power is required. Compared with the Random IRS scheme, less AN power percentage is required for the proposed method, which demonstrates that the optimization of IRS phase shifts helps impair the Eves' channel, thus makes less AN power interferes the Eves more effectively.

VI. Conclusion

In this paper, we designed the outage constrained robust transmission in the secure IRS-aided wireless communications. The statistical CSI error of cascaded IRS channel was taken into consideration for the first time in secure communications, and the OC-PM problem was formulated to minimize the transmit power by jointly optimizing the transmit beamforming vector, AN covariance, and IRS phase shifts. To solve it, an AO based method was proposed, where the optimization variables are optimized alternately. The BTI was utilized to tackle the chance constraint. The SDR scheme was utilized when designing the beamforming vector and IRS phase shifts. Specifically, the nonconvex unit modulus constrained was handled by a penalty method. Simulation results have verified the effectiveness of IRS on enhancing the physical layer security of wireless communications.

Appendix A
Proof of Theorem 1

Denote the optimal solution of Problem (25) by $(\mathbf{W}^\star, \mathbf{Z}^\star, \mathbf{x}^\star, \mathbf{y}^\star)$, and define the projection matrix as $\mathbf{P} = \mathbf{W}^{\star\frac{1}{2}}\hat{\mathbf{h}}_b\hat{\mathbf{h}}_b^H\mathbf{W}^{\star\frac{1}{2}}/\left\|\mathbf{W}^{\star\frac{1}{2}}\hat{\mathbf{h}}_b\right\|^2$, where $\hat{\mathbf{h}}_b^H = \boldsymbol{\phi}^T\mathbf{G}_{cb}$. Then, we construct a rank-one solution $\tilde{\mathbf{W}}^\star = \mathbf{W}^{\star\frac{1}{2}}\mathbf{P}\mathbf{W}^{\star\frac{1}{2}}$ and $\tilde{\mathbf{Z}}^\star = \mathbf{W}^{\star\frac{1}{2}}\mathbf{P}^\perp\mathbf{W}^{\star\frac{1}{2}} + \mathbf{Z}^\star$, where $\mathbf{P}^\perp = \mathbf{I}_{N_t} - \mathbf{P}$.

Firstly, we check the objective value of Problem (25) with the constructed solution $(\tilde{\mathbf{W}}^\star, \tilde{\mathbf{Z}}^\star)$. By substituting $(\tilde{\mathbf{W}}^\star, \tilde{\mathbf{Z}}^\star)$ into the OF of Problem (25), we have

$$\mathrm{Tr}(\tilde{\mathbf{W}}^\star + \tilde{\mathbf{Z}}^\star) = \mathrm{Tr}(\mathbf{W}^{\star\frac{1}{2}}\mathbf{W}^{\star\frac{1}{2}} + \mathbf{Z}^\star) = \mathrm{Tr}(\mathbf{W}^\star + \mathbf{Z}^\star), \tag{39}$$

which means that the objective value obtained by solution $(\tilde{\mathbf{W}}^\star, \tilde{\mathbf{Z}}^\star)$ is the same as that obtained by solution $(\mathbf{W}^\star, \mathbf{Z}^\star)$.

Secondly, we check the constraint (19e) with the constructed solution $(\tilde{\mathbf{W}}^\star, \tilde{\mathbf{Z}}^\star)$. The constraint

(19e) is equivalent to (11b). By substituting $(\tilde{\mathbf{W}}^\star, \tilde{\mathbf{Z}}^\star)$ into the left hand side of (11b), we have

$$\mathrm{Tr}([\tilde{\mathbf{W}}^\star - (\gamma-1)\tilde{\mathbf{Z}}^\star]\hat{\mathbf{h}}_b\hat{\mathbf{h}}_b^H) \tag{40a}$$

$$= \mathrm{Tr}([\mathbf{P} - (\gamma-1)\mathbf{P}^\perp]\mathbf{W}^{\star\frac{1}{2}}\hat{\mathbf{h}}_b\hat{\mathbf{h}}_b^H\mathbf{W}^{\star\frac{1}{2}}) - \mathrm{Tr}((\gamma-1)\mathbf{Z}^\star\hat{\mathbf{h}}_b\hat{\mathbf{h}}_b^H) \tag{40b}$$

