

Secrecy Rate Optimization for Intelligent Reflecting Surface Assisted MIMO System

Zheng Chu, Wanming Hao, Pei Xiao, De Mi, Zilong Liu, Mohsen Khalily, James R. Kelly, and Alexandros P. Feresidis

Abstract—This paper investigates the impact of intelligent reflecting surface (IRS) enabled wireless secure transmission. Specifically, an IRS is deployed to assist multiple-input multiple-output (MIMO) secure system to enhance the secrecy performance, and artificial noise (AN) is employed to introduce interference to degrade the reception of the eavesdropper. To improve the secrecy performance, we aim to maximize the achievable secrecy rate, subject to the transmit power constraint, by jointly designing the precoding of the secure transmission, the AN jamming, and the reflecting phase shift of the IRS. We first propose an alternative optimization algorithm (i.e., block coordinate descent (BCD) algorithm) to tackle the non-convexity of the formulated problem. This is made by deriving the transmit precoding and AN matrices via the Lagrange dual method and the phase shifts by the Majorization-Minimization (MM) algorithm. Our analysis reveals that the proposed BCD algorithm converges in a monotonically non-decreasing manner which leads to guaranteed optimal solution. Finally, we provide numerical results to validate the secrecy performance enhancement of the proposed scheme in comparison to the benchmark schemes.

Index Terms—Intelligent reflecting surface, physical-layer security, multiple-input multiple-output (MIMO), phase shift.

I. INTRODUCTION

The fifth-generation (5G) communication networks and beyond have been evolving towards machine-centric driven by a vast range of requirements such as massive connectivity, ultra-low latency, and ultra-high spectral efficiency and throughput. In particular, ultra-high data rates can be supported by a variety of contemporary techniques, e.g., massive multiple-input multiple-output (massive MIMO), relaying, millimeter wave (mmWave) communications, as well as ultra-dense networks (UDNs) [1]–[3]. Although these techniques can substantially boost the spectral efficiency, one needs to deal with

This work was supported by the U.K. Engineering and Physical Sciences Research Council under Grants EP/P008402/2 and EP/P008380/1. The work of W. Hao was supported by Scientific and Technological Key Project of Henan Province under Grant 202102210119. The authors also would like to acknowledge the support of the University of Surrey 5GIC (<http://www.surrey.ac.uk/5gic>) members for this work.

Z. Chu, P. Xiao, and D. Mi, M. Khalily are with the Institute for Communication Systems, University of Surrey, Guildford GU2 7XH, UK. (Email: andrew.chuzheng7@gmail.com, p.xiao@surrey.ac.uk, d.mi@surrey.ac.uk, m.khalily@surrey.ac.uk)

W. Hao is with the School of Information Engineering, Zhengzhou University, Zhengzhou, 450001, China. (Email: wmhao@hotmail.com)

Z. Liu is with the School of Computer Science and Electronic Engineering (CSEE), University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, UK. (Email: z.liu@essex.ac.uk)

J. Kelly is with the School of Electronic Engineering and Computer Science, Queen Mary University of London (QMUL), London E1 4NS, UK. (Email: j.kelly@qmul.ac.uk)

A. Feresidis is with the Department of Electronic, Electrical and Systems Engineering, University of Birmingham, Birmingham B15 2TT, UK. (Email: a.feresidis@bham.ac.uk)

the increase of computational complexity and hardware cost due to large amount of radio frequency (RF) chains over a high frequency band. In addition, the operation of these techniques often come at a cost of high power consumption due to the increased amount of transmitted data and emission of radio waves [4]. This motivates a novel and promising paradigm, called *smart radio environment*, which has attracted tremendous research attention in recent years. *Smart radio environment* offers a seamless wireless connectivity and the capability of processing and transmitting data via recycling the existing radio waves instead of generating new ones [4].

On the other hand, for various civilian and military applications, it has become increasingly important for secured wireless networks to safeguard the private/important information such as credit card transaction, on-line personal data, and military intelligent transmissions [5]. As a matter of fact, wireless security has become an indispensable part of 5G wireless networks. Conventionally, a reliable wireless transmission is secured via traditional cryptographic techniques operating in the network layer. However, this tends to incur large overhead as well as various challenges in terms of key distribution and management to build reliable link due to the nature of wireless transmission [6]. As an alternative approach, physical layer security has been developed to provide the secrecy capacity metric by exploiting information-theoretical fundamentals [5]. In recent years, a variety of resource allocation algorithms have been developed in physical layer security scenarios to improve the secrecy capacity. Additionally, physical layer security has also been studied in multi-antenna scenario [7]–[9].

Intelligent reflecting surface (IRS), as an enabler for *smart radio environments*, has been proposed as a novel transceiver-like technique to offer significant performance enhancement in terms of spectrum and energy efficiencies [10], [11]. An IRS is generally consisted of a planar array structure with a large number of reconfigurable passive reflecting elements, which are controlled by a communication-oriented software (e.g. IRS controller) [4], [10]. IRS unlocks a novel wireless device (WD) to achieve three-dimensional (3D) passive beamforming gains by adaptively varying the phase shifts of the reflected signals in time-varying environments [12]. IRS reflecting arrays are generally made of multiple components with small-sized, very low-cost, and low-energy consumption features in which an appropriate phase shift can efficiently improve the reflection of the intended signal without a dedicated RF signal processing, en/de-coding, or re-transmission [10].

A. State-of-the-Art

There has been a rich body of literature on the IRS assisted wireless communication networks [10], [12]–[16]. In [10], an IRS assisted multiple-input single-output (MISO) system was proposed, where the desired signals are passively received and reflected by the reflecting elements of the IRS to align the reflection link with the direct link. In [13] a class of active intelligent surface based massive MIMO was investigated by employing an array architecture to passively receive and reflect the desired signals to maximize the total received power targeting to a specific user. Also, multiple IRSs may be flexibly deployed to provide high data rates in the near-field transmission such as indoor environments [14]. In [15], [16], a series of resource allocation algorithms have been developed to jointly design transmit precoding and phase shift matrices in various MIMO scenarios, by maximizing the achievable rate via dynamically adjusting the phase shift at the IRS. From the perspective of security, IRS-aided secrecy system has recently attracted increasing attentions [17]–[20]. IRS, deployed near the legitimate users, dynamically adjusts the phase shifts to optimize the secrecy performance. In [17], [18], an IRS-aided multi-antenna secrecy system has been investigated, where the achievable secrecy rate is maximized by joint optimization of the secrecy transmit beamforming and reflecting phase shift. The power efficiency optimization problem has been formulated to meet the secrecy rate requirement in [19]. The application of artificial noise (AN) was studied to enhance the achievable secrecy rate in an secure IRS assisted system [20]. Specifically, the integration of AN with IRS can achieve a higher secrecy performance improvement compared with the benchmark schemes. This is due to the fact that an IRS assisted secure system is typically lack of sufficient spatial secure degrees of freedom (DoF) with increasing number of eavesdroppers, such that AN can be an effective means for secrecy performance improvement. In addition, we are interested to compare the IRS aided secure system to other existing secure counterparts, e.g., information-jamming-assisted secure system [21] and the relaying-assisted secure system [22]. Although these conventional transmission systems can enhance the secrecy performance, extra power consumption is needed for jamming and relaying techniques which are considered as external transceivers. The IRS assisted secure system does not need to employ an external transmitter (i.e., information jammer or relay) to generate the new or same radio wave to enhance (degrade) the received power at the legitimate user (eavesdropper) [19]. The existing state-of-the-art only focuses on single-antenna cases at the receive nodes (i.e., legitimate users and eavesdroppers). The formulated problems can be generally relaxed into the semi-definite programming (SDP) to obtain the sub-optimal transmit beamforming and the phase shift via the alternative optimization algorithm. Similar method however may not be applied in IRS aided MIMO secrecy scenario. Besides, there has been few of existing work which exploits the IRS aided MIMO secrecy system, which motivates our work in this paper.

This paper unveils the impact of the IRS on an AN aided MIMO secure system, where the secure transmit precoding,

AN jamming as well as the reflecting phase shift matrix of the IRS are jointly designed to enhance secrecy performance. *The main contributions of this paper are summarized as follows:*

- 1) We find the optimal joint design of secure transmit precoding matrix, AN jamming matrix and IRS phase shift matrix, to maximize the achievable secrecy rate, subject to transmit power and unit modulus constraints. To deal with its non-convexity, we first transform the original problem into its equivalent form via weighted minimum mean square error (WMMSE). Then the problem is further split into three sub-problems by introducing auxiliary variable matrices. We propose a block coordinate decent (BCD) algorithm to alternatively optimize the auxiliary matrices, the transmit precoding, the AN jamming matrix, and the phase shift matrix. In addition, the convergence behaviour of the proposed BCD algorithm is characterized and shown to converge to the Karush-Kuhn-Tucker (KKT) point of the original problem.
- 2) We solve the sub-problems with respect to auxiliary variable matrices for the given transmit precoding, AN matrix and phase shift matrix, where their closed-form expressions are derived by taking into consideration the first-order derivatives of these sub-problems.
- 3) We then solve the corresponding sub-problem over transmit precoding with AN matrix via the Lagrange dual method, which is derived with the semi-closed-form expressions for given phase shift matrix and auxiliary variable matrices. Also, we perform a bisection search to find the optimal dual variable.
- 4) Moreover, the main novelty of our work lies in the optimal design of the phase shift matrix at the IRS. To be specific, for the given transmit precoding, AN matrix and auxiliary variable matrices, we transform the sub-problem with respect to phase shift matrix into a quadratically constrained quadratic programming (QCQP) with unit-modulus constraint via several matrix/vector manipulations. Due to the non-convex nature of the QCQP problem, we propose a novel method based on the Majorization-Minimization (MM) algorithm [23] to iteratively find the optimal phase shift of the IRS with a closed-form expression. In addition, we prove the proposed MM algorithm converges to the KKT point.
- 5) Finally, numerical evaluations validate the benefits of the proposed algorithm in comparison to the benchmark schemes, with an emphasis on the impact of the IRS on secrecy performance enhancement in the AN aided MIMO secure system. We show that the higher transmit power and larger number of reflecting elements of the IRS will lead to a higher achievable secrecy rate. Additionally, we reveal that the number of antenna at each node and IRS-related link plays an important role in enhancing the secrecy performance.

The rest of this paper is organized as follows. The secrecy system model is described in Section II, followed by the algorithm design for the IRS aided MIMO secrecy communication with AN in Section III. The IRS aided MIMO secrecy communication without AN is described in Section

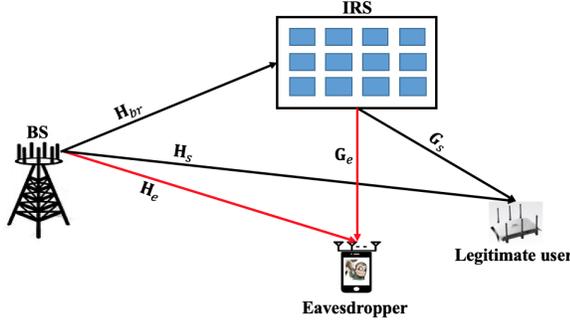


Fig. 1: An IRS aided MIMO secrecy communication.

IV. Numerical results are presented in Section V. Finally, we conclude this paper in Section VI.

Notations: We use upper case boldface letters for matrices and lower case boldface letters for vectors. $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^H$ denote the conjugate, transpose and conjugate transpose, respectively. $\text{Tr}(\cdot)$ is the trace of a matrix. $\nabla_{\mathbf{x}}f(\mathbf{x})$ represents the gradient of the function f with respect to the vector \mathbf{x} and \odot is the Hadamard product. $\mathbf{A} \succeq 0$ indicates that \mathbf{A} is a positive semi-definite matrix. \mathbf{I} and $(\cdot)^{-1}$ denote the identity matrix and the inverse of a matrix, respectively. $\text{diag}(\mathbf{A})$ is the diagonalization operation of the matrix \mathbf{A} . $|\mathbf{A}|$ denotes the determinant of \mathbf{A} and $[x]^+$ represents $\max\{x, 0\}$.

II. SYSTEM MODEL

In this paper, we consider a generic IRS assisted MIMO secrecy system as shown in Fig. 1, where a BS intends to establish a reliable communication link with a legitimate user in the presence of single eavesdropper. Also, an IRS is deployed to passively reflect the confidential information from the BS to the receivers (i.e., legitimate user/eavesdropper), each of which generally probes two path superimposed signals from both BS-user/eavesdropper and BS-IRS-user/eavesdropper links. In comparison to the traditional secrecy communication, in our paper, an AN embedded signal is generated by the BS to introduce additional interference to degrade the reception of the eavesdropper by exploiting the IRS-induced extra DoF with the necessity of the AN aided transmit beamformer. Thus, the intended signal to be transmitted at the BS can be written as

$$\mathbf{x} = \mathbf{W}\mathbf{s} + \mathbf{z}, \quad (1)$$

where $\mathbf{s} \sim \mathcal{CN}(0, \mathbf{I})$ denotes the desired signal to be transmitted, $\mathbf{W} \in \mathbb{C}^{N_T \times N_d}$ is the transmit precoding matrix, $N_d \leq \min(N_T, N_S)$ denotes the data streams observed for the user, $\mathbf{z} \sim \mathcal{CN}(0, \mathbf{Z})$ is the AN matrix, i.e., $\mathbf{Z} \in \mathbb{C}^{N_T \times N_T}$. We assume that the BS is equipped with $N_T \geq 1$ transmit antennas, the IRS with $N_R \geq 1$ reflecting units, the user with N_S , and the eavesdropper consists of N_E receive antennas. We denote $\mathbf{H}_{br} \in \mathbb{C}^{N_R \times N_T}$, $\mathbf{H}_s \in \mathbb{C}^{N_T \times N_S}$, $\mathbf{H}_{e,k} \in \mathbb{C}^{N_T \times N_E}$, $\mathbf{G}_s \in \mathbb{C}^{N_R \times N_S}$, and $\mathbf{G}_E \in \mathbb{C}^{N_R \times N_E}$ as the channel coefficients between the BS and the IRS, the BS and the user, the BS and the eavesdropper, the IRS and the user, as well as the IRS and the eavesdropper, respectively. In this paper, we assume that the channel state information (CSI) of the eavesdropper is available at the BS. This can be achieved

through different methods such as the CSI feedback method [24] or even the local oscillator power leakage from the eavesdropper receivers RF frontend [25]. The IRS elements collect all multi-path signals at a physical element, and reflect these combined signals via IRS reflecting array. We denote $\Theta = \text{diag}(\exp(j\alpha_1), \exp(j\alpha_2), \dots, \exp(j\alpha_{N_R}))$, as the diagonal matrix associated with the effective phase shifts in all IRS elements, where $\alpha_n \in [0, 2\pi]$, $\forall n \in [1, N_R]$ is the phase shift at the n -th IRS element. Thus, the received signal at the legitimate user and the eavesdropper can be written as

$$y_s = (\mathbf{H}_s^H + \mathbf{G}_s^H \Theta \mathbf{H}_{br}) (\mathbf{W}\mathbf{s} + \mathbf{z}) + \mathbf{n}_s,$$

$$y_e = (\mathbf{H}_e^H + \mathbf{G}_e^H \Theta \mathbf{H}_{br}) (\mathbf{W}\mathbf{s} + \mathbf{z}) + \mathbf{n}_e,$$

respectively, where $\mathbf{n}_s \sim \mathcal{CN}(0, \sigma_s^2 \mathbf{I})$ and $\mathbf{n}_e \sim \mathcal{CN}(0, \sigma_e^2 \mathbf{I})$. The achievable rates at the legitimate user and the eavesdropper are given by [26], [27]

$$R_s = \log \left| \mathbf{I} + \tilde{\mathbf{H}}_s \mathbf{W} \mathbf{W}^H \tilde{\mathbf{H}}_s^H \left(\mathbf{I} + \tilde{\mathbf{H}}_s \mathbf{Z} \tilde{\mathbf{H}}_s^H \right)^{-1} \right|,$$

$$R_e = \log \left| \mathbf{I} + \tilde{\mathbf{H}}_e \mathbf{W} \mathbf{W}^H \tilde{\mathbf{H}}_e^H \left(\mathbf{I} + \tilde{\mathbf{H}}_e \mathbf{Z} \tilde{\mathbf{H}}_e^H \right)^{-1} \right|,$$

where $\tilde{\mathbf{H}}_s = \frac{1}{\sigma_s} \mathbf{H}_s$, $\tilde{\mathbf{H}}_e = \frac{1}{\sigma_e} \mathbf{H}_e$, $\tilde{\mathbf{H}}_s = \mathbf{H}_s^H + \mathbf{G}_s^H \Theta \mathbf{H}_{br}$, and $\tilde{\mathbf{H}}_e = \mathbf{H}_e^H + \mathbf{G}_e^H \Theta \mathbf{H}_{br}$. Thus, the achievable secrecy rate can be written as

$$R_{sec} = [R_s - R_e]^+, \quad (2)$$

III. IRS AIDED MIMO SECRECY COMMUNICATION WITH AN

A. Problem Formulation

In order to evaluate the secrecy performance of the IRS assisted MIMO secrecy network, we aim to maximize the achievable secrecy rate subject to the transmit power constraint, where the transmit precoding, the AN matrix and the reflecting beamforming are jointly designed. Thus, this optimization problem is formulated as follows:

$$\max_{\mathbf{W}, \mathbf{Z} \succeq 0, \Theta} R_s - R_e,$$

$$s.t. \text{Tr}(\mathbf{W} \mathbf{W}^H + \mathbf{Z}) \leq P, |\exp(j\alpha_n)| = 1, \forall n = 1, \dots, N_R. \quad (3)$$

