

Robust Beamforming and Jamming for Enhancing the Physical Layer Security of Full Duplex Radios

Zhengmin Kong, Shaoshi Yang, Die Wang, Lajos Hanzo, *Fellow, IEEE*

Abstract—In this paper, we investigate the physical layer security of a full-duplex base station (BS) aided system in the worst case, where an uplink transmitter (UT) and a downlink receiver (DR) are both equipped with a single antenna, while a powerful eavesdropper is equipped with multiple antennas. For securing the confidentiality of signals transmitted from the BS and UT, an artificial noise (AN) aided secrecy beamforming scheme is proposed, which is robust to the realistic imperfect state information of both the eavesdropping channel and the residual self-interference channel. Our objective function is that of maximizing the worst-case sum secrecy rate achieved by the BS and UT, through jointly optimizing the beamforming vector of the confidential signals and the transmit covariance matrix of the AN. However, the resultant optimization problem is non-convex and non-linear. In order to efficiently obtain the solution, we transform the non-convex problem into a sequence of convex problems by adopting the block coordinate descent algorithm. We invoke a linear matrix inequality for finding its Karush-Kuhn-Tucker (KKT) solution. In order to evaluate the achievable performance, the worst-case secrecy rate is derived analytically. Furthermore, we construct another secrecy transmission scheme using the projection matrix theory for performance comparison. Our simulation results show that the proposed robust secrecy transmission scheme achieves substantial secrecy performance gains, which verifies the efficiency of the proposed method.

Index Terms—Physical layer security, full-duplex, artificial noise, secrecy beamforming, secrecy rate

I. INTRODUCTION

Full-duplex (FD) communication has the potential of doubling the spectral efficiency compared to its half-duplex counterpart. The main challenge facing FD systems is the increased self-interference imposed by the leakage from the transmitted signals to the signals received at the FD node. Sophisticated techniques, such as TX/RX antenna separation and isolation, as well as digital/analog/propagation domain interference cancellation, have been proposed for combating the self-interference [1], [2]. Thus the design of a single-channel FD system becomes feasible.

This work is financially supported by the National Natural Science Foundation of China (NSFC) (Grant No. 61801518) and the Hubei Provincial Natural Science Foundation of China (Grant No. 2017CFB661).

Z. Kong and D. Wang are with the Automation Department, Wuhan University, Wuhan 430072, China (e-mail: zmkong@whu.edu.cn, wangdie1995@whu.edu.cn).

S. Yang is with the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, and with the Key Laboratory of Universal Wireless Communications, Ministry of Education, Beijing 100876, China (e-mail: shaoshi.yang@ieee.org).

L. Hanzo is with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K. (e-mail: lh@ecs.soton.ac.uk)

On the other hand, exploiting artificial noise (AN) [3] has been widely recognized as an effective technique of improving the physical layer security (PLS) [4]–[9] of wireless communications. By transmitting AN in addition to the confidential signals, the quality of the eavesdropping channel is degraded. However, existing AN based PLS schemes cannot be used for securing uplink (UL) transmissions in conventional FDD/TDD cellular systems [10], [11], because i) the base stations (BSs) cannot transmit AN while receiving co-channel desired signals; ii) an UL transmitter (UT) typically has a single transmit antenna that can only be used for transmitting desired information-bearing signals to a BS, and no more antennas are available for transmitting AN. Notably, FD techniques provide a solution to such a dilemma, since an FD transceiver is capable of receiving the confidential signals, while simultaneously transmitting AN in order to jam potential eavesdroppers (Eves). The authors of [12] proposed a user-grouping-based fractional time model relying either on perfect channel state information (CSI) or on statistical CSI. Assuming perfect self-interference suppression, the authors of [13] studied the joint design of the confidential signal and AN for maximizing the instantaneous secrecy rate achieved by the legitimate downlink (DL) receiver (DR), while keeping the received secrecy rate at the FD BS above a predefined target. By contrast, assuming imperfect self-interference cancellation, the authors of [14] studied the joint design of the confidential signal and AN for minimizing the total transmit power, while keeping the achievable secrecy rate of the FD BS and of the DR above a predefined target. Furthermore, the authors of [15] proposed a pair of relay-assisted secure protocols by relying on the FD capability of an orthogonal frequency division multiple access (OFDMA) system.

However, none of the aforementioned contributions considered the scenario of realistic imperfect CSI, even though only imperfect eavesdropping channel information might be gleaned by the BS - if any at all - since typically the Eves sends no training signals to the BS.