$$= \mathrm{Tr}([\mathbf{P} - (\gamma-1)\mathbf{P}^\perp]\mathbf{P}\left\|\mathbf{W}^{\star\frac{1}{2}}\hat{\mathbf{h}}_b\right\|^2) - \mathrm{Tr}((\gamma-1)\mathbf{Z}^\star\hat{\mathbf{h}}_b\hat{\mathbf{h}}_b^H) \tag{40c}$$

$$= \mathrm{Tr}(\mathbf{P}\left\|\mathbf{W}^{\star\frac{1}{2}}\hat{\mathbf{h}}_b\right\|^2) - \mathrm{Tr}((\gamma-1)\mathbf{Z}^\star\hat{\mathbf{h}}_b\hat{\mathbf{h}}_b^H) \tag{40d}$$

$$= \mathrm{Tr}(\mathbf{W}^\star\hat{\mathbf{h}}_b\hat{\mathbf{h}}_b^H) - \mathrm{Tr}((\gamma-1)\mathbf{Z}^\star\hat{\mathbf{h}}_b\hat{\mathbf{h}}_b^H) \tag{40e}$$

$$= \mathrm{Tr}([\mathbf{W}^\star - (\gamma-1)\mathbf{Z}^\star]\hat{\mathbf{h}}_b\hat{\mathbf{h}}_b^H). \tag{40f}$$

Based on (40), we have

$$\mathrm{Tr}([\mathbf{W}^\star - (\gamma-1)\mathbf{Z}^\star]\hat{\mathbf{h}}_b\hat{\mathbf{h}}_b^H) \geq (\gamma-1)\sigma_b^2 \tag{41a}$$

$$\Leftrightarrow \mathrm{Tr}([\tilde{\mathbf{W}}^\star - (\gamma-1)\tilde{\mathbf{Z}}^\star]\hat{\mathbf{h}}_b\hat{\mathbf{h}}_b^H) \geq (\gamma-1)\sigma_b^2. \tag{41b}$$

Thus, the constructed solution $(\tilde{\mathbf{W}}^\star, \tilde{\mathbf{Z}}^\star)$ satisfies the constraint in (19e) (i.e., (11b)).

Thirdly, we check constraints (25b), (25c) and (25d) with the constructed solution $(\tilde{\mathbf{W}}^\star, \tilde{\mathbf{Z}}^\star)$. Since it is computationally intractable to check whether the constructed solution satisfies the constraints (25b), (25c) and (25d) directly, we instead consider the equivalent constraint (12b). By substituting $(\tilde{\mathbf{W}}^\star, \tilde{\mathbf{Z}}^\star)$ into the left hand side of (12b), we have

$$\mathrm{Tr}([\tilde{\mathbf{W}}^\star - (\beta-1)\tilde{\mathbf{Z}}^\star]\hat{\mathbf{h}}_{e,k}\hat{\mathbf{h}}_{e,k}^H) \tag{42a}$$

$$= \mathrm{Tr}(\mathbf{W}^{\star\frac{1}{2}}[\mathbf{P} - (\beta-1)\mathbf{P}^\perp]\mathbf{W}^{\star\frac{1}{2}}\hat{\mathbf{h}}_{e,k}\hat{\mathbf{h}}_{e,k}^H) - \mathrm{Tr}((\beta-1)\mathbf{Z}^\star\hat{\mathbf{h}}_{e,k}\hat{\mathbf{h}}_{e,k}^H) \tag{42b}$$

$$= \mathrm{Tr}(\mathbf{W}^{\star\frac{1}{2}}\mathbf{P}\mathbf{W}^{\star\frac{1}{2}}\hat{\mathbf{h}}_{e,k}\hat{\mathbf{h}}_{e,k}^H) - \mathrm{Tr}(\underbrace{\mathbf{W}^{\star\frac{1}{2}}(\beta-1)\mathbf{P}^\perp\mathbf{W}^{\star\frac{1}{2}}}_{\mathbf{Y}}\hat{\mathbf{h}}_{e,k}\hat{\mathbf{h}}_{e,k}^H) - \mathrm{Tr}((\beta-1)\mathbf{Z}^\star\hat{\mathbf{h}}_{e,k}\hat{\mathbf{h}}_{e,k}^H) \tag{42c}$$