By defining $\mathbf{Z} = \mathbf{V} \mathbf{V}^H$, $\mathbf{V} \in \mathbb{C}^{N_T \times N_T}$, problem (3) is equivalently expressed as

$$\max_{\mathbf{W}, \mathbf{V}, \Theta} \underbrace{\log \left| \mathbf{I} + \tilde{\mathbf{H}}_s \mathbf{W} \mathbf{W}^H \tilde{\mathbf{H}}_s^H \left(\mathbf{I} + \tilde{\mathbf{H}}_s \mathbf{V} \mathbf{V}^H \tilde{\mathbf{H}}_s^H \right)^{-1} \right|}_{A_1} + \underbrace{\log \left| \mathbf{I} + \tilde{\mathbf{H}}_e \mathbf{V} \mathbf{V}^H \tilde{\mathbf{H}}_e^H \right|}_{A_2} - \underbrace{\log \left| \mathbf{I} + \tilde{\mathbf{H}}_e \mathbf{W} \mathbf{W}^H \tilde{\mathbf{H}}_e^H + \tilde{\mathbf{H}}_e \mathbf{V} \mathbf{V}^H \tilde{\mathbf{H}}_e^H \right|}_{A_3},$$

$$s.t. \text{Tr}(\mathbf{W} \mathbf{W}^H + \mathbf{V} \mathbf{V}^H) \leq P,$$

$$|\exp(j\alpha_n)| = 1, \forall n = 1, \dots, N_R. \quad (4)$$

Problem (4) is non-convex and intractable. In order to circumvent this non-convex property, we first employ the idea of WMMSE, transforming the objective function in (4) into the equivalent counterpart, which can be designed iteratively via the BCD method [28]. To proceed, we introduce the auxiliary matrices $(\mathbf{X}_i, (i \in \{1, 2, 3\}), \mathbf{Z}_j, (j \in \{1, 2\}))$ to reformulate

A_1 , A_2 , and A_3 in the objective function in problem (4), respectively. First, let us consider the MSE matrix function of A_1 as follows:

$$\mathbf{E}_1(\mathbf{Z}_1, \mathbf{W}, \mathbf{V}) = (\mathbf{I} - \mathbf{Z}_1^H \tilde{\mathbf{H}}_s \mathbf{W})(\mathbf{I} - \mathbf{Z}_1^H \tilde{\mathbf{H}}_s \mathbf{W})^H + \mathbf{Z}_1^H (\mathbf{I} + \tilde{\mathbf{H}}_s \mathbf{V} \mathbf{V}^H \tilde{\mathbf{H}}_s^H) \mathbf{Z}_1, \quad (5)$$

thus, A_1 is equivalently modified as

$$A_1 = \max_{\mathbf{X}_1 \succeq \mathbf{0}, \mathbf{Z}_1} \log |\mathbf{X}_1| - \text{Tr} [\mathbf{X}_1 \mathbf{E}_1(\mathbf{Z}_1, \mathbf{W}, \mathbf{V})] + N_d. \quad (6)$$

Similarly, A_2 is given by

$$A_2 = \max_{\mathbf{X}_2 \succeq \mathbf{0}, \mathbf{Z}_2} \log |\mathbf{X}_2| - \text{Tr} [\mathbf{X}_2 \mathbf{E}_2(\mathbf{Z}_2, \mathbf{V})] + N_T, \quad (7)$$

where $\mathbf{E}_2(\mathbf{Z}_2, \mathbf{V}) = (\mathbf{I} - \mathbf{Z}_2^H \tilde{\mathbf{H}}_e \mathbf{V})(\mathbf{I} - \mathbf{Z}_2^H \tilde{\mathbf{H}}_e \mathbf{V})^H + \mathbf{Z}_2^H \mathbf{Z}_2$. The following lemma is required to deal with A_3 ,

Lemma 1: [26] Denoting $\mathbf{E} \in \mathbb{C}^{N \times N}$ as any positive definite matrix, we have the following function

$$-\log |\mathbf{E}| = \max_{\mathbf{Y} \succeq \mathbf{0}} \delta(\mathbf{Y}), \quad (8)$$

where $\delta(\mathbf{Y}) = \log |\mathbf{Y}| - \text{Tr}(\mathbf{Y}\mathbf{E}) + N$. Then, the optimal solution to problem (8) can be expressed as $\mathbf{Y}^{\text{opt}} = \mathbf{E}^{-1}$.

According to Lemma 1, A_3 is given by

$$-A_3 = \max_{\mathbf{X}_3 \succeq \mathbf{0}} \log |\mathbf{X}_3| - \text{Tr} [\mathbf{X}_3 \mathbf{E}_3(\mathbf{W}, \mathbf{V})] + N_E, \quad (9)$$

where $\mathbf{E}_3(\mathbf{W}, \mathbf{V}) = \mathbf{I} + \tilde{\mathbf{H}}_e \mathbf{W} \mathbf{W}^H \tilde{\mathbf{H}}_e^H + \tilde{\mathbf{H}}_e \mathbf{V} \mathbf{V}^H \tilde{\mathbf{H}}_e^H$.

Remark 1: It can be verified that A_1 , A_2 , and A_3 are concave functions with respect to each matrix of \mathbf{W} , \mathbf{V} , $\mathbf{X}_i (i \in \{1, 2, 3\})$, and $\mathbf{Z}_j (j \in \{1, 2\})$ given the other matrices. Also, combining A_1 , A_2 and A_3 provides the lower bound of the achievable secrecy rate R_{sec} in (2).

We substitute (6)-(9) into (4), which is equivalently rewritten as

$$\begin{aligned} & \max_{\Omega} \log |\mathbf{X}_1| - \text{Tr} \left[\mathbf{X}_1 \left((\mathbf{I} - \mathbf{Z}_1^H \tilde{\mathbf{H}}_s \mathbf{W})(\mathbf{I} - \mathbf{Z}_1^H \tilde{\mathbf{H}}_s \mathbf{W})^H \right. \right. \\ & \quad \left. \left. + \mathbf{Z}_1^H (\mathbf{I} + \tilde{\mathbf{H}}_s \mathbf{V} \mathbf{V}^H \tilde{\mathbf{H}}_s^H) \mathbf{Z}_1 \right) \right] + \log |\mathbf{X}_2| \\ & - \text{Tr} \left[\mathbf{X}_2 \left((\mathbf{I} - \mathbf{Z}_2^H \tilde{\mathbf{H}}_e \mathbf{V})(\mathbf{I} - \mathbf{Z}_2^H \tilde{\mathbf{H}}_e \mathbf{V})^H + \mathbf{Z}_2^H \mathbf{Z}_2 \right) \right] \\ & + \log |\mathbf{X}_3| - \text{Tr} \left[\mathbf{X}_3 (\mathbf{I} + \tilde{\mathbf{H}}_e \mathbf{W} \mathbf{W}^H \tilde{\mathbf{H}}_e^H + \tilde{\mathbf{H}}_e \mathbf{V} \mathbf{V}^H \tilde{\mathbf{H}}_e^H) \right], \\ & \text{s.t. } \text{Tr}(\mathbf{W} \mathbf{W}^H + \mathbf{V} \mathbf{V}^H) \leq P, \\ & |\exp(j\alpha_n)| = 1, \forall n = 1, \dots, N_R, \\ & \Omega = \{\mathbf{W}, \mathbf{V}, \mathbf{X}_1 \succeq \mathbf{0}, \mathbf{X}_2 \succeq \mathbf{0}, \mathbf{X}_3 \succeq \mathbf{0}, \mathbf{Z}_1, \mathbf{Z}_2, \Theta\}. \quad (10) \end{aligned}$$

Problem (10) is still non-convex with respect to the transmit precoding matrices (\mathbf{W} , \mathbf{V}), the auxiliary matrices ($\mathbf{X}_i, (i \in \{1, 2, 3\})$, $\mathbf{Z}_j, (j \in \{1, 2\})$), as well as the phase shift matrix Θ . In the sequel, we propose the BCD algorithm to iteratively update these variable matrices in problem (10). Specifically, problem (10) is separated into the three sub-problems (or sub-iterations), each of which aims to obtain the optimal one (or a group) of variable matrices given others, respectively.

B. Sub-iteration 1: Optimizing \mathbf{Z}_j and \mathbf{X}_i

In this subsection, we first solve (10) to optimize $\mathbf{Z}_j, (j \in \{1, 2\})$ and $\mathbf{X}_i, (i \in \{1, 2, 3\})$ given \mathbf{W} , \mathbf{V} , and Θ . First, (10) is equivalently modified into the following sub-problems over $\mathbf{Z}_j, (j \in \{1, 2\})$ as

$$\mathbf{Z}_1 = \arg \min_{\mathbf{Z}_1} \text{Tr} [\mathbf{X}_1 \mathbf{E}_1(\mathbf{Z}_1, \mathbf{W}, \mathbf{V})], \quad (11a)$$

$$\mathbf{Z}_2 = \arg \min_{\mathbf{Z}_2} \text{Tr} [\mathbf{X}_2 \mathbf{E}_2(\mathbf{Z}_2, \mathbf{W}, \mathbf{V})]. \quad (11b)$$

In order to solve (11a) and (11b), we take into consideration their own first-order derivative, respectively, and the closed-form solution of $\mathbf{Z}_j (j \in \{1, 2\})$ is given by

$$\mathbf{Z}_1 = \left(\mathbf{I} + \tilde{\mathbf{H}}_s \mathbf{V} \mathbf{V}^H \tilde{\mathbf{H}}_s^H + \tilde{\mathbf{H}}_s \mathbf{W} \mathbf{W}^H \tilde{\mathbf{H}}_s^H \right)^{-1} \tilde{\mathbf{H}}_s \mathbf{W}, \quad (12a)$$

$$\mathbf{Z}_2 = \left(\mathbf{I} + \tilde{\mathbf{H}}_e \mathbf{V} \mathbf{V}^H \tilde{\mathbf{H}}_e^H \right)^{-1} \tilde{\mathbf{H}}_e \mathbf{V}. \quad (12b)$$

Next, we solve (10) to optimize $\mathbf{X}_i, (i \in \{1, 2, 3\})$ given \mathbf{W} , \mathbf{V} , $\mathbf{Z}_j, (j \in \{1, 2\})$, and Θ . It is observed that the matrices $\mathbf{X}_i, (i \in \{1, 2, 3\})$ are independent with each other in the objective function of (10). Thus, by exploiting Lemma 1, the closed-form solutions of $\mathbf{X}_i, (i \in \{1, 2, 3\})$ is derived as

$$\begin{aligned} \mathbf{X}_1 &= \left[(\mathbf{I} - \mathbf{Z}_1^H \tilde{\mathbf{H}}_s \mathbf{W})(\mathbf{I} - \mathbf{Z}_1^H \tilde{\mathbf{H}}_s \mathbf{W})^H + \mathbf{Z}_1^H (\mathbf{I} \right. \\ & \quad \left. + \tilde{\mathbf{H}}_s \mathbf{V} \mathbf{V}^H \tilde{\mathbf{H}}_s^H) \mathbf{Z}_1 \right]^{-1} \\ & \stackrel{a}{=} \mathbf{I} + \mathbf{W}^H \tilde{\mathbf{H}}_s^H (\mathbf{I} + \tilde{\mathbf{H}}_s \mathbf{V} \mathbf{V}^H \tilde{\mathbf{H}}_s^H)^{-1} \tilde{\mathbf{H}}_s \mathbf{W}, \quad (13a) \end{aligned}$$

$$\begin{aligned} \mathbf{X}_2 &= \left[(\mathbf{I} - \mathbf{Z}_2^H \tilde{\mathbf{H}}_e \mathbf{V})(\mathbf{I} - \mathbf{Z}_2^H \tilde{\mathbf{H}}_e \mathbf{V})^H + \mathbf{Z}_2^H \mathbf{Z}_2 \right]^{-1} \\ & \stackrel{b}{=} \mathbf{I} + \mathbf{V}^H \tilde{\mathbf{H}}_e^H \tilde{\mathbf{H}}_e \mathbf{V}, \quad (13b) \end{aligned}$$

$$\mathbf{X}_3 = \left(\mathbf{I} + \tilde{\mathbf{H}}_e \mathbf{W} \mathbf{W}^H \tilde{\mathbf{H}}_e^H + \tilde{\mathbf{H}}_e \mathbf{V} \mathbf{V}^H \tilde{\mathbf{H}}_e^H \right)^{-1}, \quad (13c)$$

where $\stackrel{a}{=}$ and $\stackrel{b}{=}$ denotes that (12a) and (12b) substitute into (13a) and (13b), respectively.