Although the authors of [16], [17] analyzed the secrecy outage probability in an imperfect eavesdropper CSI knowledge scenario for a FD relay system, the eavesdropper was only equipped with a single antenna and the source was unable to communicate with the destination directly.

Hence, more realistic practical problems have to be considered. More specifically, the security in the DL and UL have to be considered simultaneously, and the sum secrecy rate of the DL and UL should be optimized in the face of a

“sophisticated/strong” eavesdropper having imperfect CSI¹.

In contrast to [12]–[16], in this paper, our new contribution is that we jointly design the beamforming vector of the confidential signal and the covariance matrix of the AN, in order to maximize the worst-case² sum secrecy rate achieved by the FD BS and the UT, under the realistic assumption that only imperfect state information of the residual self-interference channel and of the eavesdropping channel is available to the FD BS.

To the best of our knowledge, there is no open literature addressing the simultaneous optimization of the sum secrecy rate of the DL and UL in the face of realistic CSI error concerning a “sophisticated/strong” eavesdropper’s channel in the above-mentioned worst-case FD scenario. Since the worst-case optimization problem is considered, in the paper, we adopt the deterministic model of [18], [19] to characterize the imperfect channel state information (CSI). Because the resultant objective function considered is non-convex, the global optimum is hard to obtain. As a low-complexity suboptimal method, the block coordinate descent (BCD) algorithm of [20, Section 2.7] is adopted to transform the non-convex and non-linear problem into a sequence of convex problems. In other words, without using any approximation method such as the classic semi-definite relaxation (SDR) technique, we transform our non-convex and non-linear problem into a more well-behaved form and conceive an efficient algorithm for finding its Karush-Kuhn-Tucke(KKT) solution. Then the locally optimal solution of the original problem is obtained. Our analysis and simulation results reveal valuable insights into how several relevant design parameters affect the achievable secrecy performance of the system considered; compared to projection matrix theory, the proposed robust secrecy transmission scheme is capable of achieving substantial secrecy performance gains.

The rest of this paper is organized as follows. In Section II, the system model and the problem formulation are described. In Section III, we transform the original non-convex problem into a sequence of convex problems and invoke an iterative BCD algorithm for finding its Karush-Kuhn-Tucke(KKT) solution. Our simulation results are provided in Section IV, and the conclusions are offered in Section V.

Notation: $(\cdot)^T$, $(\cdot)^*$ and $(\cdot)^H$ denote the transpose, conjugate and conjugate transpose, respectively. $(\cdot)^{-1}$, $\text{tr}(\cdot)$, and $\|\cdot\|_F$ denote the inverse, trace, and Frobenius norm of a matrix, respectively. $\ln(\cdot)$ denotes the natural logarithm, and \otimes represents the Kronecker product.

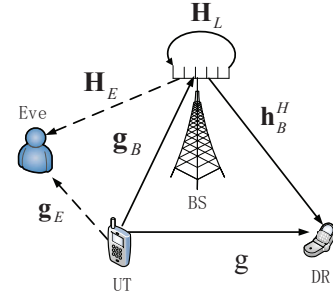


Fig. 1. A cellular system composed of a multi-antenna FD BS, a single-antenna UT, a single-antenna DR, and a multi-antenna Eve.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. Transmission Model

We consider an FD cellular system composed of an FD BS, a single-antenna legitimate DR, a single-antenna legitimate UT and a powerful multi-antenna Eve, as illustrated in Fig. 1. The UT and BS are scheduled for UL and DL transmission respectively, in the same frequency band. We assume that the Eve has N_E receiving antennas, while the FD BS has N_T transmitting antennas and N_R receiving antennas to facilitate simultaneous transmission and reception. Additionally, we assume $N_T \geq N_E + 1$, so that the BS has sufficient degrees of freedom to transmit the AN. As shown in Fig. 1, the DL channels from the BS to the DR and to the Eve are denoted by $\mathbf{h}_B^H \in \mathbb{C}^{1 \times N_T}$ and $\mathbf{H}_E \in \mathbb{C}^{N_E \times N_T}$, respectively, while the UL channels from the UT to the BS, the Eve and the DR are denoted by $\mathbf{g}_B \in \mathbb{C}^{N_R \times 1}$, $\mathbf{g}_E \in \mathbb{C}^{N_E \times 1}$, and g , respectively. Finally, $\mathbf{H}_L \in \mathbb{C}^{N_R \times N_T}$ denotes the residual self-interference channel due to the imperfect self-interference cancellation [2].