$$\leq \mathrm{Tr}(\mathbf{W}^{\star\frac{1}{2}}\mathbf{P}\mathbf{W}^{\star\frac{1}{2}}\hat{\mathbf{h}}_{e,k}\hat{\mathbf{h}}_{e,k}^H) - \mathrm{Tr}((\beta-1)\mathbf{Z}^\star\hat{\mathbf{h}}_{e,k}\hat{\mathbf{h}}_{e,k}^H) \tag{42d}$$

$$= \hat{\mathbf{h}}_{e,k}^H\mathbf{W}^{\star\frac{1}{2}}\frac{\mathbf{W}^{\star\frac{1}{2}}\hat{\mathbf{h}}_b\hat{\mathbf{h}}_b^H\mathbf{W}^{\star\frac{1}{2}}}{\left\|\mathbf{W}^{\star\frac{1}{2}}\hat{\mathbf{h}}_b\right\|^2}\mathbf{W}^{\star\frac{1}{2}}\hat{\mathbf{h}}_{e,k} - \mathrm{Tr}((\beta-1)\mathbf{Z}^\star\hat{\mathbf{h}}_{e,k}\hat{\mathbf{h}}_{e,k}^H) \tag{42e}$$

$$= \hat{\mathbf{h}}_b^H\mathbf{W}^{\star\frac{1}{2}}\frac{\mathbf{W}^{\star\frac{1}{2}}\hat{\mathbf{h}}_{e,k}\hat{\mathbf{h}}_{e,k}^H\mathbf{W}^{\star\frac{1}{2}}}{\left\|\mathbf{W}^{\star\frac{1}{2}}\hat{\mathbf{h}}_b\right\|^2}\mathbf{W}^{\star\frac{1}{2}}\hat{\mathbf{h}}_b - \mathrm{Tr}((\beta-1)\mathbf{Z}^\star\hat{\mathbf{h}}_{e,k}\hat{\mathbf{h}}_{e,k}^H) \tag{42f}$$

$$\leq \lambda_{\max}(\mathbf{W}^{\star\frac{1}{2}}\hat{\mathbf{h}}_{e,k}\hat{\mathbf{h}}_{e,k}^H\mathbf{W}^{\star\frac{1}{2}}) - \mathrm{Tr}((\beta-1)\mathbf{Z}^\star\hat{\mathbf{h}}_{e,k}\hat{\mathbf{h}}_{e,k}^H) \tag{42g}$$

$$\Leftrightarrow \leq \hat{\mathbf{h}}_{e,k}^H\mathbf{W}^\star\hat{\mathbf{h}}_{e,k} - \mathrm{Tr}((\beta-1)\mathbf{Z}^\star\hat{\mathbf{h}}_{e,k}\hat{\mathbf{h}}_{e,k}^H) \tag{42h}$$

$$\Leftrightarrow \leq \hat{\mathbf{h}}_{e,k}^H[\mathbf{W}^\star - (\beta-1)\mathbf{Z}^\star]\hat{\mathbf{h}}_{e,k}, \tag{42i}$$

where the second term of (42c) is a nonnegative value due to $\mathbf{Y} \succeq \mathbf{0}$ and $\beta \geq 1$. Based on (42i), we have

$$\hat{\mathbf{h}}_{e,k}^H[\tilde{\mathbf{W}}^\star - (\beta-1)\tilde{\mathbf{Z}}^\star]\hat{\mathbf{h}}_{e,k} \leq \hat{\mathbf{h}}_{e,k}^H[\mathbf{W}^\star - (\beta-1)\mathbf{Z}^\star]\hat{\mathbf{h}}_{e,k} \leq (\beta-1)\sigma_{e,k}^2, \tag{43}$$

which implies that the constructed solution $(\tilde{\mathbf{W}}^\star, \tilde{\mathbf{Z}}^\star)$ satisfies constraint (12b) and then satisfies constraints (25b), (25c) and (25d).

Fourthly, it is obvious that the constraints (10d) and (10e) are satisfied with the constructed solution $(\tilde{\mathbf{W}}^\star, \tilde{\mathbf{Z}}^\star)$ since $\tilde{\mathbf{W}}^\star \succeq \mathbf{0}$ and $\tilde{\mathbf{Z}}^\star \succeq \mathbf{0}$.