C. Sub-iteration 2: Optimizing \mathbf{W} and \mathbf{V}

In this subsection, we solve problem (10) to optimally design \mathbf{W} and \mathbf{V} given $\mathbf{X}_i, (i \in \{1, 2, 3\})$, $\mathbf{Z}_j (j \in \{1, 2\})$, and Θ . To proceed, problem (10) is equivalently rewritten with respect to \mathbf{W} and \mathbf{V} , as

$$\begin{aligned} & \min_{\mathbf{W} \succeq \mathbf{0}, \mathbf{V} \succeq \mathbf{0}} \text{Tr} \left(\mathbf{X}_1 \mathbf{Z}_1^H \tilde{\mathbf{H}}_s \mathbf{W} \mathbf{W}^H \tilde{\mathbf{H}}_s^H \mathbf{Z}_1 \right) \\ & + \text{Tr} \left(\mathbf{X}_1 \mathbf{Z}_1^H \tilde{\mathbf{H}}_s \mathbf{V} \mathbf{V}^H \tilde{\mathbf{H}}_s^H \mathbf{Z}_1 \right) - \text{Tr} \left(\mathbf{X}_1 \mathbf{Z}_1^H \tilde{\mathbf{H}}_s \mathbf{W} \right) \\ & - \text{Tr} \left(\mathbf{X}_1 \mathbf{W}^H \tilde{\mathbf{H}}_s \mathbf{Z}_1 \right) - \text{Tr} \left(\mathbf{X}_2 \mathbf{Z}_2^H \tilde{\mathbf{H}}_e \mathbf{V} \right) \\ & - \text{Tr} \left(\mathbf{X}_2 \mathbf{V}^H \tilde{\mathbf{H}}_e^H \mathbf{Z}_2 \right) + \text{Tr} \left(\mathbf{X}_2 \mathbf{Z}_2^H \tilde{\mathbf{H}}_e \mathbf{V} \mathbf{V}^H \tilde{\mathbf{H}}_e^H \mathbf{Z}_2 \right) \\ & + \text{Tr} \left(\mathbf{X}_3 \tilde{\mathbf{H}}_e \mathbf{W} \mathbf{W}^H \tilde{\mathbf{H}}_e^H \right) + \text{Tr} \left(\mathbf{X}_3 \tilde{\mathbf{H}}_e \mathbf{V} \mathbf{V}^H \tilde{\mathbf{H}}_e^H \right) \\ & \text{s.t. } \text{Tr}(\mathbf{W} \mathbf{W}^H) + \text{Tr}(\mathbf{V} \mathbf{V}^H) \leq P. \quad (14) \end{aligned}$$

Problem (14) is a convex problem with respect to \mathbf{W} and \mathbf{V} , which can be solved via Lagrange dual problem. Let us write the Lagrangian function to problem (14) as follows:

$$\begin{aligned} \mathcal{L}(\mathbf{W}, \mathbf{V}, \mu) &= \text{Tr} \left(\mathbf{X}_1 \mathbf{Z}_1^H \tilde{\mathbf{H}}_s \mathbf{W} \mathbf{W}^H \tilde{\mathbf{H}}_s^H \mathbf{Z}_1 \right) \\ & + \text{Tr} \left(\mathbf{X}_1 \mathbf{Z}_1^H \tilde{\mathbf{H}}_s \mathbf{V} \mathbf{V}^H \tilde{\mathbf{H}}_s^H \mathbf{Z}_1 \right) - \text{Tr} \left(\mathbf{X}_1 \mathbf{Z}_1^H \tilde{\mathbf{H}}_s \mathbf{W} \right) \\ & - \text{Tr} \left(\mathbf{X}_1 \mathbf{W}^H \tilde{\mathbf{H}}_s \mathbf{Z}_1 \right) - \text{Tr} \left(\mathbf{X}_2 \mathbf{Z}_2^H \tilde{\mathbf{H}}_e \mathbf{V} \right) \\ & - \text{Tr} \left(\mathbf{X}_2 \mathbf{V}^H \tilde{\mathbf{H}}_e^H \mathbf{Z}_2 \right) + \text{Tr} \left(\mathbf{X}_2 \mathbf{Z}_2^H \tilde{\mathbf{H}}_e \mathbf{V} \mathbf{V}^H \tilde{\mathbf{H}}_e^H \mathbf{Z}_2 \right) \\ & + \text{Tr} \left(\mathbf{X}_3 \tilde{\mathbf{H}}_e \mathbf{W} \mathbf{W}^H \tilde{\mathbf{H}}_e^H \right) + \text{Tr} \left(\mathbf{X}_3 \tilde{\mathbf{H}}_e \mathbf{V} \mathbf{V}^H \tilde{\mathbf{H}}_e^H \right) \\ & + \mu [\text{Tr}(\mathbf{W} \mathbf{W}^H) + \text{Tr}(\mathbf{V} \mathbf{V}^H) - P], \quad (15) \end{aligned}$$

and its dual problem is given by

$$\max_{\mu \geq 0} \min_{\mathbf{W} \succeq \mathbf{0}, \mathbf{V} \succeq \mathbf{0}} \mathcal{L}(\mathbf{W}, \mathbf{V}, \mu). \quad (16)$$

In order to solve the dual problem (16), we consider a two-level algorithm. Particularly, the inner-level problem is

considered to optimize \mathbf{W} and \mathbf{V} for given $\mu > 0$, and the outer level problem is a single-variable optimization problem with respect to μ , which is obtained via one-dimensional search (i.e., bisection method). To proceed, we first solve the inner-level problem to optimize \mathbf{W} and \mathbf{V} , which can be derived in terms of closed-form expression by taking into consideration the first-order derivative of (15) as

$$\mathbf{W} = \left(\mu \mathbf{I} + \tilde{\mathbf{H}}_s^H \mathbf{Z}_1 \mathbf{X}_1 \mathbf{Z}_1^H \tilde{\mathbf{H}}_s + \tilde{\mathbf{H}}_e^H \mathbf{X}_3 \tilde{\mathbf{H}}_e \right)^{-1} \tilde{\mathbf{H}}_s^H \mathbf{Z}_1 \mathbf{X}_1, \quad (17)$$

$$\mathbf{V} = \left(\mu \mathbf{I} + \tilde{\mathbf{H}}_s^H \mathbf{Z}_1 \mathbf{X}_1 \mathbf{Z}_1^H \tilde{\mathbf{H}}_s + \tilde{\mathbf{H}}_e^H \mathbf{Z}_2 \mathbf{X}_2 \mathbf{Z}_2^H \tilde{\mathbf{H}}_e + \tilde{\mathbf{H}}_e^H \mathbf{X}_3 \tilde{\mathbf{H}}_e \right)^{-1} \tilde{\mathbf{H}}_e^H \mathbf{Z}_2 \mathbf{X}_2. \quad (18)$$

We further simplify the solutions of \mathbf{W} and \mathbf{V} via eigen-decomposition. Specifically, let $\Sigma = \tilde{\mathbf{H}}_s^H \mathbf{Z}_1 \mathbf{X}_1 \mathbf{Z}_1^H \tilde{\mathbf{H}}_s + \tilde{\mathbf{H}}_e^H \mathbf{X}_3 \tilde{\mathbf{H}}_e$ and $\Delta = \tilde{\mathbf{H}}_s^H \mathbf{Z}_1 \mathbf{X}_1 \mathbf{Z}_1^H \tilde{\mathbf{H}}_s + \tilde{\mathbf{H}}_e^H \mathbf{Z}_2 \mathbf{X}_2 \mathbf{Z}_2^H \tilde{\mathbf{H}}_e + \tilde{\mathbf{H}}_e^H \mathbf{X}_3 \tilde{\mathbf{H}}_e$, we have the following eigen-decomposition:

$$\Sigma = \mathbf{R} \Lambda \mathbf{R}^H, \quad \Delta = \mathbf{S} \Xi \mathbf{S}^H, \quad (19)$$

where \mathbf{R} (or \mathbf{S}) denotes a unitary matrix consisting of the orthonormal eigenvectors of Σ (or Δ) and Λ (or Ξ) is a diagonal matrix, each diagonal element of which is the eigenvalue of Σ (or Δ). Thus, for given $\mu > 0$, the optimal solution of \mathbf{W} and \mathbf{V} can be re-expressed as

$$\mathbf{W} = \mathbf{R} (\mu \mathbf{I} + \Lambda)^{-1} \mathbf{R}^H \tilde{\mathbf{H}}_s^H \mathbf{Z}_1 \mathbf{X}_1, \quad (20a)$$

$$\mathbf{V} = \mathbf{S} (\mu \mathbf{I} + \Xi)^{-1} \mathbf{S}^H \tilde{\mathbf{H}}_e^H \mathbf{Z}_2 \mathbf{X}_2. \quad (20b)$$

Then, we solve the outer-level problem to achieve the optimal solution of dual variable μ . With (20a) and (20b), the optimal solution to problem (14) can be obtained by searching the optimal dual variable μ such that the complementary slackness condition holds

$$\mu [\text{Tr}(\mathbf{W}\mathbf{W}^H) + \text{Tr}(\mathbf{V}\mathbf{V}^H) - P] = 0. \quad (21)$$

The optimal solution to dual variable μ can be obtained by performing a bisection search, as it can be shown that the term $f(\mathbf{W}(\mu), \mathbf{V}(\mu)) = \text{Tr}(\mathbf{W}\mathbf{W}^H) + \text{Tr}(\mathbf{V}\mathbf{V}^H)$ is a monotonically non-increasing function with respect to μ [26]. The detailed procedure to solve problem (14) is summarized as *Algorithm 1*. It has been proven in [27] that *Algorithm 1* monotonically converges to a KKT point of problem (14).

Algorithm 1: Bisection search to solve problem (14)

- 1) **Initialization:** κ is a small value to denote the algorithm accuracy; **lower** bound μ_{low} and **upper** bound μ_{up} .
 - 2) **Calculate** $\mu = \mu_{\text{low}} + \mu_{\text{up}}$.
 - 3) **Obtain** the optimal solution of \mathbf{W} and \mathbf{V} .
 - 4) **If** $\text{Tr}(\mathbf{W}\mathbf{W}^H + \mathbf{V}\mathbf{V}^H) \geq P$, $\mu_{\text{low}} = \mu$, **otherwise**, $\mu_{\text{up}} = \mu$.
 - 5) **If** $|\mu_{\text{up}} - \mu_{\text{low}}| \leq \kappa$, **terminate**, **otherwise**, go to **Step 2**.
 - 6) **Output:** \mathbf{W} and \mathbf{V} .
-

D. Sub-iteration 3: Optimizing Θ

In this subsection, we solve problem (10) to optimize Θ for given \mathbf{W} , \mathbf{V} , $\mathbf{X}_i (i \in \{1, 2, 3\})$, and $\mathbf{Z}_j (j \in \{1, 2\})$. We first

equivalently rewrite problem (10) as follows:

$$\begin{aligned} \min_{\Theta} & \underbrace{\text{Tr}(\mathbf{X}_1 \mathbf{Z}_1^H \tilde{\mathbf{H}}_s \mathbf{W} \mathbf{W}^H \tilde{\mathbf{H}}_s^H \mathbf{Z}_1)}_{B_1} + \underbrace{\text{Tr}(\mathbf{X}_1 \mathbf{Z}_1^H \tilde{\mathbf{H}}_s \mathbf{V} \mathbf{V}^H \tilde{\mathbf{H}}_s^H \mathbf{Z}_1)}_{B_2} \\ & - \left[\underbrace{\text{Tr}(\mathbf{X}_1 \mathbf{Z}_1^H \tilde{\mathbf{H}}_s \mathbf{W})}_{B_2} + \underbrace{\text{Tr}(\mathbf{X}_1 \mathbf{W}^H \tilde{\mathbf{H}}_s \mathbf{Z}_1)}_{B_1} \right] \\ & + \underbrace{\text{Tr}(\mathbf{X}_2 \mathbf{Z}_2^H \tilde{\mathbf{H}}_e \mathbf{V} \mathbf{V}^H \tilde{\mathbf{H}}_e^H \mathbf{Z}_2)}_{B_3} \\ & - \left[\underbrace{\text{Tr}(\mathbf{X}_2 \mathbf{Z}_2^H \tilde{\mathbf{H}}_e \mathbf{V})}_{B_3} + \underbrace{\text{Tr}(\mathbf{X}_2 \mathbf{V}^H \tilde{\mathbf{H}}_e^H \mathbf{Z}_2)}_{B_4} \right] \\ & + \underbrace{\text{Tr}(\mathbf{X}_3 \tilde{\mathbf{H}}_e \mathbf{W} \mathbf{W}^H \tilde{\mathbf{H}}_e^H)}_{B_5} + \underbrace{\text{Tr}(\mathbf{X}_3 \tilde{\mathbf{H}}_e \mathbf{V} \mathbf{V}^H \tilde{\mathbf{H}}_e^H)}_{B_5}, \end{aligned} \quad (22)$$

$$s.t. \quad |\exp(j\alpha_n)| = 1, \forall n = 1, \dots, N_R, \quad (22)$$

where $\tilde{\mathbf{H}}_s = \frac{1}{\sigma_s} \tilde{\mathbf{H}}_s$, $\tilde{\mathbf{H}}_e = \frac{1}{\sigma_e} \tilde{\mathbf{H}}_e$, $\tilde{\mathbf{H}}_s = \mathbf{H}_s^H + \mathbf{G}_s^H \Theta \mathbf{H}_{br}$, and $\tilde{\mathbf{H}}_e = \mathbf{H}_e^H + \mathbf{G}_e^H \Theta \mathbf{H}_{br}$. Problem (22) is a non-convex problem and intractable with respect to Θ due to its unit modulus constraint, which cannot be solved directly. In order to tackle this non-convexity, we propose an MM algorithm to approximately derive the optimal solution of phase shift. To facilitate the MM algorithm, we first take into consideration the equivalent transformation of the objective function in (22). Let $\gamma_s = \frac{1}{\sigma_s^2}$, $\gamma_e = \frac{1}{\sigma_e^2}$, $\bar{\mathbf{W}} = \mathbf{W}\mathbf{W}^H$, and $\bar{\mathbf{V}} = \mathbf{V}\mathbf{V}^H$, B_1 , B_2 , B_3 , B_4 , and B_5 in problem (22) can be transformed via exploiting a series of mathematical manipulations, respectively,

• B_1 :

$$\begin{aligned} B_1 & = \gamma_s \text{Tr}(\Theta^H \mathbf{G}_s \mathbf{Z}_1 \mathbf{X}_1 \mathbf{Z}_1^H \mathbf{G}_s^H \Theta \mathbf{H}_{br} \bar{\mathbf{W}} \mathbf{H}_{br}^H) \\ & + \gamma_s \text{Tr}(\Theta^H \mathbf{G}_s \mathbf{Z}_1 \mathbf{X}_1 \mathbf{Z}_1^H \mathbf{H}_s^H \bar{\mathbf{W}} \mathbf{H}_s^H \Theta) \\ & + \gamma_s \text{Tr}(\mathbf{H}_{br} \bar{\mathbf{W}} \mathbf{H}_s \mathbf{Z}_1 \mathbf{X}_1 \mathbf{Z}_1^H \mathbf{G}_s^H \Theta) \\ & + \gamma_s \text{Tr}(\mathbf{X}_1 \mathbf{Z}_1^H \mathbf{H}_s^H \bar{\mathbf{W}} \mathbf{H}_s \mathbf{Z}_1) \\ & + \gamma_s \text{Tr}(\Theta^H \mathbf{G}_s \mathbf{Z}_1 \mathbf{X}_1 \mathbf{Z}_1^H \mathbf{G}_s^H \Theta \mathbf{H}_{br} \bar{\mathbf{V}} \mathbf{H}_{br}^H) \\ & + \gamma_s \text{Tr}(\Theta^H \mathbf{G}_s \mathbf{Z}_1 \mathbf{X}_1 \mathbf{Z}_1^H \mathbf{H}_s^H \bar{\mathbf{V}} \mathbf{H}_s^H \Theta) \\ & + \gamma_s \text{Tr}(\mathbf{H}_{br} \bar{\mathbf{V}} \mathbf{H}_s \mathbf{Z}_1 \mathbf{X}_1 \mathbf{Z}_1^H \mathbf{G}_s^H \Theta) \\ & + \gamma_s \text{Tr}(\mathbf{X}_1 \mathbf{Z}_1^H \mathbf{H}_s^H \bar{\mathbf{V}} \mathbf{H}_s \mathbf{Z}_1) \\ & = \text{Tr}(\Theta^H \mathbf{R}_s \Theta \mathbf{T}_{1,s}) + \text{Tr}(\Theta^H \mathbf{Q}_{1,s}^H) \\ & + \text{Tr}(\Theta \mathbf{Q}_{1,s}) + \text{Tr}(\Theta^H \mathbf{R}_s \Theta \mathbf{T}_{2,s}) \\ & + \text{Tr}(\Theta^H \mathbf{Q}_{2,s}^H) + \text{Tr}(\Theta \mathbf{Q}_{2,s}) + c_s, \end{aligned} \quad (23)$$

where

$$\begin{aligned} \mathbf{R}_s & = \gamma_s \mathbf{G}_s \mathbf{Z}_1 \mathbf{X}_1 \mathbf{Z}_1^H \mathbf{G}_s^H, \quad \mathbf{T}_{1,s} = \mathbf{H}_{br} \bar{\mathbf{W}} \mathbf{H}_{br}^H, \\ \mathbf{Q}_{1,s} & = \gamma_s \mathbf{H}_{br} \bar{\mathbf{W}} \mathbf{H}_s \mathbf{Z}_1 \mathbf{X}_1 \mathbf{Z}_1^H \mathbf{G}_s^H, \\ c_s & = \gamma_s \text{Tr}(\mathbf{X}_1 \mathbf{Z}_1^H \mathbf{H}_s^H \bar{\mathbf{W}} \mathbf{H}_s \mathbf{Z}_1) \\ & + \gamma_s \text{Tr}(\mathbf{X}_1 \mathbf{Z}_1^H \mathbf{H}_s^H \bar{\mathbf{V}} \mathbf{H}_s \mathbf{Z}_1), \\ \mathbf{T}_{2,s} & = \mathbf{H}_{br} \bar{\mathbf{V}} \mathbf{H}_{br}^H, \quad \mathbf{Q}_{2,s} = \gamma_s \mathbf{H}_{br} \bar{\mathbf{V}} \mathbf{H}_s \mathbf{Z}_1 \mathbf{X}_1 \mathbf{Z}_1^H \mathbf{G}_s^H. \end{aligned}$$