In the DL transmission, the BS transmits the confidential information-bearing signals together with the AN to protect the confidential information from wiretapping. The baseband signal from the BS can be expressed as

$$\mathbf{x}_b \triangleq \mathbf{v}s + \mathbf{n}_a, \quad (1)$$

where $s \sim \mathcal{CN}(0, 1)$ ³ is the confidential information-bearing signal, $\mathbf{v} \in \mathbb{C}^{N_T \times 1}$ denotes the beamforming vector, and $\mathbf{n}_a \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Omega})$ denotes the AN. The total available power of the BS is denoted by P_{tot} , while the design of \mathbf{v} and \mathbf{n}_a must satisfy the following power constraint:

$$\text{tr}(\mathbf{v}\mathbf{v}^H + \mathbf{\Omega}) \leq P_{\text{tot}}. \quad (2)$$

¹The eavesdropper is said to be “sophisticated/strong”, for example when it has multiple antennas or when the number of eavesdropper antennas is higher than that of each legitimate UT and DR, or alternatively, when it sends no training signals to the BS.

²In this treatise, we consider the worst-case scenario, where the UT and DR are equipped with a single antenna, while the “sophisticated/strong” eavesdropper is equipped with multiple antennas. Furthermore, when only imperfect eavesdropping channel knowledge is available to the BS. In such cases, the capacity of the eavesdropper is potentially higher than that of the legitimate user. Naturally, this scenario is more challenging than the case where the legitimate users have the same capability as the eavesdropper.

³This assumption is widely adopted in the studies on physical layer security, which facilitates the calculation of the secrecy rate [13], [14], [18], [19].

Let us assume $z \sim \mathcal{CN}(0,1)$ is the confidential signal transmitted from the UT and its transmit power is P_t . In a scheduled slot, the signals received at the BS, the DR and the Eve can be expressed as

$$y_B = \mathbf{r}^H \left(\sqrt{P_t} \mathbf{g}_B z + \mathbf{H}_L \mathbf{x}_b + \mathbf{n}_B \right), \quad (3)$$

$$y_D = \mathbf{h}_B^H \mathbf{x}_b + \sqrt{P_t} g z + n_D, \quad (4)$$

$$\mathbf{y}_E = \mathbf{H}_E \mathbf{x}_b + \sqrt{P_t} \mathbf{g}_E z + \mathbf{n}_E, \quad (5)$$

respectively, where \mathbf{n}_B , n_D , \mathbf{n}_E denote the received noise having a zero mean and unit variance/identity covariance matrix. Since the perfect state information of \mathbf{H}_L is unavailable to the BS, the maximum ratio combining (MRC) receiver, instead of the minimum mean-square error (MMSE) receiver, is assumed to be adopted at the BS for maximizing the signal-to-noise ratio (SNR) of the received confidential signal. The MRC receiver adopted is defined by $\mathbf{r} = \frac{\mathbf{g}_B}{\|\mathbf{g}_B\|_F}$, hence we have $\|\mathbf{r}\|_F = 1$.

B. Channel State Information Model

We consider the quasi-stationary flat-fading channels and assume that the UL/DL pair of channels exhibits reciprocity. At the beginning of the scheduled slot, the BS obtains the CSI of all legitimate channels via training. In practice, the DR broadcasts training sequences to facilitate the UL channel estimation, thus \mathbf{h}_B^H may be estimated at the BS and the inter-terminal interference channel g may be estimated at the UT. Then, the UT can embed g into its own training sequence and broadcast the composite sequence. As a result, both g and \mathbf{g}_B can be acquired at the BS. By contrast, the residual self-interference channel \mathbf{H}_L after adopting digital/analog domain cancellation techniques remains unknown [2]. Additionally, due to the lack of explicit cooperation between the legitimate network nodes and the Eve, perfect knowledge of the wiretap channels is difficult to obtain at the legitimate network nodes. Hence, we assume that the BS can only acquire imperfect estimates of \mathbf{H}_E and \mathbf{g}_E , and we have

$$\bar{\mathbf{H}}_E = \mathbf{H}_E + \Delta \mathbf{H}_E, \quad \bar{\mathbf{g}}_E = \mathbf{g}_E + \Delta \mathbf{g}_E, \quad (6)$$