Therefore, we conclude that $(\tilde{\mathbf{W}}^\star, \tilde{\mathbf{Z}}^\star)$ is also an optimal solution of Problem (25) with rank one. The proof is completed.

Appendix B
Proof of Proposition 1

We denote the OF of Problem (28) by $f(\mathbf{E}, \boldsymbol{\delta}, \mathbf{x}, \mathbf{y})$ and the optimal solution of Problem (28) by $(\mathbf{E}^\star, \boldsymbol{\delta}^\star, \mathbf{x}^\star, \mathbf{y}^\star)$, then we have $f(\mathbf{E}, \boldsymbol{\delta}, \mathbf{x}, \mathbf{y}) \leq f(\mathbf{E}^\star, \boldsymbol{\delta}^\star, \mathbf{x}^\star, \mathbf{y}^\star)$ for all the variables $\mathbf{E}, \boldsymbol{\delta}, \mathbf{x}, \mathbf{y}$ that satisfy the constraints (27b), (27c), (27d), (27e), (27f), (28c), (28d) and $\log(1+M) - \log\det(\mathbf{I}+\mathbf{E}) = 0$ (i.e., $\text{rank}(\mathbf{E}) = 1$). Assume that $(\mathbf{E}_s, \boldsymbol{\delta}_s, \mathbf{x}_s, \mathbf{y}_s)$ is the optimal solution of Problem (31) for penalty factor κ_s, we have $g(\mathbf{E}_s, \boldsymbol{\delta}_s, \mathbf{x}_s, \mathbf{y}_s; \kappa_s) \geq g(\mathbf{E}^\star, \boldsymbol{\delta}^\star, \mathbf{x}^\star, \mathbf{y}^\star; \kappa_s)$, which leads to the inequality

$$\begin{aligned} &f(\mathbf{E}_s, \boldsymbol{\delta}_s, \mathbf{x}_s, \mathbf{y}_s) + \frac{1}{\kappa_s}[\log(1+M) - \log\det(\mathbf{I}+\mathbf{E}_s)] \\ &\geq f(\mathbf{E}^\star, \boldsymbol{\delta}^\star, \mathbf{x}^\star, \mathbf{y}^\star) + \frac{1}{\kappa_s}[\log(1+M) - \log\det(\mathbf{I}+\mathbf{E}^\star)] \qquad (44a) \\ &= f(\mathbf{E}^\star, \boldsymbol{\delta}^\star, \mathbf{x}^\star, \mathbf{y}^\star), \qquad (44b) \end{aligned}$$

where the (44b) holds as $(\mathbf{E}^\star, \boldsymbol{\delta}^\star, \mathbf{x}^\star, \mathbf{y}^\star)$ is the optimal solution of Problem (28) and therefore the rank-one constraint is satisfied. The inequality in (44) can be rewritten as

$$[\log(1+M) - \log\det(\mathbf{I}+\mathbf{E}_s)] \geq \kappa_s[f(\mathbf{E}^\star, \boldsymbol{\delta}^\star, \mathbf{x}^\star, \mathbf{y}^\star) - f(\mathbf{E}_s, \boldsymbol{\delta}_s, \mathbf{x}_s, \mathbf{y}_s)]. \tag{45}$$

Assume that $\{\bar{\mathbf{E}}, \bar{\boldsymbol{\delta}}, \bar{\mathbf{x}}, \bar{\mathbf{y}}\}$ is a limit point of sequence $\{\mathbf{E}_s, \boldsymbol{\delta}_s, \mathbf{x}_s, \mathbf{y}_s; \kappa_s\}$ for $\kappa_s \to 0$, i.e., $\lim\limits_{\kappa_s \to 0} \mathbf{E}_s = \bar{\mathbf{E}}$, $\lim\limits_{\kappa_s \to 0} \boldsymbol{\delta}_s = \bar{\boldsymbol{\delta}}$, $\lim\limits_{\kappa_s \to 0} \mathbf{x}_s = \bar{\mathbf{x}}$ and $\lim\limits_{\kappa_s \to 0} \mathbf{y}_s = \bar{\mathbf{y}}$. By taking the limit operation $\kappa_s \to 0$ on both sides of (45), we have