- B_2 :

$$\begin{aligned}
B_2 &= \sqrt{\gamma_s} \text{Tr} (\Theta \mathbf{H}_{br} \mathbf{W} \mathbf{X}_1 \mathbf{Z}_1^H \mathbf{G}_s^H) \\
&\quad + \sqrt{\gamma_s} \text{Tr} (\mathbf{X}_1 \mathbf{Z}_1^H \mathbf{H}_s^H \mathbf{W}) \\
&\quad + \sqrt{\gamma_s} \text{Tr} (\Theta^H \mathbf{G}_s \mathbf{Z}_1 \mathbf{X}_1 \mathbf{W}^H \mathbf{H}_{br}^H) \\
&\quad + \sqrt{\gamma_s} \text{Tr} (\mathbf{X}_1 \mathbf{W}^H \mathbf{H}_s \mathbf{Z}_1) \\
&= \text{Tr} (\Theta \mathbf{Q}_{0,s}) + \text{Tr} (\Theta^H \mathbf{Q}_{0,s}^H) + c_{0,s}, \quad (24)
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{Q}_{0,s} &= \sqrt{\gamma_s} \mathbf{H}_{br} \mathbf{W} \mathbf{X}_1 \mathbf{Z}_1^H \mathbf{G}_s^H, \\
c_{0,s} &= \sqrt{\gamma_s} \text{Tr} (\mathbf{X}_1 \mathbf{Z}_1^H \mathbf{H}_s^H \mathbf{W}) + \sqrt{\gamma_s} \text{Tr} (\mathbf{X}_1 \mathbf{W}^H \mathbf{H}_s \mathbf{Z}_1), \\
\mathbf{q}_{0,s} &= \left[(\mathbf{Q}_{0,s})_{(1,1)}, \dots, (\mathbf{Q}_{0,s})_{N_R, N_R} \right]^T.
\end{aligned}$$

- B_3 :

$$\begin{aligned}
B_3 &= \gamma_e \text{Tr} (\Theta^H \mathbf{G}_e \mathbf{Z}_2 \mathbf{X}_2 \mathbf{Z}_2^H \mathbf{G}_e^H \Theta \mathbf{H}_{br} \bar{\mathbf{V}} \mathbf{H}_{br}^H) \\
&\quad + \gamma_e \text{Tr} (\Theta^H \mathbf{G}_e \mathbf{Z}_2 \mathbf{X}_2 \mathbf{Z}_2^H \mathbf{H}_e^H \bar{\mathbf{V}} \mathbf{H}_{br}^H) \\
&\quad + \gamma_e \text{Tr} (\mathbf{H}_{br} \bar{\mathbf{V}} \mathbf{H}_e \mathbf{Z}_2 \mathbf{X}_2 \mathbf{Z}_2^H \mathbf{G}_e^H \Theta) \\
&\quad + \gamma_e \text{Tr} (\mathbf{X}_2 \mathbf{Z}_2^H \mathbf{H}_e^H \bar{\mathbf{V}} \mathbf{H}_e \mathbf{Z}_2) \\
&= \text{Tr} (\Theta^H \mathbf{R}_{0,e} \Theta \mathbf{T}_{0,br}) + \text{Tr} (\Theta^H \mathbf{Q}_{0,e}^H) \\
&\quad + \text{Tr} (\Theta \mathbf{Q}_{0,e}) + c_{0,e}, \quad (25)
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{R}_{0,e} &= \gamma_e \mathbf{G}_e \mathbf{Z}_2 \mathbf{X}_2 \mathbf{Z}_2^H \mathbf{G}_e^H, \quad \mathbf{T}_{0,br} = \mathbf{H}_{br} \bar{\mathbf{V}} \mathbf{H}_{br}^H, \\
\mathbf{Q}_{0,e} &= \gamma_e \mathbf{H}_{br} \bar{\mathbf{V}} \mathbf{H}_e \mathbf{Z}_2 \mathbf{X}_2 \mathbf{Z}_2^H \mathbf{G}_e^H, \\
c_{0,e} &= \gamma_e \text{Tr} (\mathbf{X}_2 \mathbf{Z}_2^H \mathbf{H}_e^H \bar{\mathbf{V}} \mathbf{H}_e \mathbf{Z}_2).
\end{aligned}$$

- B_4 :

$$\begin{aligned}
B_4 &= \sqrt{\gamma_e} \text{Tr} (\Theta \mathbf{H}_{br} \mathbf{V} \mathbf{X}_2 \mathbf{Z}_2^H \mathbf{G}_e^H) \\
&\quad + \sqrt{\gamma_e} \text{Tr} (\mathbf{X}_2 \mathbf{Z}_2^H \mathbf{H}_e^H \mathbf{V}) \\
&\quad + \sqrt{\gamma_e} \text{Tr} (\Theta^H \mathbf{G}_e \mathbf{Z}_2 \mathbf{X}_2 \mathbf{V}^H \mathbf{H}_{br}^H) \\
&\quad + \sqrt{\gamma_e} \text{Tr} (\mathbf{X}_2 \mathbf{V}^H \mathbf{H}_e \mathbf{Z}_2) \\
&= \text{Tr} (\Theta \mathbf{Q}_e) + \text{Tr} (\Theta^H \mathbf{Q}_e^H) + s_e, \quad (26)
\end{aligned}$$

where $\mathbf{Q}_e = \sqrt{\gamma_e} \mathbf{H}_{br} \mathbf{V} \mathbf{X}_2 \mathbf{Z}_2^H \mathbf{G}_e^H$, $s_e = \sqrt{\gamma_e} \text{Tr} (\mathbf{X}_2 \mathbf{Z}_2^H \mathbf{H}_e^H \mathbf{V}) + \sqrt{\gamma_e} \text{Tr} (\mathbf{X}_2 \mathbf{V}^H \mathbf{H}_e \mathbf{Z}_2)$.

- B_5 :

$$\begin{aligned}
B_5 &= \gamma_e \text{Tr} (\Theta \mathbf{G}_e \mathbf{X}_3 \mathbf{G}_e^H \Theta \mathbf{H}_{br} \bar{\mathbf{W}} \mathbf{H}_{br}^H) \\
&\quad + \gamma_e \text{Tr} (\Theta^H \mathbf{G}_e \mathbf{X}_3 \mathbf{H}_e^H \bar{\mathbf{W}} \mathbf{H}_{br}^H) \\
&\quad + \gamma_e \text{Tr} (\mathbf{H}_{br} \bar{\mathbf{W}} \mathbf{H}_e \mathbf{X}_3 \mathbf{G}_e^H \Theta) \\
&\quad + \gamma_e \text{Tr} (\mathbf{X}_3 \mathbf{H}_e^H \bar{\mathbf{W}} \mathbf{H}_e) \\
&\quad + \gamma_e \text{Tr} (\Theta \mathbf{G}_e \mathbf{X}_3 \mathbf{G}_e^H \Theta \mathbf{H}_{br} \bar{\mathbf{V}} \mathbf{H}_{br}^H) \\
&\quad + \gamma_e \text{Tr} (\Theta^H \mathbf{G}_e \mathbf{X}_3 \mathbf{H}_e^H \bar{\mathbf{V}} \mathbf{H}_{br}^H) \\
&\quad + \gamma_e \text{Tr} (\mathbf{H}_{br} \bar{\mathbf{V}} \mathbf{H}_e \mathbf{X}_3 \mathbf{G}_e^H \Theta) \\
&\quad + \gamma_e \text{Tr} (\mathbf{X}_3 \mathbf{H}_e^H \bar{\mathbf{V}} \mathbf{H}_e) \\
&= \text{Tr} (\Theta^H \mathbf{R}_e \Theta \mathbf{T}_{1,br}) + \text{Tr} (\Theta^H \mathbf{Q}_{1,e}^H) \\
&\quad + \text{Tr} (\Theta \mathbf{Q}_{1,e}) + \text{Tr} (\Theta^H \mathbf{R}_e \Theta \mathbf{T}_{2,br}) \\
&\quad + \text{Tr} (\Theta^H \mathbf{Q}_{2,e}^H) + \text{Tr} (\Theta \mathbf{Q}_{2,e}) + c_{1,e}, \quad (27)
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{R}_e &= \gamma_e \mathbf{G}_e \mathbf{X}_3 \mathbf{G}_e^H, \quad \mathbf{T}_{1,br} = \mathbf{H}_{br} \bar{\mathbf{W}} \mathbf{H}_{br}^H, \\
\mathbf{Q}_{1,e} &= \gamma_e \mathbf{H}_{br} \bar{\mathbf{W}} \mathbf{H}_e \mathbf{X}_3 \mathbf{G}_e^H, \quad \mathbf{T}_{2,br} = \mathbf{H}_{br} \bar{\mathbf{V}} \mathbf{H}_{br}^H, \\
\mathbf{Q}_{2,e} &= \gamma_e \mathbf{H}_{br} \bar{\mathbf{V}} \mathbf{H}_e \mathbf{X}_3 \mathbf{G}_e^H, \\
c_{1,e} &= \gamma_e \text{Tr} (\mathbf{X}_3 \mathbf{H}_e^H \bar{\mathbf{W}} \mathbf{H}_e) + \gamma_e \text{Tr} (\mathbf{X}_3 \mathbf{H}_e^H \bar{\mathbf{V}} \mathbf{H}_e).
\end{aligned}$$

By exploiting (23)-(27), problem (22) is equivalently mod-

ified by omitting the constant terms c_s , $c_{0,s}$, $c_{0,e}$, s_e , and $c_{1,e}$ as

$$\begin{aligned}
\min_{\Theta} \text{Tr} [\Theta^H \mathbf{R}_s \Theta \mathbf{T}_s] &+ \text{Tr} [\Theta^H \mathbf{Q}_s^H] + \text{Tr} [\Theta \mathbf{Q}_s] \\
&+ \text{Tr} (\Theta^H \mathbf{R}_{0,e} \Theta \mathbf{T}_{0,br}) + \text{Tr} [\Theta^H \mathbf{Q}_e^H] \\
&+ \text{Tr} [\Theta \mathbf{Q}_e] + \text{Tr} [\Theta^H \mathbf{R}_e \Theta \mathbf{T}_{br}], \quad (28) \\
\text{s.t. } |\exp(j\alpha_n)| &= 1, \forall n = 1, \dots, N_R.
\end{aligned}$$

where $\mathbf{T}_s = \mathbf{T}_{1,s} + \mathbf{T}_{2,s}$, $\mathbf{Q}_s = \mathbf{Q}_{1,s} + \mathbf{Q}_{2,s} - \mathbf{Q}_{0,s}$, $\mathbf{Q}_e = \mathbf{Q}_{0,e} + \mathbf{Q}_{1,e} + \mathbf{Q}_{2,e} - \mathbf{Q}_e$, and $\mathbf{T}_{br} = \mathbf{T}_{1,br} + \mathbf{T}_{2,br}$. Problem (28) is still hard to solve, thus, we resort to the following lemma to address the objective function in (28).

Lemma 2: [29] Let $\mathbf{C}_1 \in \mathbb{C}^{m \times m}$ and $\mathbf{C}_2 \in \mathbb{C}^{m \times m}$ be matrices, and $\mathbf{1} = [1, \dots, 1]^T$ be a $m \times 1$ vector. Assuming that $\mathbf{D} \in \mathbb{C}^{m \times m}$ is a diagonal matrix $\mathbf{D} = \text{diag}(d_1, \dots, d_2)$, and $\mathbf{d} = \mathbf{D}\mathbf{1}$, we have the following matrix identities:

$$\begin{aligned}
\text{Tr}(\mathbf{C}_1 \mathbf{D} \mathbf{C}_2 \mathbf{D}^H) &= \mathbf{d}^H (\mathbf{C}_1 \odot \mathbf{C}_2^T) \mathbf{d}, \\
\text{Tr}(\mathbf{D} \mathbf{C}_2) &= \mathbf{1}^T (\mathbf{D} \odot \mathbf{C}_2^T) \mathbf{1} = \mathbf{d}^T \mathbf{c}_2, \\
\text{Tr}(\mathbf{D}^H \mathbf{C}_2^H) &= \mathbf{c}_2^H \mathbf{d}^*,
\end{aligned}$$

where $\mathbf{c}_2 = [(\mathbf{C}_2)_{(1,1)}, \dots, (\mathbf{C}_2)_{(m,m)}]^T$.