where $\bar{\mathbf{H}}_E$ and $\bar{\mathbf{g}}_E$ denote the estimates of \mathbf{H}_E and \mathbf{g}_E , respectively, while $\Delta \mathbf{H}_E$ and $\Delta \mathbf{g}_E$ denote the corresponding estimation error [21]–[25]. This model is sufficiently generic for characterizing different types of wiretap channels. More specifically, when the Eve is constituted by an active subscriber, the perfect CSI of the wiretap channel will be available to the associated network nodes, indicating $\Delta \mathbf{H}_E = \mathbf{0}$ and $\Delta \mathbf{g}_E = \mathbf{0}$. On the other hand, when the Eve is completely passive, the BS and UT may know nothing about the CSI of the wiretap channels, implying that $\bar{\mathbf{H}}_E$ and $\bar{\mathbf{g}}_E$ can take arbitrary values.

We characterize the uncertainty associated with the CSI estimation and residual self-interference channel by the widely adopted *deterministic uncertainty model* [18], [19]. In this model, \mathbf{H}_L , $\Delta \mathbf{H}_E$, and $\Delta \mathbf{g}_E$ are bounded by the sets $\xi_{\mathbf{H}_L} \triangleq \{\mathbf{H}_L : \|\mathbf{H}_L\|_F \leq \delta_{\mathbf{H}_L}\}$, $\xi_{\mathbf{H}_E} \triangleq \{\Delta \mathbf{H}_E : \|\Delta \mathbf{H}_E\|_F \leq \delta_{\mathbf{H}_E}\}$ and $\xi_{\mathbf{g}_E} \triangleq \{\Delta \mathbf{g}_E : \|\Delta \mathbf{g}_E\|_F \leq \delta_{\mathbf{g}_E}\}$, respectively, where $\delta_{\mathbf{H}_L}$, $\delta_{\mathbf{H}_E}$ and $\delta_{\mathbf{g}_E}$ are some given constants. Relying on

this deterministic model, we optimize the system's secrecy performance under the worst-case channel conditions. This approach guarantees the absolute robustness of the system design, since the achievable secrecy performance would be no worse than that of the worst case.

C. Problem Formulation

We seek to optimize the beamformer \mathbf{v} and the covariance matrix $\mathbf{\Omega}$ of AN in order to maximize the worst-case sum secrecy rate. Defining $\mathbf{Q} = \mathbf{v}\mathbf{v}^H$, we can obtain

$$\begin{aligned} \max_{\mathbf{v}, \mathbf{\Omega} \succeq \mathbf{0}} \Delta \mathbf{H}_E \in \xi_{\mathbf{H}_E}, \mathbf{H}_L \in \xi_{\mathbf{H}_L}, \Delta \mathbf{g}_E \in \xi_{\mathbf{g}_E} & \ln(1 + \eta_1) + \ln(1 + \eta_2) \\ & - \ln \det(\mathbf{I}_{N_E} + \mathbf{Z}\mathbf{N}^{-1}) \\ \text{s.t.} & \text{rank}(\mathbf{Q}) = 1, \\ & \text{tr}(\mathbf{Q} + \mathbf{\Omega}) \leq P_{\text{tot}}, \end{aligned} \quad (7)$$

where $\eta_1 \triangleq \frac{\mathbf{h}_B^H \mathbf{Q} \mathbf{h}_B}{P_t |g|^2 + \mathbf{h}_B^H \mathbf{\Omega} \mathbf{h}_B + 1}$, $\eta_2 \triangleq \frac{P_t \|\mathbf{g}_B\|_F^2}{\mathbf{r}^H \mathbf{H}_L (\mathbf{Q} + \mathbf{\Omega}) \mathbf{H}_L^H \mathbf{r} + 1}$, $\mathbf{Z} \triangleq \mathbf{H}_E \mathbf{Q} \mathbf{H}_E^H + P_t \mathbf{g}_E \mathbf{g}_E^H$, $\mathbf{N} \triangleq \mathbf{H}_E \mathbf{\Omega} \mathbf{H}_E^H + \mathbf{I}_{N_E}$.

III. ROBUST JOINT BEAMFORMING AND AN DESIGN: AN ITERATIVE APPROACH

The max-min problem (7) is non-convex. In the following, let us transform (7) into a sequence of convex programming problems.