$$\begin{aligned} &[\log(1+M) - \log\det(\mathbf{I}+\bar{\mathbf{E}})] = \lim_{\kappa_s \to 0}[\log(1+M) - \log\det(\mathbf{I}+\mathbf{E}_s)] \qquad (46a) \\ &\geq \lim_{\kappa_s \to 0} \kappa_s[f(\mathbf{E}^\star, \boldsymbol{\delta}^\star, \mathbf{x}^\star, \mathbf{y}^\star) - f(\mathbf{E}_s, \boldsymbol{\delta}_s, \mathbf{x}_s, \mathbf{y}_s)] = 0, \qquad (46b) \end{aligned}$$

where the (46a) holds due to the continuity of the function $\log(1+M)-\log\det(\mathbf{I}+\mathbf{E})$. According to (46) and Lemma 2, we have $[\log(1+M)-\log\det(\mathbf{I}+\bar{\mathbf{E}})]=0$, and $(\bar{\mathbf{E}},\bar{\boldsymbol{\delta}},\bar{\mathbf{x}},\bar{\mathbf{y}})$ are feasible for Problem (28).

Furthermore, by taking the limit operation $\kappa_s \to 0$ on both sides of (44), and utilizing the non-negativity of κ_s, we have

$$f(\bar{\mathbf{E}},\bar{\boldsymbol{\delta}},\bar{\mathbf{x}},\bar{\mathbf{y}})+\lim_{\kappa_s\to 0}\frac{1}{\kappa_s}[\log(1+M)-\log\det(\mathbf{I}+\mathbf{E}_s)] \geq f(\mathbf{E}^\star,\boldsymbol{\delta}^\star,\mathbf{x}^\star,\mathbf{y}^\star) \tag{47a}$$

$$\Leftrightarrow f(\bar{\mathbf{E}},\bar{\boldsymbol{\delta}},\bar{\mathbf{x}},\bar{\mathbf{y}}) \geq f(\mathbf{E}^\star,\boldsymbol{\delta}^\star,\mathbf{x}^\star,\mathbf{y}^\star), \tag{47b}$$

where the (47b) holds because of $\log(1+M)-\log\det(\mathbf{I}+\mathbf{E}_s)\leq 0$ according to Lemma 2. It is seen from (47b) that $(\bar{\mathbf{E}},\bar{\boldsymbol{\delta}},\bar{\mathbf{x}},\bar{\mathbf{y}})$ is a feasible point for Problem (28) whose OF value is no smaller than that of the optimal solution $(\mathbf{E}^\star,\boldsymbol{\delta}^\star,\mathbf{x}^\star,\mathbf{y}^\star)$. Therefore the limit solution $(\bar{\mathbf{E}},\bar{\boldsymbol{\delta}},\bar{\mathbf{x}},\bar{\mathbf{y}})$ is also an optimal solution for Problem (28).

REFERENCES

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

[2] J. Zhao and Y. Liu, "A survey of intelligent reflecting surfaces (IRSs): Towards 6G wireless communication networks." [Online]. Available: https://arxiv.org/abs/1907.04789

[3] T. J. Cui, M. Q. Qi, X. Wan, J. Zhao, and Q. Cheng, "Coding metamaterials, digital metamaterials and programmable metamaterials," *Light: Science & Applications*, vol. 3, no. 10, pp. 218–218, 2014.

[4] M. Di Renzo, K. Ntontin, J. Song, F. H. Danufane, X. Qian, F. Lazarakis, J. De Rosny, D.-T. Phan-Huy, O. Simeone, R. Zhang *et al.*, "Reconfigurable intelligent surfaces vs. relaying: Differences, similarities, and performance comparison," *IEEE Open Journal of the Communications Society*, vol. 1, pp. 798–807, 2020.

[5] X. Yu, D. Xu, and R. Schober, "MISO wireless communication systems via intelligent reflecting surfaces," in *2019 IEEE/CIC International Conference on Communications in China (ICCC)*, 2019, pp. 735–740.

[6] Y. Yang, B. Zheng, S. Zhang, and R. Zhang, "Intelligent reflecting surface meets OFDM: Protocol design and rate maximization," *IEEE Transactions on Communications*, vol. early access, pp. 1–1, 2020.

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

[8] H. Guo, Y. Liang, J. Chen, and E. G. Larsson, "Weighted sum-rate maximization for intelligent reflecting surface enhanced wireless networks," in *2019 IEEE Global Communications Conference (GLOBECOM)*, 2019, pp. 1–6.