By exploiting Lemma 2, (28) is equivalent to

$$\begin{aligned}
\min_{\Theta} \theta^H (\mathbf{R}_s \odot \mathbf{T}_s^T) \theta &+ \theta^H \mathbf{q}_s^* + \mathbf{q}_s^T \theta + \theta^H (\mathbf{R}_{0,e} \odot \mathbf{T}_{0,br}^T) \theta \\
&+ \theta^H \mathbf{q}_e^* + \mathbf{q}_e^T \theta + \theta^H (\mathbf{R}_e \odot \mathbf{T}_{br}^T) \theta \\
\text{s.t. } |\theta_n| &= 1, n = 1, \dots, N_R, \quad (29)
\end{aligned}$$

where $\theta = [\theta_1, \dots, \theta_{N_R}]^T$, $\theta_n = \exp(j * \alpha_n)$, and

$$\begin{aligned}
\mathbf{q}_s &= [(\mathbf{Q}_s)_{(1,1)}, \dots, (\mathbf{Q}_s)_{(N_R, N_R)}]^T, \\
\mathbf{q}_e &= [(\mathbf{Q}_e)_{(1,1)}, \dots, (\mathbf{Q}_e)_{(N_R, N_R)}]^T.
\end{aligned}$$

To proceed, some mathematical manipulations are employed to equivalently transform (29) into the following form

$$\min_{\theta} \theta^H \Phi \theta + 2\Re\{\theta^H \mathbf{q}^*\} \quad (30a)$$

$$\text{s.t. } |\theta_n| = 1, n = 1, \dots, N_R, \quad (30b)$$

where $\Phi = (\mathbf{R}_s \odot \mathbf{T}_s^T) + (\mathbf{R}_{0,e} \odot \mathbf{T}_{0,br}^T) + (\mathbf{R}_e \odot \mathbf{T}_{br}^T)$, $\mathbf{q} = \mathbf{q}_s + \mathbf{q}_e$. Problem (30) is still non-convex and intractable due to unit modulus equality constraint (30b). In order to tackle this issue, the MM algorithm is employed, and a sequence of tractable sub-problems are considered to iteratively solve problem (30) via approximating its objective function and constraint set [23]. Now, let us first take into consideration the following problem:

$$\min_{\mathbf{t}} f_0(\mathbf{t}), \quad \text{s.t. } f_i(\mathbf{t}) \leq 0, \quad i = 1, \dots, L. \quad (31)$$

We approximate both the objective function and the feasible constraint set of problem (31) at each iteration.¹ Thus, the following convex sub-problem can be solved at the m -th iteration.

$$\min_{\mathbf{x}} g_0(\mathbf{x}|\mathbf{x}^{(m)}), \quad \text{s.t. } g_i(\mathbf{x}|\mathbf{x}^{(m)}) \leq 0, \quad i = 1, \dots, L, \quad (32)$$

where $g_i(*|\mathbf{x}^{(m)})$, $\forall m = 0, \dots, L$ denotes a convex function which guarantees the following conditions:

$$\begin{aligned}
g_i(\mathbf{x}^{(m)}|\mathbf{x}^{(m)}) &= f_i(\mathbf{x}^{(m)}), \\
g_i(\mathbf{x}|\mathbf{x}^{(m)}) &\geq f_i(\mathbf{x}), \\
\nabla g_i(\mathbf{x}^{(m)}|\mathbf{x}^{(m)}) &= \nabla f_i(\mathbf{x}^{(m)}).
\end{aligned} \quad (33)$$

¹Here we assume that f_i is differential [23].

The sequence $\mathbf{x}^{(m)}$ incur a monotonically decreasing $f_0(\mathbf{x})$ which converges to a KKT point [23]. In other words, the sub-problem (32) is introduced based on the upper bound of the objective function in (31) via a convex surrogate function, and feasible set in (31) is approximated via convexifications.

Proposition 1: [30] The objective function (30a) is approximated in the following for any given $\boldsymbol{\theta}^{(m)}$ at the m -th iteration and for any feasible $\boldsymbol{\theta}$.

$$\begin{aligned} f(\boldsymbol{\theta}) &= \boldsymbol{\theta}^H \boldsymbol{\Phi} \boldsymbol{\theta} + 2\Re\{\boldsymbol{\theta}^H \mathbf{q}^*\} \\ &\leq \boldsymbol{\theta}^H \boldsymbol{\Upsilon} \boldsymbol{\theta} - 2\Re\{\boldsymbol{\theta}^H (\boldsymbol{\Upsilon} - \boldsymbol{\Phi}) \boldsymbol{\theta}^{(m)}\} \\ &\quad + \left(\boldsymbol{\theta}^{(m)}\right)^H (\boldsymbol{\Upsilon} - \boldsymbol{\Phi}) \boldsymbol{\theta}^{(m)} + 2\Re\{\boldsymbol{\theta}^H \mathbf{q}^*\} = g(\boldsymbol{\theta} | \boldsymbol{\theta}^{(m)}), \end{aligned} \quad (34)$$

where $\boldsymbol{\Upsilon} = \lambda_{\max} \mathbf{I}_{N_R}$ and λ_{\max} denotes the maximum eigenvalue of $\boldsymbol{\Phi}$.

Proposition 1 constructs a surrogate function of (30a), and it is easily verified that $g(\boldsymbol{\theta} | \boldsymbol{\theta}^{(m)})$ in (34) guarantees the conditions in (33).

Problem (30) can be approximated at the m -th iteration as
$$\min_{\boldsymbol{\theta}} g(\boldsymbol{\theta} | \boldsymbol{\theta}^{(m)}), \quad s.t. \quad (30b). \quad (35)$$

Define $\boldsymbol{\theta}^H \boldsymbol{\theta} = N_R$ and $\boldsymbol{\theta}^H \boldsymbol{\Upsilon} \boldsymbol{\theta} = N_R \lambda_{\max}$, problem (35) is equivalent to

$$\min_{\boldsymbol{\theta}} -2\Re\{\boldsymbol{\theta}^H ((\lambda_{\max} \mathbf{I}_{N_R} - \boldsymbol{\Phi}) \boldsymbol{\theta}^{(m)} - \mathbf{q}^*)\}, \quad s.t. \quad (30b). \quad (36)$$

The optimal solution of problem (36) at the $(m+1)$ -th iteration is expressed in the following closed-form,

$$\boldsymbol{\theta}^{(m+1)} = \exp\left(j \arg[(\lambda_{\max} \mathbf{I}_{N_R} - \boldsymbol{\Phi}) \boldsymbol{\theta}^{(m)} - \mathbf{q}^*]\right). \quad (37)$$

Algorithm 2: Proposed MM algorithm to solve (30)

- 1) **Initialization:** L is the maximum iteration number, δ is a small number for accuracy, $\boldsymbol{\theta}^{(0)}$ denotes feasible solution.
- 2) **Repeat:** m
 - **Obtain** $f(\boldsymbol{\theta}^{(m)})$ via (30).
 - **Obtain** $\boldsymbol{\theta}^{(m+1)}$ via (37).
 - **Obtain** $f(\boldsymbol{\theta}^{(m+1)})$.
- 3) **if** $\frac{|f(\boldsymbol{\theta}^{(m+1)}) - f(\boldsymbol{\theta}^{(m)})|}{f(\boldsymbol{\theta}^{(m+1)})} \leq \delta$ or $m > L$, **end loop**.
- 4) **Otherwise:** $m = m + 1$, **go to Step 2**.
- 5) **Output:** Calculate $\boldsymbol{\theta}^{\text{opt}}$ via (37).

On the basis of the above method to solve problem (22), we proceed to analyze the properties of *Algorithm 2* via the following *theorem*.

Theorem 1: The objective value in (30), denoted as $f(\boldsymbol{\theta})$, has a non-increasing trend and guaranteed convergence, and the converged solution generated via *Algorithm 2* is a KKT point of problem (22).

Proof: Refer to Appendix A. ■

E. Block Coordinate Descent Algorithm

In this subsection, we present the BCD algorithm to solve problem (4) on the basis of above analyses in Section III-B, III-C, and (III-D), which is summarized in *Algorithm 3*.

Algorithm 3: Proposed BCD algorithm to solve problem (4).

- 1) **Initialization:** Maximum iteration number L_1 ; Precoding $\mathbf{W}^{(0)}$ and $\mathbf{V}^{(0)}$; **Obtain** $\mathbf{X}_i^{(0)}$, ($i \in \{1, 2, 3\}$), $\mathbf{Z}_j^{(0)}$, $j \in \{1, 2\}$ via (13a)-(13c) and (12a), (12b), respectively; κ is a small value to denote algorithm accuracy.
- 2) **Repeat:**
 - a) **Given** $\mathbf{X}_i^{(n)}$ ($i \in \{1, 2, 3\}$), $\mathbf{Z}_j^{(n)}$ ($j \in \{1, 2\}$), $\boldsymbol{\Theta}^{(n)}$, **obtain** $\mathbf{W}^{(n+1)}$ and $\mathbf{V}^{(n+1)}$ via *Algorithm 1*.
 - b) **Given** $\mathbf{X}_i^{(n)}$ ($i \in \{1, 2, 3\}$), $\mathbf{Z}_j^{(n)}$ ($j \in \{1, 2\}$), $\mathbf{W}^{(n+1)}$, and $\mathbf{V}^{(n+1)}$, **obtain** $\boldsymbol{\theta}^{(n+1)}$ via *Algorithm 2*. **Calculate** $\boldsymbol{\Theta}^{(n+1)} = \text{diag}(\boldsymbol{\theta}^{(n+1)})$
 - c) **Given** $\mathbf{W}^{(n+1)}$, $\mathbf{V}^{(n+1)}$, and $\boldsymbol{\Theta}^{(n+1)}$, **obtain** $\mathbf{Z}_j^{(n+1)}$ ($j \in \{1, 2\}$) via (12a) and (12b), respectively.
 - d) **Given** $\mathbf{W}^{(n+1)}$, $\mathbf{V}^{(n+1)}$, $\boldsymbol{\Theta}^{(n+1)}$, and $\mathbf{Z}_j^{(n+1)}$ ($j \in \{1, 2\}$), **obtain** $\mathbf{X}_i^{(n+1)}$ ($i \in \{1, 2, 3\}$) via (13a), (13b), and (13c), respectively.
- 3) **If** $\frac{|R_{\text{sec}}(\mathbf{W}^{(n+1)}, \mathbf{V}^{(n+1)}, \boldsymbol{\Theta}^{(n+1)}) - R_{\text{sec}}(\mathbf{W}^{(n)}, \mathbf{V}^{(n)}, \boldsymbol{\Theta}^{(n)})|}{R_{\text{sec}}(\mathbf{W}^{(n+1)}, \mathbf{V}^{(n+1)}, \boldsymbol{\Theta}^{(n+1)})} < \kappa$ or achieve to the maximum iteration number L_1 , **terminate**; **Otherwise**, $n = n + 1$ and go to **step 2**.
- 4) **Output:** $(\mathbf{W}^{\text{opt}}, \mathbf{V}^{\text{opt}}, \mathbf{X}_i^{\text{opt}}, \mathbf{Z}_j^{\text{opt}}, \boldsymbol{\Theta}^{\text{opt}})$ and $R_{\text{sec}}(\mathbf{W}^{\text{opt}}, \mathbf{V}^{\text{opt}}, \boldsymbol{\Theta}^{\text{opt}})$.

To proceed, the following *theorem* is required to analyze the properties of the convergence and the optimal solution shown in *Algorithm 3*.

Theorem 2: The objective function value sequence $R_{\text{sec}}(\mathbf{W}^{(n)}, \mathbf{V}^{(n)}, \boldsymbol{\Theta}^{(n)})$, ($n = 1, 2, \dots, L_1$) produces a non-descending trend and can guarantee its solution convergence. Also, the converged solution of the iterative sequences generated via *Algorithm 3* is a KKT point of problem (4).

Proof: Refer to Appendix A. ■

Theorem 2 indicates that the proposed algorithm monotonically converges to a stationary point of problem (4). The monotonic convergence guarantees an improved objective value with random initialization. Moreover, the computational complexity of the proposed BCD algorithm is analyzed. In Sub-iterations 1, the computational complexity of computing $\mathbf{Z}_1, \mathbf{Z}_1, \mathbf{X}_1, \mathbf{X}_2$ and \mathbf{X}_3 is $\mathcal{O}(2N_T^3 + 3N_E^3)$. The computational complexity of *Algorithm 1* in Sub-iterations 2 is $\mathcal{O}((N_T^3 + N_E^3) \log_2(1/\kappa))$, where $\mathcal{O}(\log_2(1/\kappa))$ is the complexity of dual variable μ . The computational complexity of the MM algorithm in Sub-iteration 3 is $\mathcal{O}(N_R^3 + TN_R^3)$, where T is the iterative number. Finally, the overall computing complexity of the proposed BCD algorithm is $\mathcal{O}(L_1(2N_T^3 + 3N_E^3 + N_R^3 + TN_R^3 + (N_T^3 + N_E^3) \log_2(1/\kappa)))$.

F. Feasibility Conditions For Positive Secrecy Rate

In this subsection, we check the feasibility conditions that the optimization problem (3) can be solved when it can achieve a positive achievable secrecy rate to validate (3). Hence, we first check whether problem (3) is feasible by taking into consideration the AN jamming matrix $\mathbf{Z} \succeq \mathbf{0}$ satisfying $\mathbf{Z} = \frac{\eta P}{N_T} \mathbf{I}$ and $\text{Tr}(\mathbf{Z}) = \eta P$,² where $\eta \in [0, 1]$. To this end,

²Note that the AN power is uniformly distributed to each symbol, which has been shown to be the best solution for the BS [31].

we conduct the following power minimization problem

$$\begin{aligned} & \min_{\bar{\mathbf{W}} \succeq \mathbf{0}, \Theta} \text{Tr}(\bar{\mathbf{W}}), \\ & \text{s.t. } \log \left| \mathbf{I} + \hat{\mathbf{H}}_s \bar{\mathbf{W}} \hat{\mathbf{H}}_s^H \right| - \log \left| \mathbf{I} + \hat{\mathbf{H}}_e \bar{\mathbf{W}} \hat{\mathbf{H}}_e^H \right| \geq \bar{R}, \\ & \text{Tr}(\bar{\mathbf{W}}) \leq (1 - \eta)P, \end{aligned} \quad (38)$$

where $\hat{\mathbf{H}}_i = \mathbf{C}_i^H \bar{\mathbf{H}}_i$, $\mathbf{C}_i \mathbf{C}_i^H = (\mathbf{I} + \bar{\mathbf{H}}_i \mathbf{Z} \bar{\mathbf{H}}_i^H)^{-1}$, $\forall i \in \{s, e\}$, and $\bar{R} > 0$ denotes the target secrecy rate for the system. Provided problem (38) is feasible, the positive achievable secrecy rate can be obtained. Now, let us solve problem (38) to optimize $\bar{\mathbf{W}}$ and Θ alternatively to achieve a sub-optimal solution. To proceed, we divide (38) into the two following subproblems to optimize $\bar{\mathbf{W}}$ and Θ separately,

1) Optimize $\bar{\mathbf{W}}$ for given Θ :

$$\begin{aligned} & \min_{\bar{\mathbf{W}} \succeq \mathbf{0}} \text{Tr}(\bar{\mathbf{W}}), \\ & \text{s.t. } \log \left| \mathbf{I} + \hat{\mathbf{H}}_s \bar{\mathbf{W}} \hat{\mathbf{H}}_s^H \right| - \log \left| \mathbf{I} + \hat{\mathbf{H}}_e \bar{\mathbf{W}} \hat{\mathbf{H}}_e^H \right| \geq \bar{R}, \\ & \text{Tr}(\bar{\mathbf{W}}) \leq (1 - \eta)P. \end{aligned} \quad (39)$$

The sub-problem (39) is readily solved via the First-order Taylor approximation [6].

2) Optimize Θ for given $\bar{\mathbf{W}}$: The sub-problem can be equivalently reformulated as problem (30) [10], which can be solved via the MM algorithm shown in Section III-D.

Thus, we solve problems (39) and (30) alternatively until convergence.