Employing [18, Proposition 1] and [26, Lemma 4.1], we obtain the following equivalent relation:

$$\begin{aligned} & - \ln(P_t |g|^2 + \mathbf{h}_B^H \mathbf{\Omega} \mathbf{h}_B + 1) = \\ \max_{a_1 > 0} & [-a_1 (P_t |g|^2 + \mathbf{h}_B^H \mathbf{\Omega} \mathbf{h}_B + 1) + \ln a_1 + 1], \end{aligned} \quad (8)$$

$$\begin{aligned} & - \ln[1 + \mathbf{r}^H \mathbf{H}_L (\mathbf{v}\mathbf{v}^H + \mathbf{\Omega}) \mathbf{H}_L^H \mathbf{r}] = \\ \max_{a_2 > 0} & \{-a_2 [1 + \mathbf{r}^H \mathbf{H}_L (\mathbf{v}\mathbf{v}^H + \mathbf{\Omega}) \mathbf{H}_L^H \mathbf{r}] + \ln a_2 + 1\}, \end{aligned} \quad (9)$$

$$\begin{aligned} & - \ln \det[\mathbf{I}_{N_E} + \mathbf{H}_E (\mathbf{v}\mathbf{v}^H + \mathbf{\Omega}) \mathbf{H}_E^H + P_t \mathbf{g}_E \mathbf{g}_E^H] = \\ N_E + \max_{\mathbf{W}_E \succeq \mathbf{0}} & [\ln \det(\mathbf{W}_E) - \text{tr}(\mathbf{W}_E \mathbf{X})], \end{aligned} \quad (10)$$

where $\mathbf{X} \triangleq \mathbf{I}_{N_E} + \mathbf{H}_E (\mathbf{v}\mathbf{v}^H + \mathbf{\Omega}) \mathbf{H}_E^H + P_t \mathbf{g}_E \mathbf{g}_E^H$. Then, expanding the objective function of (7) and invoking the above relationship, we have

$$\max_{\substack{a_1 > 0, a_2 > 0, \mathbf{W}_E \succeq \mathbf{0} \\ \mathbf{Q} \succeq \mathbf{0}, \mathbf{\Omega} \succeq \mathbf{0}}} \Xi(a_1, a_2, \mathbf{W}_E, \mathbf{Q}, \mathbf{\Omega}) \quad (11a)$$

$$\text{s.t.} \quad \text{rank}(\mathbf{Q}) = 1, \quad (11b)$$

$$\text{tr}(\mathbf{Q} + \mathbf{\Omega}) \leq P_{\text{tot}}, \quad (11c)$$

where the objective function (11a) is given in the equation (12).

$$\begin{aligned} \Xi(a_1, a_2, \mathbf{W}_E, \mathbf{Q}, \Omega) &\triangleq \ln(\mathbf{h}_B^H \mathbf{v} \mathbf{v}^H \mathbf{h}_B + P_t |g|^2 + \mathbf{h}_B^H \Omega \mathbf{h}_B + 1) - a_1 (P_t |g|^2 + \mathbf{h}_B^H \Omega \mathbf{h}_B + 1) + \ln a_1 + 1 + \\ &\ln(P_t \|\mathbf{g}_B\|_F^2 + \mathbf{r}^H \mathbf{H}_L (\mathbf{v} \mathbf{v}^H + \Omega) \mathbf{H}_L^H \mathbf{r} + 1) - a_2 [1 + \mathbf{r}^H \mathbf{H}_L (\mathbf{v} \mathbf{v}^H + \Omega) \mathbf{H}_L^H \mathbf{r}] + \ln a_2 + 1 - N_E + \ln \det(\mathbf{W}_E) \\ &- \text{tr}(\mathbf{W}_E \mathbf{X}) + \ln \det(\mathbf{I}_{N_E} + \mathbf{H}_E \Omega \mathbf{H}_E^H) \end{aligned} \quad (12)$$

$$\begin{aligned} F(a_1, a_2, \alpha, \beta, \gamma, \mathbf{W}_E, \mathbf{Q}, \Omega) &\triangleq \ln(1 + P_t |g|^2 + \mathbf{h}_B^H (\mathbf{Q} + \Omega) \mathbf{h}_B) - a_1 (P_t |g|^2 + \mathbf{h}_B^H \Omega \mathbf{h}_B + 1) + \ln a_1 + 1 \\ &- a_2 (1 + \alpha) + \ln a_2 + 1 + \ln(1 + P_t \|\mathbf{g}_B\|_F^2 + \alpha) + \ln \det(\mathbf{W}_E) - \text{tr}(\mathbf{W}_E) - \beta - P_t \gamma + \ln \det(\mathbf{I}_{N_E} + \mathbf{M}) \end{aligned} \quad (17)$$