[9] G. Zhou, C. Pan, H. Ren, K. Wang, and A. Nallanathan, "Intelligent reflecting surface aided multigroup multicast MISO communication systems," *IEEE Transactions on Signal Processing*, vol. 68, pp. 3236–3251, 2020.

[10] C. Pan, H. Ren, K. Wang, W. Xu, M. Elkashlan, A. Nallanathan, and L. Hanzo, "Multicell MIMO communications relying on intelligent reflecting surfaces," *IEEE Transactions on Wireless Communications*, vol. 19, no. 8, pp. 5218–5233, 2020.

[11] C. Pan, H. Ren, K. Wang, M. Elkashlan, A. Nallanathan, J. Wang, and L. Hanzo, "Intelligent reflecting surface aided MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1719–1734, 2020.

[12] T. Bai, C. Pan, Y. Deng, M. Elkashlan, A. Nallanathan, and L. Hanzo, "Latency minimization for intelligent reflecting surface aided mobile edge computing," *IEEE Journal on Selected Areas in Communications*, vol. early access, pp. 1–1, 2020.

[13] D. Xu, X. Yu, Y. Sun, D. W. K. Ng, and R. Schober, "Resource allocation for IRS-assisted full-duplex cognitive radio systems," *IEEE Transactions on Communications*, vol. early access, pp. 1–1, 2020.

[14] Y.-S. Shiu, S. Y. Chang, H.-C. Wu, S. C.-H. Huang, and H.-H. Chen, "Physical layer security in wireless networks: A tutorial," *IEEE wireless Communications*, vol. 18, no. 2, pp. 66–74, 2011.

[15] X. Chen, D. W. K. Ng, W. H. Gerstacker, and H.-H. Chen, "A survey on multiple-antenna techniques for physical layer security," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 1027–1053, 2016.

[16] J. Li, A. P. Petropulu, and S. Weber, "On cooperative relaying schemes for wireless physical layer security," *IEEE Transactions on signal processing*, vol. 59, no. 10, pp. 4985–4997, 2011.

[17] Y. Sun, D. W. K. Ng, J. Zhu, and R. Schober, "Robust and secure resource allocation for full-duplex MISO multicarrier NOMA systems," *IEEE Transactions on Communications*, vol. 66, no. 9, pp. 4119–4137, 2018.

[18] L. Dong, Z. Han, A. P. Petropulu, and H. V. Poor, "Improving wireless physical layer security via cooperating relays," *IEEE Transactions on signal processing*, vol. 58, no. 3, pp. 1875–1888, 2009.

[19] X. Yu, D. Xu, and R. Schober, "Enabling secure wireless communications via intelligent reflecting surfaces." [Online]. Available: https://arxiv.org/abs/1904.09573

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

[21] Z. Chu, W. Hao, P. Xiao, and J. Shi, "Intelligent reflecting surface aided multi-antenna secure transmission," *IEEE Wireless Communications Letters*, vol. 9, no. 1, pp. 108–112, 2020.

[22] H. Shen, W. Xu, S. Gong, Z. He, and C. Zhao, "Secrecy rate maximization for intelligent reflecting surface assisted multi-antenna communications," *IEEE Communications Letters*, vol. 23, no. 9, pp. 1488–1492, 2019.

[23] J. Chen, Y.-C. Liang, Y. Pei, and H. Guo, "Intelligent reflecting surface: A programmable wireless environment for physical layer security," *IEEE Access*, vol. 7, pp. 82 599–82 612, 2019.

[24] H. Yang, Z. Xiong, J. Zhao, D. Niyato, L. Xiao, and Q. Wu, "Deep reinforcement learning based intelligent reflecting surface for secure wireless communications," *IEEE Transactions on Wireless Communications*, vol. early access, pp. 1–1, 2020.

[25] X. Guan, Q. Wu, and R. Zhang, "Intelligent reflecting surface assisted secrecy communication: Is artificial noise helpful or not?" *IEEE Wireless Communications Letters*, vol. 9, no. 6, pp. 778–782, 2020.

[26] D. Xu, X. Yu, Y. Sun, D. W. K. Ng, and R. Schober, "Resource allocation for secure IRS-assisted multiuser MISO systems," in *2019 IEEE Globecom Workshops (GC Wkshps)*. IEEE, 2019, pp. 1–6.