IV. BENCHMARK SCHEME: IRS AIDED MIMO SECRECY COMMUNICATION WITHOUT AN

In this section, we consider a benchmark scheme, i.e., IRS aided MIMO secrecy communication without AN. In this scenario, the intended signal is expressed as $\mathbf{x} = \mathbf{W}\mathbf{s}$, and the achievable secrecy rate is given by

$$\begin{aligned} \bar{R}_{sec}(\mathbf{W}, \Theta) = & \left[\log \left| \mathbf{I} + \tilde{\mathbf{H}}_s \mathbf{W} \mathbf{W}^H \tilde{\mathbf{H}}_s^H \right| \right. \\ & \left. - \log \left| \mathbf{I} + \tilde{\mathbf{H}}_e \mathbf{W} \mathbf{W}^H \tilde{\mathbf{H}}_e^H \right| \right]^+. \end{aligned} \quad (40)$$

Thus, its secrecy rate maximization problem is written as

$$\begin{aligned} & \max_{\mathbf{W}, \Theta} \bar{R}_{sec}, \\ & \text{s.t. } \text{Tr}(\mathbf{W} \mathbf{W}^H) \leq P, |\exp(j\alpha_n)| = 1, \forall n = 1, \dots, N_R. \end{aligned} \quad (41)$$

According to the idea of WMMSE transformation and Lemma 1 shown in Section III-A, problem (41) can be equivalently modified as

$$\begin{aligned} & \max_{\mathbf{W}, \mathbf{X}_1, \mathbf{Z}_1, \Theta} \log |\mathbf{X}_1| - \text{Tr} \left[\mathbf{X}_1 \left((\mathbf{I} - \mathbf{Z}_1^H \tilde{\mathbf{H}}_s \mathbf{W}) (\mathbf{I} - \mathbf{Z}_1^H \tilde{\mathbf{H}}_s \mathbf{W})^H \right. \right. \\ & \left. \left. + \mathbf{Z}_1^H \mathbf{Z}_1 \right) \right] + \log |\mathbf{X}_2| - \text{Tr} \left[\mathbf{X}_2 (\mathbf{I} + \tilde{\mathbf{H}}_e \mathbf{W} \mathbf{W}^H \tilde{\mathbf{H}}_e^H) \right]. \end{aligned} \quad (42)$$

Similar to the mathematical manipulations shown in Section III-B-Section III-D, we can derive the optimal solutions for \mathbf{Z}_1 , $\mathbf{X}_i (i \in \{1, 2\})$, \mathbf{W} , and Θ . The BCD algorithm in Algorithm 3 is similarly considered to deal with problem (42). Specifically, at each iteration, \mathbf{W} is first updated for given $\mathbf{X}_i (i \in \{1, 2\})$, \mathbf{Z}_1 , and Θ which is obtained via solving problem (42) similar to Algorithm 1. Next, Θ is updated for given \mathbf{W} , $\mathbf{X}_i (i \in \{1, 2\})$, and \mathbf{Z}_1 via the MM algorithm similar to Algorithm 2. Then, $\mathbf{X}_i (i \in \{1, 2\})$ and \mathbf{Z}_1 are updated, respectively, in terms of closed-form expressions for

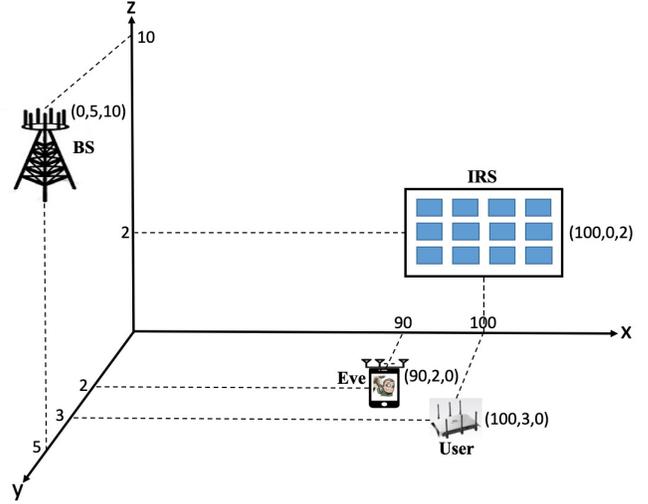


Fig. 2: System deployment

given \mathbf{W} and Θ . It can be shown that the objective value sequence generated by the BCD algorithm employed to solve problem (42) produces a monotonically non-decreasing trend and its solution is guaranteed to converge to satisfy the KKT conditions of problem (41).

V. SIMULATION RESULTS

In this section, the simulation results are presented to demonstrate the performance of the proposed algorithm. The system deployment is shown in Fig. 2, where the 3-D coordinates of the BS, the IRS, the legitimate user, and the eavesdropper are (0, 5, 10), (100, 0, 2), (100, 3, 0), and (90, 2, 0) in meter (m), respectively. It is assumed that the BS is equipped with $N_T = 5$ transmit antennas, the number of the reflecting elements at the IRS is $N_R = 30$, and the legitimate user and the eavesdropper consist of $N_S = 4$ and $N_E = 3$ receive antennas, respectively, unless otherwise specified. All small-scale channel coefficients are generated as the identical and circularly complex Gaussian random variable with zero mean and unit variance. Moreover, the large-scale path loss depends upon the distance between any two nodes, which is expressed as $\text{PL} = A \left(\frac{d}{d_0} \right)^{-\beta}$, where $A = -30$ dB denotes the path loss at the reference distance $d_0 = 1$ m, and d denotes the distance of the communication link between the BS and the IRS (i.e., d_{BS-IRS}), the BS and the legitimate user (i.e., $d_{BS-user}$), the BS and the eavesdropper (i.e., d_{BS-eve}), the IRS and the legitimate user (i.e., $d_{IRS-user}$), and the IRS and the eavesdropper (i.e., $d_{IRS-eve}$), all of which can be calculated based on the coordinates shown in Fig. 2. In addition, the path loss exponents of the BS-IRS link, the BS-user link, the BS-eve link, the IRS-user link and the IRS-eve link are set to $\beta_{BS-IRS} = 2.5$, $\beta_{BS-user} = \beta_{BS-eve} = 4$, and $\beta_{IRS-user} = \beta_{IRS-eve} = 2$, respectively, unless otherwise specified. The noise power are set to be $\sigma_s^2 = \sigma_e^2 = -105$ dBm, unless otherwise specified.

First, we evaluate the convergence properties of the BCD algorithm with different transmit powers P and different reflecting elements N_R in Fig. 3 and Fig. 4, respectively. From these two figures, it is observed that the achievable secrecy

rate has an non-decreasing trend with iteration, and the BCD algorithm achieves the convergence less than 10 iterations in most cases, which validates *Theorem 2*. In addition, larger transmit power or larger size reflecting elements plays a positive role to increase the achievable secrecy rate.

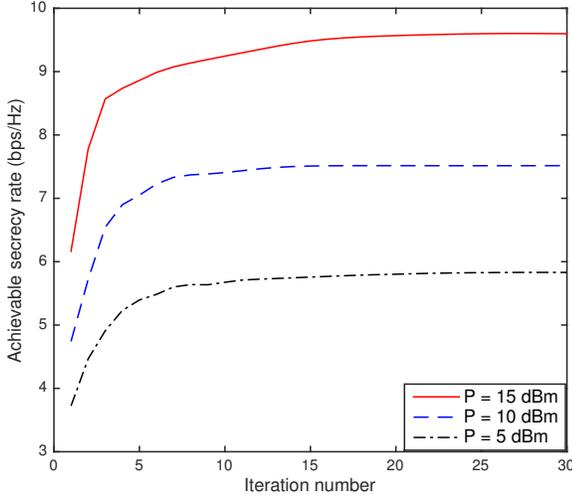


Fig. 3: Convergence performance of proposed BCD algorithm with different target transmit power.

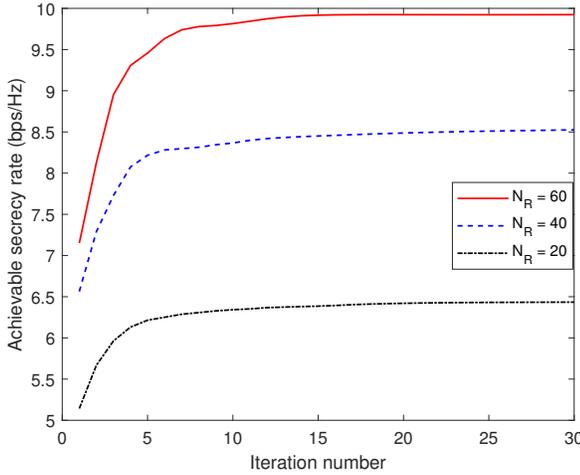


Fig. 4: Convergence performance of proposed BCD algorithm with different reflecting elements N_R .

To demonstrate the advantage of the proposed scheme (denoted by “MIMO IRS with AN” in the simulation results), we include the following three baseline schemes for comparison:

- 1) *Baseline 1* [32]: We consider an IRS aided MIMO secrecy system without the aided of AN, (i.e., $\mathbf{Z} = \mathbf{0}$) where the secure precoding and the phase shift matrices are jointly optimized via the BCD algorithm. This scheme is denoted as “MIMO IRS without AN”.
- 2) *Baseline 2* [26], [27]: The secrecy precoding and the AN precoding are jointly designed without the aided of the IRS in the MIMO secrecy system, which is denoted as “MIMO with AN and without IRS”.
- 3) *Baseline 3* [6]: We consider a classic MIMO wiretap channel without the aided of the AN and the IRS, which

is denoted by “MIMO without AN and without IRS”.

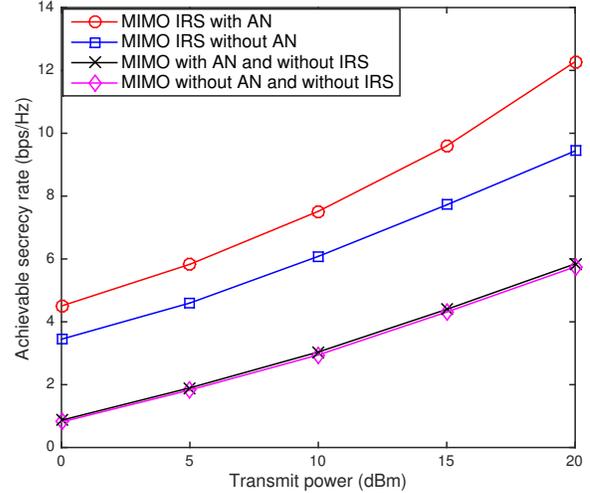


Fig. 5: Achievable secrecy rate versus transmit power.

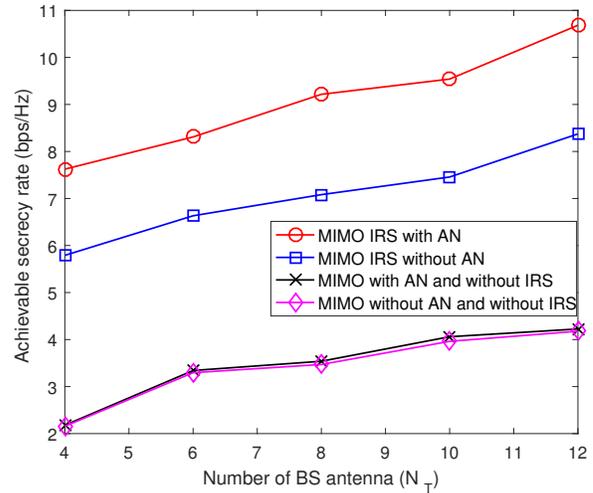


Fig. 6: Achievable secrecy rate versus the number of BS antenna N_T .

Next, we evaluate the achievable secrecy rate performance versus the maximum transmit power P in Fig. 5. One can observe from the figure that achievable secrecy rate increases with the maximum transmit power P , which confirms the property shown in Fig. 3. Also, we observe that the proposed scheme outperforms the other three baseline schemes,³ manifesting the positive role played by the IRS to enhance the achievable secrecy performance.⁴ Moreover, the scheme with AN and without IRS slightly outperforms that without AN and IRS. This is due to the fact that the power allocation of AN may not be sufficient to significantly improve the achievable secrecy rate such that the AN only introduce a slight secrecy performance improvement. However, it is evident that IRS

³“MIMO IRS without AN”, “MIMO with AN and without IRS”, and “MIMO without AN and without IRS”.

⁴The AN has been considered as one of most effective schemes to improve the achievable secrecy performance [27].

achieves a better secrecy performance than its counterpart without AN.

Fig. 6 shows that the achievable secrecy rate performance versus the number of the BS transmit antenna N_T , where it can be easily observed that a larger number of the BS transmit antenna leads to a higher achievable secrecy rate. Additionally, similar behaviour and conclusion from Fig. 5 are also observed here with different BS transmit antennas N_T . Fig. 7 describes the impact of the number of eavesdropper antenna N_E on the achievable secrecy rate performance. As expected, the achievable secrecy rate has a decreasing trend with the number of eavesdropper antenna N_E . Moreover, the proposed scheme has a better performance than other three baseline schemes due to the positive impact from the AN and the IRS.

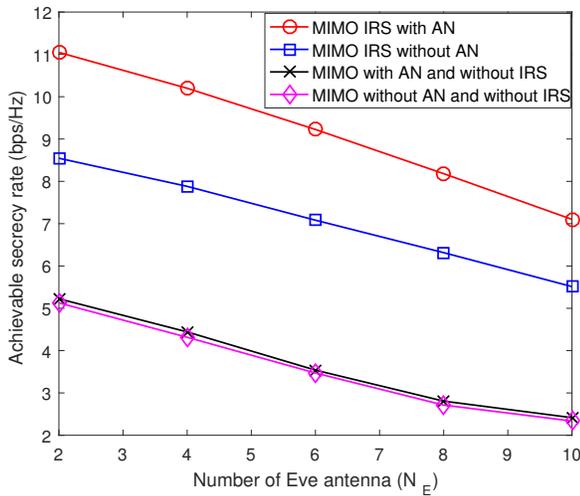


Fig. 7: Achievable secrecy rate versus the number of eavesdropper antenna N_E .

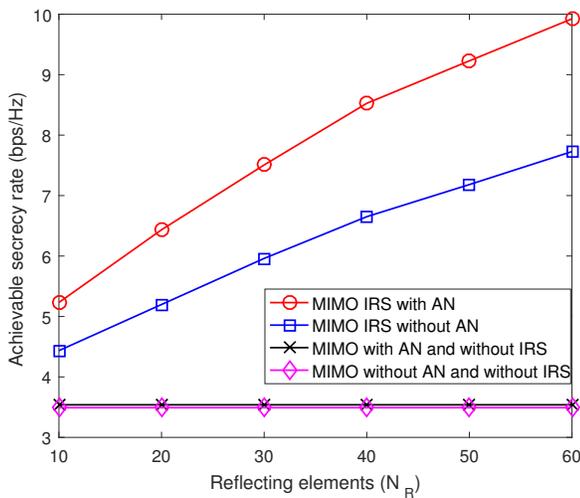


Fig. 8: Achievable secrecy rate versus reflecting elements N_R .

In Fig. 8, we evaluate the achievable secrecy rate performance versus the reflecting elements at the IRS N_R . With the help of IRS, the achievable secrecy rate exhibits an increasing behaviour with the reflecting elements N_R , whereas

the schemes without IRS remains constant with N_R in terms of the achievable secrecy rate. This is due to the fact that a larger number of reflecting elements brings a higher DoF to optimize secrecy performance, and a more constructive reflecting signal is produced to enhance the reception at the legitimate user. In addition, the performance gain between the IRS-aided schemes⁵ and the schemes without IRS⁶ becomes larger, which highlights the advantage of the IRS via joint optimization of the secure transmit precoding, the AN matrix and the phase shift matrix.

Finally, we evaluate the impact of the path loss exponent of the reflecting link (i.e., the IRS-user/eve link) at the IRS on the achievable secrecy rate performance. It aims to investigate the impact of the large-scale fading channel in the reflecting link on the secrecy system performance. Fig. 9 shows the achievable secrecy rate versus the path loss exponent of the IRS-user link $\beta_{IRS-user}$. It is apparent from the figure that the achievable secrecy rate exhibits a decreasing trend with $\beta_{IRS-user}$, owing to the fact that a larger-scale fading will result in a weaker signal reflected from the IRS, diminishing the benefits of the IRS. Additionally, the achievable secrecy rate versus the path loss exponent of the IRS-eve link $\beta_{IRS-eve}$ is shown in Fig. 10, where we observe that the achievable secrecy rate increases with $\beta_{IRS-eve}$. This follows from the fact that the weaker reflecting signal is introduced to degrade the reception at the eavesdropper with large-scale fading channel between the IRS and the eavesdropper. Actually, these two figures reveal an engineering insights that the IRS should be carefully deployed to enhance the system's secrecy performance with less blocking objects in the legitimate link or more blocking objects in the eavesdropping link.