$$\lambda_\beta \begin{bmatrix} \mathbf{I}_{N_T \times N_E} & \mathbf{0} \\ \mathbf{0} & -\delta_{\mathbf{H}_E}^2 \end{bmatrix} - \begin{bmatrix} (\mathbf{Q} + \Omega)^T \otimes \mathbf{W}_E & ((\mathbf{Q} + \Omega)^T \otimes \mathbf{W}_E) \text{vec}(\bar{\mathbf{H}}_E) \\ \text{vec}(\bar{\mathbf{H}}_E)^H ((\mathbf{Q} + \Omega)^T \otimes \mathbf{W}_E) & \text{vec}(\bar{\mathbf{H}}_E)^H ((\mathbf{Q} + \Omega)^T \otimes \mathbf{W}_E) \text{vec}(\bar{\mathbf{H}}_E) - \beta \end{bmatrix} \succeq \mathbf{0}, \exists \lambda_\beta \geq 0. \quad (19)$$

By introducing optimization variables α, β, γ , the problem (11) can be reformulated as

$$\begin{aligned} &\max_{\substack{a_1 > 0, a_2 > 0, \mathbf{W}_E \succeq \mathbf{0} \\ \alpha > 0, \beta > 0, \mathbf{Q} \succeq \mathbf{0}, \Omega \succeq \mathbf{0}}} F(a_1, a_2, \alpha, \beta, \gamma, \mathbf{W}_E, \mathbf{Q}, \Omega) \\ \text{s.t. } &\alpha \geq \mathbf{r}^H \mathbf{H}_L (\mathbf{Q} + \Omega) \mathbf{H}_L^H \mathbf{r}, \quad \forall \mathbf{H}_L \in \xi_{\mathbf{H}_L}; \quad (13a) \\ &\beta \geq \text{tr}(\mathbf{W}_E (\mathbf{H}_E (\mathbf{Q} + \Omega) \mathbf{H}_E^H)), \quad \forall \Delta \mathbf{H}_E \in \xi_{\Delta \mathbf{H}_E}; \quad (13b) \\ &\gamma \geq \text{tr}(\mathbf{W}_E (P_t \mathbf{g}_E \mathbf{g}_E^H)), \quad \forall \Delta \mathbf{g}_E \in \xi_{\Delta \mathbf{g}_E}; \quad (13c) \\ &\mathbf{M} \preceq \mathbf{H}_E \Omega \mathbf{H}_E^H \quad \forall \Delta \mathbf{H}_E \in \xi_{\Delta \mathbf{H}_E}; \quad (13d) \\ &\text{rank}(\mathbf{Q}) = 1; \quad (13e) \\ &\text{tr}(\mathbf{Q} + \Omega) \leq P_{\text{tot}}; \quad (13f) \end{aligned}$$

where the objective function $F(a_1, a_2, \alpha, \beta, \gamma, \mathbf{W}_E, \mathbf{Q}, \Omega)$ is defined in (17).

Let us first transform the constraint (13a) into a more convenient formulation:

$$\begin{aligned} \mathbf{r}^H \mathbf{H}_L (\mathbf{Q} + \Omega) \mathbf{H}_L^H \mathbf{r} &= \text{tr}(\mathbf{r} \mathbf{r}^H \mathbf{H}_L (\mathbf{Q} + \Omega) \mathbf{H}_L^H) \\ &\stackrel{(a)}{=} (\text{vec}(\mathbf{H}_L))^H ((\mathbf{Q} + \Omega)^T \otimes \mathbf{r} \mathbf{r}^H) \text{vec}(\mathbf{H}_L), \end{aligned} \quad (14)$$

where the equation (a) is due to the following equation:

$$\text{tr}(\mathbf{ABCD}) = (\text{vec}(\mathbf{D}^T))^T (\mathbf{C}^T \otimes \mathbf{A}) \text{vec}(\mathbf{B}). \quad (15)$$