[27] S. Hong, C. Pan, H. Ren, K. Wang, and A. Nallanathan, "Artificial-noise-aided secure mimo wireless communications via intelligent reflecting surface," *IEEE Transactions on Communications*, vol. early access, pp. 1–1, 2020.

[28] A. Taha, M. Alrabeiah, and A. Alkhateeb, "Enabling large intelligent surfaces with compressive sensing and deep learning." [Online]. Available: https://arxiv.org/abs/1904.10136

[29] S. Zhang and R. Zhang, "Capacity characterization for intelligent reflecting surface aided MIMO communication," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1823–1838, 2020.

[30] Z. Zhou, N. Ge, Z. Wang, and L. Hanzo, "Joint transmit precoding and reconfigurable intelligent surface phase adjustment: A decomposition-aided channel estimation approach." [Online]. Available: https://www.researchgate.net/publication/337824343

[31] Z. Wang, L. Liu, and S. Cui, "Channel estimation for intelligent reflecting surface assisted multiuser communications: Framework, algorithms, and analysis," *IEEE Transactions on Wireless Communications*, vol. early access, pp. 1–1, 2020.

[32] P. Wang, J. Fang, H. Duan, and H. Li, "Compressed channel estimation for intelligent reflecting surface-assisted millimeter wave systems," *IEEE Signal Processing Letters*, vol. 27, pp. 905–909, 2020.

[33] J. Chen, Y.-C. Liang, H. V. Cheng, and W. Yu, "Channel estimation for reconfigurable intelligent surface aided multi-user MIMO systems." [Online]. Available: https://arxiv.org/abs/1912.03619

[34] G. Zhou, C. Pan, H. Ren, K. Wang, M. D. Renzo, and A. Nallanathan, "Robust beamforming design for intelligent reflecting surface aided MISO communication systems," *IEEE Wireless Communications Letters*, vol. 9, no. 10, pp. 1658–1662, 2020.

[35] X. Yu, D. Xu, Y. Sun, D. W. K. Ng, and R. Schober, "Robust and secure wireless communications via intelligent reflecting surfaces," *IEEE Journal on Selected Areas in Communications*, vol. early access, pp. 1–1, 2020.

[36] G. Zhou, C. Pan, H. Ren, K. Wang, and A. Nallanathan, "A framework of robust transmission design for IRS-aided MISO communications with imperfect cascaded channels," *IEEE Transactions on Signal Processing*, vol. 68, pp. 5092–5106, 2020.

[37] K.-Y. Wang, A. M.-C. So, T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "Outage constrained robust transmit optimization for multiuser MISO downlinks: Tractable approximations by conic optimization," *IEEE Transactions on Signal Processing*, vol. 62, no. 21, pp. 5690–5705, 2014.

[38] Z. Chu, H. Xing, M. Johnston, and S. Le Goff, "Secrecy rate optimizations for a MISO secrecy channel with multiple multiantenna eavesdroppers," *IEEE Transactions on Wireless Communications*, vol. 15, no. 1, pp. 283–297, 2015.

[39] S. Ma, M. Hong, E. Song, X. Wang, and D. Sun, "Outage constrained robust secure transmission for miso wiretap channels," *IEEE Transactions on Wireless Communications*, vol. 13, no. 10, pp. 5558–5570, 2014.

[40] Y. Yuan and Z. Ding, "Outage constrained secrecy rate maximization design with swipt in mimo-cr systems," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 6, pp. 5475–5480, 2017.

[41] Q. Li and W.-K. Ma, "Optimal and robust transmit designs for MISO channel secrecy by semidefinite programming," *IEEE Transactions on Signal Processing*, vol. 59, no. 8, pp. 3799–3812, 2011.

[42] J. Nocedal and S. Wright, *Numerical optimization*. Springer Science & Business Media, 2006.

[43] T. Jiang and Y. Shi, "Over-the-air computation via intelligent reflecting surfaces," in *2019 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2019, pp. 1–6.

[44] A. Ben-Tal and A. Nemirovski, *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*. SIAM, 2001.

[45] W.-C. Liao, T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "QoS-based transmit beamforming in the presence of eavesdroppers: An optimized artificial-noise-aided approach," *IEEE Transactions on Signal Processing*, vol. 59, no. 3, pp. 1202–1216, 2010.