VI. CONCLUSION

This paper investigated the IRS aided secure MIMO wireless networks, where the secure precoder, the AN jamming precoder and the phase shift matrix at the IRS have been jointly designed to maximize the achievable secrecy rate with the required transmit power budget. Due to the non-convexity of the formulated problem, we proposed a BCD algorithm to alternatively optimize the secure precoder, the AN jamming precoder and the phase shift matrix. Specifically, we derived the closed-form expression of the secure precoder and the AN jamming precoder via the WMMSE algorithm and KKT conditions, and the phase shift has been derived via the MM algorithm to obtain its closed-form solution. We also analyzed the convergence property of the proposed BCD algorithm. Finally, simulation results have been presented to demonstrate the superiority of the proposed scheme over the baseline schemes, which highlights the positive role of the IRS for the secrecy performance enhancement in the MIMO wireless communication systems.

⁵“MIMO IRS with AN” and “MIMO IRS without AN”.

⁶“MIMO with AN and without IRS”, and “MIMO without AN and without IRS”.

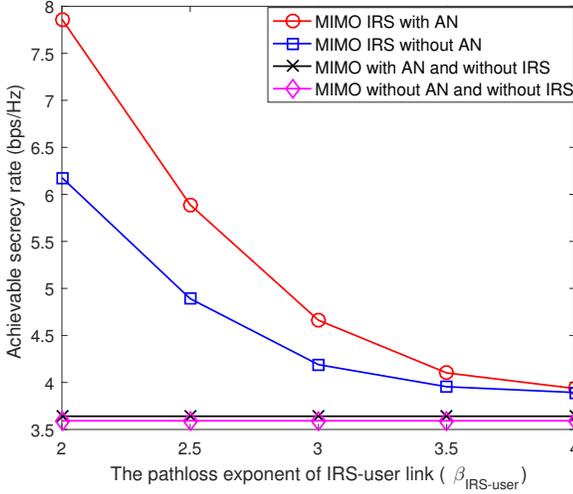


Fig. 9: Achievable secrecy rate versus the path loss exponent of the IRS reflecting link.

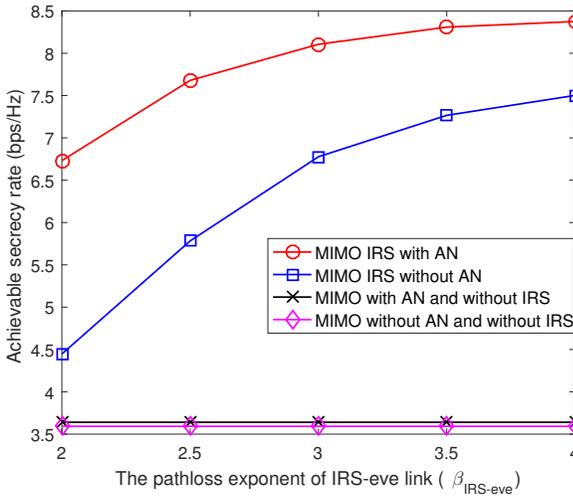


Fig. 10: Achievable secrecy rate versus the path loss exponent of the IRS reflecting link.

APPENDIX

PROOF OF *Theorem 1*

First, we show that the objective value in (30), denoted as $f(\theta)$, has a non-increasing trend and guaranteed convergence. Also, $\theta^{(m)}$ is also the optimal solution to problem (35) due to the equivalence between (35) and (36). Thus, $g(\theta^{(m+1)}|\theta^{(m)}) \leq g(\theta^{(m)}|\theta^{(m)})$ holds. In addition, by exploiting these three conditions in (33) to easily achieve the inequality $f(\theta^{(m)}) \geq f(\theta^{(m+1)})$. Moreover, the unit modulus constraint incurs a lower bound of $f(\theta^{(m)})$, which ensures the convergence of *Algorithm 2*. The first part of *Theorem 1* has been completed.

On the other hand, we will show that the converged solution generated via *Algorithm 2* is a KKT point of problem (22). It means that its optimal solution, denoted by θ^{opt} , has to satisfy the KKT conditions of problem (35). Thus, the following

Lagrange function is considered:

$$\mathcal{L}(\theta, \lambda) = g(\theta|\theta^{(m)}) + \sum_{n=1}^{N_R} \lambda_n (|\theta_n| - 1), \quad (43)$$

where $\lambda = [\lambda_1, \dots, \lambda_{N_R}]$ denotes the dual variables associated with the unit modulus constraint. It can be verified that there exists an optimal λ^{opt} to guarantee that the following KKT conditions hold:

$$\begin{aligned} \nabla_{\theta^{\text{opt}}} \mathcal{L}(\theta, \lambda)|_{\theta=\theta^{\text{opt}}} &= \nabla_{\theta^{\text{opt}}} g(\theta|\theta^{\text{opt}}) + \sum_{n=1}^{N_R} \lambda_n (\nabla_{\theta^{\text{opt}}} |\theta_n|)|_{\theta=\theta^{\text{opt}}} = \mathbf{0}, \quad (44a) \end{aligned}$$

$$\lambda_n (|\theta_n^{\text{opt}}| - 1) = 0, \forall n. \quad (44b)$$

By applying the second condition in (33),

$$\nabla_{\theta^{\text{opt}}} g(\theta|\theta^{\text{opt}})|_{\theta=\theta^{\text{opt}}} = \nabla_{\theta^{\text{opt}}} f(\theta)|_{\theta=\theta^{\text{opt}}}. \quad (45)$$

Denote the objective value of (22) as $\rho(\theta)$, it can be easily verified that $\nabla_{\theta^{\text{opt}}} \rho(\theta)|_{\theta=\theta^{\text{opt}}} = \nabla_{\theta^{\text{opt}}} f(\theta)|_{\theta=\theta^{\text{opt}}}$ due to a series of equivalent mathematical manipulations employed from (22) to (30). To proceed, $\nabla_{\theta^{\text{opt}}} g(\theta|\theta^{\text{opt}})|_{\theta=\theta^{\text{opt}}} = \nabla_{\theta^{\text{opt}}} \rho(\theta)|_{\theta=\theta^{\text{opt}}}$ holds. Therefore, we have

$$\begin{aligned} \nabla_{\theta^{\text{opt}}} \mathcal{L}(\theta, \lambda)|_{\theta=\theta^{\text{opt}}} &= \nabla_{\theta^{\text{opt}}} \rho(\theta)|_{\theta=\theta^{\text{opt}}} + \sum_{n=1}^{N_R} \lambda_n (\nabla_{\theta^{\text{opt}}} |\theta_n|)|_{\theta=\theta^{\text{opt}}} = \mathbf{0}, \quad (46) \end{aligned}$$

which integrates with (44b) to guarantee the KKT conditions of problem (22). The second part of *Theorem 1* has been completed.

PROOF OF *Theorem 2*

In order to prove *Theorem 2*, we denote the objective function of problem (10) as $\tilde{R}_{\text{sec}}(\mathbf{W}, \mathbf{V}, \mathbf{X}_i, \mathbf{Z}_j, \Theta)$. Also, let us denote $(\mathbf{W}^{(n+1)}, \mathbf{V}^{(n+1)}, \mathbf{X}_i^{(n+1)}, \mathbf{Z}_j^{(n+1)}, \Theta^{(n+1)})$ as the updated solution at the $(m+1)$ -th iteration, which can be obtained via Step 2 of *Algorithm 3*. It can be verified that the following relation holds

$$\begin{aligned} &R_{\text{sec}}(\mathbf{W}^{(n)}, \mathbf{V}^{(n)}, \Theta^{(n)}) \\ &= \tilde{R}_{\text{sec}}(\mathbf{W}^{(n)}, \mathbf{V}^{(n)}, \mathbf{X}_i^{(n)}, \mathbf{Z}_j^{(n)}, \Theta^{(n)}) \\ &\leq \tilde{R}_{\text{sec}}(\mathbf{W}^{(n+1)}, \mathbf{V}^{(n+1)}, \mathbf{X}_i^{(n)}, \mathbf{Z}_j^{(n)}, \Theta^{(n)}) \\ &\leq \tilde{R}_{\text{sec}}(\mathbf{W}^{(n+1)}, \mathbf{V}^{(n+1)}, \mathbf{X}_i^{(n)}, \mathbf{Z}_j^{(n)}, \Theta^{(n+1)}) \\ &\leq \tilde{R}_{\text{sec}}(\mathbf{W}^{(n+1)}, \mathbf{V}^{(n+1)}, \mathbf{X}_i^{(n+1)}, \mathbf{Z}_j^{(n+1)}, \Theta^{(n+1)}) \\ &= R_{\text{sec}}(\mathbf{W}^{(n+1)}, \mathbf{V}^{(n+1)}, \Theta^{(n+1)}). \quad (47) \end{aligned}$$

where the first and last equalities are due to (13a)-(13c), the first inequality is due to Step 2-a of *Algorithm 3*, the second inequality is on the basis of Step 2-b of *Algorithm 3*, and the third inequality follows Step 2-c and 2-d of *Algorithm 3*, respectively. In addition, each solution sequence $(\mathbf{W}^{(n)}, \mathbf{V}^{(n)})$, $(n = 1, 2, \dots, L_1)$ is the feasible solution to problem (10). Thus, (47) incurs a monotonically increasing trend of the objective value sequence $R_{\text{sec}}(\mathbf{W}^{(n)}, \mathbf{V}^{(n)}, \Theta^{(n)})$. Moreover, $R_{\text{sec}}(\mathbf{W}^{(n)}, \mathbf{V}^{(n)}, \Theta^{(n)})$ is upper-bounded due to the transmit power constraint in (10).

Next, we will show that the converged solution $(\mathbf{W}^{\text{opt}}, \mathbf{V}^{\text{opt}}, \mathbf{X}_i^{\text{opt}}, \mathbf{Z}_j^{\text{opt}}, \Theta^{\text{opt}})$ secures the KKT conditions of problem (4). It can be readily shown that the optimal solution $(\mathbf{W}^{\text{opt}}, \mathbf{V}^{\text{opt}})$ satisfies the KKT conditions of problem (14)

[27]. We denote the objective function of problem (14) as $C(\mathbf{W}, \mathbf{V}, \Theta^{\text{opt}})$, and write the Lagrangian function of (14) as follows:

$$\mathcal{L}(\mathbf{W}, \mathbf{V}, \tau) = C(\mathbf{W}, \mathbf{V}, \Theta^{\text{opt}}) + \tau (\text{Tr}(\mathbf{W}\mathbf{W}^H) + \text{Tr}(\mathbf{V}\mathbf{V}^H) - P), \quad (48)$$

where $\tau \geq 0$ denotes the dual variable associated with the transmit power constraint. The Slater's condition holds (14) and $(\mathbf{W}^{\text{opt}}, \mathbf{V}^{\text{opt}})$ is a KKT point of (14) [33]. Thus, there exists an optimal τ^{opt} to guarantee the following KKT conditions:

$$\begin{aligned} \nabla_{\mathbf{W}^{\text{opt}}} \mathcal{L}(\mathbf{W}, \mathbf{V}, \tau) \Big|_{\mathbf{W}=\mathbf{W}^{\text{opt}}} \\ = \nabla_{\mathbf{W}^{\text{opt}}} C(\mathbf{W}, \mathbf{V}, \Theta^{\text{opt}}) \Big|_{\mathbf{W}=\mathbf{W}^{\text{opt}}} + 2\tau^{\text{opt}} \mathbf{W}^{\text{opt}} = \mathbf{0}, \end{aligned} \quad (49)$$

$$\begin{aligned} \nabla_{\mathbf{V}^{\text{opt}}} \mathcal{L}(\mathbf{W}, \mathbf{V}, \tau) \Big|_{\mathbf{V}=\mathbf{V}^{\text{opt}}} \\ = \nabla_{\mathbf{V}^{\text{opt}}} C(\mathbf{W}, \mathbf{V}, \Theta^{\text{opt}}) \Big|_{\mathbf{V}=\mathbf{V}^{\text{opt}}} + 2\tau^{\text{opt}} \mathbf{V}^{\text{opt}} = \mathbf{0}, \end{aligned} \quad (50)$$

$$\tau^{\text{opt}} (\text{Tr}(\mathbf{W}\mathbf{W}^H) + \text{Tr}(\mathbf{V}\mathbf{V}^H) - P) = 0. \quad (51)$$

We can also verify the following relations:

$$\begin{aligned} \nabla_{\mathbf{W}^{\text{opt}}} \tilde{R}_{sec}(\mathbf{W}, \mathbf{V}, \mathbf{X}_i^{\text{opt}}, \mathbf{Z}_j^{\text{opt}}, \Theta^{\text{opt}}) \Big|_{\mathbf{W}=\mathbf{W}^{\text{opt}}} \\ = \nabla_{\mathbf{W}^{\text{opt}}} C(\mathbf{W}, \mathbf{V}, \Theta^{\text{opt}}) \Big|_{\mathbf{W}=\mathbf{W}^{\text{opt}}}, \end{aligned} \quad (52)$$

$$\begin{aligned} \nabla_{\mathbf{V}^{\text{opt}}} \tilde{R}_{sec}(\mathbf{W}, \mathbf{V}, \mathbf{X}_i^{\text{opt}}, \mathbf{Z}_j^{\text{opt}}, \Theta^{\text{opt}}) \Big|_{\mathbf{V}=\mathbf{V}^{\text{opt}}} \\ = \nabla_{\mathbf{V}^{\text{opt}}} C(\mathbf{W}, \mathbf{V}, \Theta^{\text{opt}}) \Big|_{\mathbf{V}=\mathbf{V}^{\text{opt}}}. \end{aligned} \quad (53)$$

To proceed, we exploit the left hand side (LHS) of (52) and (53) via taking the first-derivatives of $\mathbf{W} = \mathbf{W}^{\text{opt}}$ and $\mathbf{V} = \mathbf{V}^{\text{opt}}$. Now, we consider the first derivative of \tilde{R}_{sec} with respect to $\mathbf{W} = \mathbf{W}^{\text{opt}}$ as follows:

$$\begin{aligned} \nabla_{\mathbf{W}^{\text{opt}}} \tilde{R}_{sec}(\mathbf{W}, \mathbf{V}, \mathbf{X}_i^{\text{opt}}, \mathbf{Z}_j^{\text{opt}}, \Theta^{\text{opt}}) \Big|_{\mathbf{W}=\mathbf{W}^{\text{opt}}} \\ = -\text{Tr} \left[\mathbf{X}_1 \left(\nabla_{\mathbf{W}^{\text{opt}}} \mathbf{E}_1(\mathbf{Z}_1^{\text{opt}}, \mathbf{W}, \mathbf{V}, \Theta^{\text{opt}}) \Big|_{\mathbf{W}=\mathbf{W}^{\text{opt}}} \right) \right] \\ - \text{Tr} \left[\mathbf{X}_3 \left(\nabla_{\mathbf{W}^{\text{opt}}} \mathbf{E}_3(\mathbf{W}, \mathbf{V}, \Theta^{\text{opt}}) \Big|_{\mathbf{W}=\mathbf{W}^{\text{opt}}} \right) \right]. \end{aligned} \quad (54)$$