With similar procedures, the constraint (13b) - (13c) can be transformed into a more convenient formulation, and the non-

convex problem (7) can be equivalently reformulated as:

$$\begin{aligned} &\max_{\substack{a_1 > 0, a_2 > 0, \mathbf{W}_E \succeq \mathbf{0} \\ \alpha > 0, \beta > 0, \mathbf{Q} \succeq \mathbf{0}, \Omega \succeq \mathbf{0}}} F(a_1, a_2, \alpha, \beta, \gamma, \mathbf{W}_E, \mathbf{Q}, \Omega) \\ \text{s.t. } &\alpha \geq (\text{vec}(\mathbf{H}_L))^H ((\mathbf{Q} + \Omega)^T \otimes \mathbf{r} \mathbf{r}^H) \text{vec}(\mathbf{H}_L), \\ &\forall \mathbf{H}_L \in \xi_{\mathbf{H}_L}; \quad (16a) \\ &\beta \geq (\text{vec}(\Delta \mathbf{H}_E))^H ((\mathbf{Q} + \Omega)^T \otimes \mathbf{W}_E^H) \text{vec}(\Delta \mathbf{H}_E) + \\ &2\text{Re}((\text{vec}(\Delta \mathbf{H}_E))^H ((\mathbf{Q} + \Omega)^T \otimes \mathbf{W}_E^H) \text{vec}(\bar{\mathbf{H}}_E)) + \\ &\text{vec}(\bar{\mathbf{H}}_E)^H ((\mathbf{Q} + \Omega)^T \otimes \mathbf{W}_E^H) \text{vec}(\bar{\mathbf{H}}_E), \\ &\forall \Delta \mathbf{H}_E \in \xi_{\Delta \mathbf{H}_E}; \quad (16b) \\ &\gamma \geq \Delta \mathbf{g}_E^H \mathbf{W}_E \Delta \mathbf{g}_E + 2\text{Re}(\Delta \mathbf{g}_E^H \mathbf{W}_E \bar{\mathbf{g}}_E) + \bar{\mathbf{g}}_E^H \mathbf{W}_E \bar{\mathbf{g}}_E, \\ &\forall \Delta \mathbf{g}_E \in \xi_{\Delta \mathbf{g}_E}; \quad (16c) \\ &\mathbf{M} \preceq \Delta \mathbf{H}_E \Omega \Delta \mathbf{H}_E^H + \Delta \mathbf{H}_E \Omega \bar{\mathbf{H}}_E^H + \bar{\mathbf{H}}_E \Omega \Delta \mathbf{H}_E^H \\ &+ \bar{\mathbf{H}}_E \Omega \bar{\mathbf{H}}_E^H, \forall \Delta \mathbf{H}_E \in \xi_{\Delta \mathbf{H}_E}; \quad (16d) \\ &\text{rank}(\mathbf{Q}) = 1; \quad (16e) \\ &\text{tr}(\mathbf{Q} + \Omega) \leq P_{\text{tot}}. \quad (16f) \end{aligned}$$

The resultant problem (16) still remains challenging to solve. To obtain a more tractable problem, the following reformulation is carried out.

According to the eigenvalue equation, the constraint (16a) can be rewritten as

$$\alpha \geq \delta_{\mathbf{H}_L}^2 \lambda_{\max}((\mathbf{Q} + \Omega)^T \otimes \mathbf{r} \mathbf{r}^H), \quad (18)$$

where λ_{\max} is the maximal eigenvalue of $(\mathbf{Q} + \Omega)^T \otimes \mathbf{r} \mathbf{r}^H$.

According to the S-Procedure [28], the constraints (16b) and (16c) hold if and only if there exist $\lambda_\beta \geq 0$ and $\lambda_\gamma \geq 0$ such that (19) and (20) hold:

$$\begin{aligned} &\lambda_\gamma \begin{bmatrix} \mathbf{I}_{N_T \times N_E} & \mathbf{0} \\ \mathbf{0} & -\delta_{\mathbf{g}_E}^2 \end{bmatrix} - \begin{bmatrix} \mathbf{W}_E & \mathbf{W}_E \bar{\mathbf{g}}_E \\ \bar{\mathbf{g}}_E^H \mathbf{W}_E & \bar{\mathbf{g}}_E^H \mathbf{W}_E \bar{\mathbf{g}}_E - \gamma \end{bmatrix} \succeq \mathbf{0}, \\ &\exists \lambda_\gamma \geq 0. \end{aligned} \quad (20)$$