Substituting the closed-form expressions of \mathbf{X}_1 and \mathbf{X}_3 shown in (13a) and (13c) into (54), yields

$$\begin{aligned} \nabla_{\mathbf{W}^{\text{opt}}} \tilde{R}_{sec}(\mathbf{W}, \mathbf{V}, \mathbf{X}_i^{\text{opt}}, \mathbf{Z}_j^{\text{opt}}, \Theta^{\text{opt}}) \Big|_{\mathbf{W}=\mathbf{W}^{\text{opt}}} \\ = -\text{Tr} \left[\mathbf{E}_1(\mathbf{Z}_1^{\text{opt}}, \mathbf{W}, \mathbf{V}, \Theta^{\text{opt}})^{-1} \right. \\ \left. \left(\nabla_{\mathbf{W}^{\text{opt}}} \mathbf{E}_1(\mathbf{Z}_1^{\text{opt}}, \mathbf{W}, \mathbf{V}, \Theta^{\text{opt}}) \Big|_{\mathbf{W}=\mathbf{W}^{\text{opt}}} \right) \right] \\ - \text{Tr} \left[\mathbf{E}_3(\mathbf{W}, \mathbf{V}, \Theta^{\text{opt}})^{-1} \left(\nabla_{\mathbf{W}^{\text{opt}}} \mathbf{E}_3(\mathbf{W}, \mathbf{V}, \Theta^{\text{opt}}) \Big|_{\mathbf{W}=\mathbf{W}^{\text{opt}}} \right) \right], \end{aligned} \quad (55)$$

By applying matrix differentiation identity, we have

$$\begin{aligned} \nabla_{\mathbf{W}^{\text{opt}}} \tilde{R}_{sec}(\mathbf{W}, \mathbf{V}, \mathbf{X}_i^{\text{opt}}, \mathbf{Z}_j^{\text{opt}}, \Theta^{\text{opt}}) \Big|_{\mathbf{W}=\mathbf{W}^{\text{opt}}} \\ = -\nabla_{\mathbf{W}^{\text{opt}}} \log |\mathbf{C}_1(\mathbf{Z}_1^{\text{opt}}, \mathbf{W}, \mathbf{V}, \Theta^{\text{opt}})| \Big|_{\mathbf{W}=\mathbf{W}^{\text{opt}}} \\ - \nabla_{\mathbf{W}^{\text{opt}}} \log |\mathbf{E}_3(\mathbf{W}, \mathbf{V}, \Theta^{\text{opt}})| \Big|_{\mathbf{W}=\mathbf{W}^{\text{opt}}}, \\ = \nabla_{\mathbf{W}^{\text{opt}}} \log |\mathbf{X}_1(\mathbf{Z}_1^{\text{opt}}, \mathbf{W}, \mathbf{V}, \Theta^{\text{opt}})| \Big|_{\mathbf{W}=\mathbf{W}^{\text{opt}}} \\ + \nabla_{\mathbf{W}^{\text{opt}}} \log |\mathbf{X}_3(\mathbf{W}, \mathbf{V}, \Theta^{\text{opt}})| \Big|_{\mathbf{W}=\mathbf{W}^{\text{opt}}} \\ = \nabla_{\mathbf{W}^{\text{opt}}} R_{sec}(\mathbf{W}, \mathbf{V}, \Theta^{\text{opt}}) \Big|_{\mathbf{W}=\mathbf{W}^{\text{opt}}}. \end{aligned} \quad (56)$$

Similarly, we have

$$\begin{aligned} \nabla_{\mathbf{V}^{\text{opt}}} \tilde{R}_{sec}(\mathbf{W}, \mathbf{V}, \mathbf{X}_i^{\text{opt}}, \mathbf{Z}_j^{\text{opt}}, \Theta^{\text{opt}}) \Big|_{\mathbf{V}=\mathbf{V}^{\text{opt}}} \\ = \nabla_{\mathbf{V}^{\text{opt}}} R_{sec}(\mathbf{W}, \mathbf{V}, \Theta^{\text{opt}}) \Big|_{\mathbf{V}=\mathbf{V}^{\text{opt}}}, \end{aligned} \quad (57)$$

$$\begin{aligned} \nabla_{\Theta^{\text{opt}}} \tilde{R}_{sec}(\mathbf{W}^{\text{opt}}, \mathbf{V}^{\text{opt}}, \mathbf{X}_i^{\text{opt}}, \mathbf{Z}_j^{\text{opt}}, \Theta) \Big|_{\Theta=\Theta^{\text{opt}}} \\ = \nabla_{\Theta^{\text{opt}}} R_{sec}(\mathbf{W}^{\text{opt}}, \mathbf{V}^{\text{opt}}, \Theta) \Big|_{\Theta=\Theta^{\text{opt}}}. \end{aligned} \quad (58)$$

Thus, combining with (52) and (53), we obtain

$$\begin{aligned} \nabla_{\mathbf{W}^{\text{opt}}} R_{sec}(\mathbf{W}, \mathbf{V}, \Theta^{\text{opt}}) \Big|_{\mathbf{W}=\mathbf{W}^{\text{opt}}} \\ = \nabla_{\mathbf{W}^{\text{opt}}} C(\mathbf{W}, \mathbf{V}, \Theta^{\text{opt}}) \Big|_{\mathbf{W}=\mathbf{W}^{\text{opt}}}, \end{aligned} \quad (59)$$

$$\begin{aligned} \nabla_{\mathbf{V}^{\text{opt}}} R_{sec}(\mathbf{W}, \mathbf{V}, \Theta^{\text{opt}}) \Big|_{\mathbf{V}=\mathbf{V}^{\text{opt}}} \\ = \nabla_{\mathbf{V}^{\text{opt}}} C(\mathbf{W}, \mathbf{V}, \Theta^{\text{opt}}) \Big|_{\mathbf{V}=\mathbf{V}^{\text{opt}}}. \end{aligned} \quad (60)$$

We substitute (59) and (60) into (49) and (50), respectively,

$$\begin{aligned} \nabla_{\mathbf{W}^{\text{opt}}} \mathcal{L}(\mathbf{W}, \mathbf{V}, \tau) \Big|_{\mathbf{W}=\mathbf{W}^{\text{opt}}} \\ = \nabla_{\mathbf{W}^{\text{opt}}} R_{sec}(\mathbf{W}, \mathbf{V}, \Theta^{\text{opt}}) \Big|_{\mathbf{W}=\mathbf{W}^{\text{opt}}} + 2\tau^{\text{opt}} \mathbf{W}^{\text{opt}} = \mathbf{0}, \end{aligned} \quad (61)$$

$$\begin{aligned} \nabla_{\mathbf{V}^{\text{opt}}} \mathcal{L}(\mathbf{W}, \mathbf{V}, \tau) \Big|_{\mathbf{V}=\mathbf{V}^{\text{opt}}} \\ = \nabla_{\mathbf{V}^{\text{opt}}} R_{sec}(\mathbf{W}, \mathbf{V}, \Theta^{\text{opt}}) \Big|_{\mathbf{V}=\mathbf{V}^{\text{opt}}} + 2\tau^{\text{opt}} \mathbf{V}^{\text{opt}} = \mathbf{0}. \end{aligned} \quad (62)$$

On the other hand, we have proved that Θ^{opt} can satisfy the KKT conditions of problem (22) such that (44b) and (46) hold, and we can readily obtain

$$\nabla_{\Theta^{\text{opt}}} \tilde{R}_{sec}(\mathbf{W}^{\text{opt}}, \mathbf{V}^{\text{opt}}, \mathbf{X}_i^{\text{opt}}, \mathbf{Z}_j^{\text{opt}}, \Theta) \Big|_{\Theta=\Theta^{\text{opt}}} = \nabla_{\Theta^{\text{opt}}} \rho(\Theta) \Big|_{\Theta=\Theta^{\text{opt}}}. \quad (63)$$

Combining (63) with (58), the following relation holds

$$\nabla_{\Theta^{\text{opt}}} R_{sec}(\mathbf{W}^{\text{opt}}, \mathbf{V}^{\text{opt}}, \Theta) \Big|_{\Theta=\Theta^{\text{opt}}} = \nabla_{\Theta^{\text{opt}}} \rho(\Theta) \Big|_{\Theta=\Theta^{\text{opt}}}. \quad (64)$$

We substitute (64) into (46), it is easily verified that

$$\begin{aligned} \nabla_{\Theta^{\text{opt}}} \mathcal{L}(\theta, \lambda) \Big|_{\theta=\theta^{\text{opt}}} = \nabla_{\Theta^{\text{opt}}} R_{sec}(\mathbf{W}^{\text{opt}}, \mathbf{V}^{\text{opt}}, \Theta) \Big|_{\Theta=\theta^{\text{opt}}} \\ + \sum_{n=1}^{N_R} \lambda_n (\nabla_{\theta^{\text{opt}}} |\theta_n|) \Big|_{\theta=\theta^{\text{opt}}} = \mathbf{0}. \end{aligned} \quad (65)$$

Therefore, the KKT conditions (44b), (51), (61), (62), and (65) are exactly the KKT conditions of problem (4), which

has completed the proof of *Theorem 2*.

REFERENCES

- [1] Z. Gao, L. Dai, D. Mi, Z. Wang, M. A. Imran, and M. Z. Shakir, "MmWave massive-MIMO-based wireless backhaul for the 5G ultra-dense network," *IEEE Wireless Commun.*, vol. 22, no. 5, pp. 13–21, Oct. 2015.
- [2] X. Chen, C. Zhong, C. Yuen, and H. Chen, "Multi-antenna relay aided wireless physical layer security," *IEEE Commun. Mag.*, vol. 53, no. 12, pp. 40–46, Dec. 2015.
- [3] S. Chen, F. Qin, B. Hu, X. Li, and Z. Chen, "User-centric ultra-dense networks for 5G: challenges, methodologies, and directions," *IEEE Wireless Commun.*, vol. 23, no. 2, pp. 78–85, Apr. 2016.
- [4] Renzo, Marco Di, et al., "Smart radio environments empowered by reconfigurable AI meta-surfaces: an idea whose time has come," *EURASIP J. Wirel. Commun. Netw.*, no. 129, pp. 1–20, May 2019.
- [5] Y. Wu, A. Khisti, C. Xiao, G. Caire, K. Wong, and X. Gao, "A survey of physical layer security techniques for 5G wireless networks and challenges ahead," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 4, pp. 679–695, Apr. 2018.
- [6] K. Cumanan, Z. Ding, B. Sharif, G. Y. Tian, and K. K. Leung, "Secrecy rate optimizations for a MIMO secrecy channel with a multiple-antenna eavesdropper," *IEEE Trans. Veh. Technol.*, vol. 63, no. 4, pp. 1678–1690, May 2014.
- [7] Q. Li and W.-K. Ma, "Optimal and robust transmit designs for MISO channel secrecy by semidefinite programming," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3799–3812, Aug. 2011.
- [8] Z. Chu, K. Cumanan, Z. Ding, M. Johnston, and S. Le Goff, "Secrecy rate optimizations for a MIMO secrecy channel with a cooperative jammer," *IEEE Trans. Veh. Technol.*, vol. 64, no. 5, pp. 1833–1847, May 2015.
- [9] F. Zhou, Z. Li, J. Cheng, Q. Li, and J. Si, "Robust AN-aided beamforming and power splitting design for secure MISO cognitive radio with SWIPT," *IEEE Trans. Wireless Commun.*, vol. 16, no. 4, pp. 2450–2464, Apr. 2017.
- [10] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network: Joint active and passive beamforming design," *CoRR*, vol. abs/1809.01423, 2018.
- [11] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Commun. Mag.*, vol. 58, no. 1, pp. 106–112, 2020.
- [12] Q. Wu and R. Zhang, "Joint active and passive beamforming optimization for intelligent reflecting surface assisted SWIPT under QoS constraints," *IEEE J. Sel. Area Commun.*, vol. 38, no. 8, pp. 1735–1748, 2020.
- [13] S. Hu, F. Rusek, and O. Edfors, "Beyond massive MIMO: The potential of positioning with large intelligent surfaces," *IEEE Trans. Signal Process.*, vol. 66, no. 7, pp. 1761–1774, Apr. 2018.
- [14] G. Zhou, C. Pan, H. Ren, K. Wang, and A. Nallanathan, "Intelligent reflecting surface aided multigroup multicast miso communication systems," *IEEE Trans. Signal Process.*, vol. 68, pp. 3236–3251, 2020.
- [15] C. Pan, H. Ren, K. Wang, W. Xu, M. ElKashlan, A. Nallanathan, and L. Hanzo, "Multicell MIMO communications relying on intelligent reflecting surfaces," *IEEE Trans. Wireless Commun.*, vol. 19, no. 8, pp. 5218–5233, 2020.
- [16] 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 J. Sel. Area Commun.*, vol. 38, no. 8, pp. 1719–1734, 2020.
- [17] H. Shen, W. Xu, S. Gong, Z. He, and C. Zhao, "Secrecy rate maximization for intelligent reflecting surface assisted multi-antenna communications," *IEEE Commun. Lett.*, vol. 23, no. 9, pp. 1488–1492, Sep. 2019.
- [18] M. Cui, G. Zhang, and R. Zhang, "Secure wireless communication via intelligent reflecting surface," *IEEE Wireless Commun. Lett.*, vol. 8, no. 5, pp. 1410–1414, Oct. 2019.
- [19] Z. Chu, W. Hao, P. Xiao, and J. Shi, "Intelligent reflecting surface aided multi-antenna secure transmission," *IEEE Wireless Commun. Lett.*, vol. 9, no. 1, pp. 108–112, Jan. 2020.
- [20] X. Guan, Q. Wu, and R. Zhang, "Intelligent reflecting surface assisted secrecy communication: Is artificial noise helpful or not?," *IEEE Wireless Commun. Lett.*, vol. 9, no. 6, pp. 778–782, 2020.
- [21] H. Zhang and L. Duan, "Going beyond secrecy rate via information jamming," in *IEEE GLOBECOM*, Abu Dhabi, UAE, pp. 1–6, Dec. 2018.
- [22] G. Chen, Y. Gong, P. Xiao, and J. A. Chambers, "Physical layer network security in the full-duplex relay system," *IEEE Trans. Inf. Forensics Secur.*, vol. 10, no. 3, pp. 574–583, Mar. 2015.
- [23] Y. Sun, P. Babu, and D. P. Palomar, "Majorization-minimization algorithms in signal processing, communications, and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 794–816, Feb. 2017.
- [24] G. Geraci, M. Egan, J. Yuan, A. Razi, and I. B. Collings, "Secrecy sum-rates for multi-user MIMO regularized channel inversion precoding," *IEEE Trans. Commun.*, vol. 60, no. 11, pp. 3472–3482, 2012.
- [25] A. Mukherjee and A. L. Swindlehurst, "Detecting passive eavesdroppers in the MIMO wiretap channel," in *IEEE, ICASSP, Kyoto, Japan*, pp. 2809–2812, 2012.
- [26] Q. Li, M. Hong, H. Wai, Y. Liu, W. Ma, and Z. Luo, "Transmit solutions for MIMO wiretap channels using alternating optimization," *IEEE J. Sel. Area Commun.*, vol. 31, no. 9, pp. 1714–1727, Sept. 2013.
- [27] Q. Shi, W. Xu, J. Wu, E. Song, and Y. Wang, "Secure beamforming for MIMO broadcasting with wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2841–2853, May 2015.
- [28] Q. Shi, M. Razaviyayn, Z. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, Sep. 2011.
- [29] X.-D. Zhang, *Matrix Analysis and Applications*. Cambridge University Press, 2017.
- [30] J. Song, P. Babu, and D. P. Palomar, "Sequence design to minimize the weighted integrated and peak sidelobe levels," *IEEE Trans. Signal Process.*, vol. 64, no. 8, pp. 2051–2064, Apr. 2016.
- [31] W. Liao, T. Chang, W. Ma, and C. Chi, "QoS-based transmit beamforming in the presence of eavesdroppers: An optimized artificial-noise-aided approach," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1202–1216, Mar. 2011.
- [32] H. Shen, W. Xu, S. Gong, Z. He, and C. Zhao, "Secrecy rate maximization for intelligent reflecting surface assisted multi-antenna communications," *IEEE Commun. Lett.*, vol. 23, no. 9, pp. 1488–1492, Sept. 2019.
- [33] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.