Then, invoking the robust quadratic matrix inequality of [29], the constraint (16d) holds if and only if there exists a value $\lambda_M \geq 0$ such that

$$\begin{bmatrix} \bar{\mathbf{H}}_E \boldsymbol{\Omega} \bar{\mathbf{H}}_E^H - \mathbf{M} & \bar{\mathbf{H}}_E \boldsymbol{\Omega} \\ \boldsymbol{\Omega} \bar{\mathbf{H}}_E^H & \boldsymbol{\Omega} \end{bmatrix} - \lambda_M \begin{bmatrix} \mathbf{I}_{N_E} & \mathbf{0} \\ \mathbf{0} & -\frac{1}{\delta_{\mathbf{H}_E}^2} \end{bmatrix} \succeq \mathbf{0},$$

$\exists \lambda_M \geq 0.$ (21)

Finally, invoking the classic semidefinite relaxation (SDR) technique [5], [27], we drop the constraint (16e) to obtain the rank-relaxation version of (16).

As a result, the robust joint design of the confidential signal's beamforming vector \mathbf{v} and the covariance matrix $\boldsymbol{\Omega}$ of the AN can be reformulated as

$$\begin{aligned} & \max_{\substack{a_1 > 0, a_2 > 0, \mathbf{W}_E \succeq \mathbf{0} \\ \alpha > 0, \beta > 0, \mathbf{Q} \succeq \mathbf{0}, \boldsymbol{\Omega} \succeq \mathbf{0}}} F(a_1, a_2, \alpha, \beta, \gamma, \mathbf{W}_E, \mathbf{Q}, \boldsymbol{\Omega}) \\ \text{s.t.} & \quad (18), (19), (20), (21) \text{ and } (16f). \end{aligned} \quad (22)$$

Now, although the problem (22) itself still remains non-convex and its global optimum is difficult to obtain, it can be decomposed into two sub-problems that are convex. Then, we can adopt the efficient BCD algorithm of [20, Section 2.7] to find a sequence of locally optimal solutions of (22). More specifically, in this method, all optimization variables are decoupled into two blocks: $\{\alpha, \beta, \mathbf{Q}, \boldsymbol{\Omega}\}$ and $\{a_1, a_2, \mathbf{W}_E\}$. Note that when the block $\{a_1, a_2, \mathbf{W}_E\}$ is fixed, the problem (22) becomes convex with respect to $\{\alpha, \beta, \mathbf{Q}, \boldsymbol{\Omega}\}$, and when the block $\{\alpha, \beta, \mathbf{Q}, \boldsymbol{\Omega}\}$ is fixed, the problem (22) becomes convex with respect to $\{a_1, a_2, \mathbf{W}_E\}$. Hence, the two blocks of variables can be optimized in turn with low complexity by fixing one block and optimizing the other.⁴

The BCD algorithm employed is summarized in Algorithm 1. Since Algorithm 1 yields non-descending objective function values, it must converge subject to the plausible constraint that physically the secrecy rate is finite.

Given Algorithm 1, the optimal \mathbf{Q} can be found by solving (22). However, there is no guarantee that $\text{rank}(\mathbf{Q}) = 1$ and that \mathbf{v} can be derived from \mathbf{Q} without any performance deterioration. In the following, we propose a simple method of constructing \mathbf{v} from \mathbf{Q} with good guaranteed performance, as summarized in Algorithm 2.

Proposition 1. *The secrecy performance achieved by $\hat{\mathbf{Q}}$ constructed in Algorithm 2 is not inferior to the one achieved by \mathbf{Q} .*

Proof. The objective function value of the problem (7) represents the achievable secrecy rate performance. Then, let us replace $\mathbf{v}\mathbf{v}^H$ by $\hat{\mathbf{Q}}$ and \mathbf{Q} , respectively, in the problem (7) to give the proof.

Firstly, we have the following linear matrix inequality

$$\mathbf{Q} - \hat{\mathbf{Q}} = \mathbf{Q}^{1/2} \left(\mathbf{I} - \frac{\mathbf{Q}^{1/2} \mathbf{h}_B \mathbf{h}_B^H \mathbf{Q}^{1/2}}{\mathbf{h}_B^H \mathbf{Q} \mathbf{h}_B} \right) \mathbf{Q}^{1/2} \succeq \mathbf{0}. \quad (23)$$

⁴As pointed out in [31], the combination of substantially increased computing power, sophisticated algorithms, and new coding approaches have jointly made it possible to solve modest-sized convex optimization problems on microsecond or millisecond time scales and with strict completion deadlines. This enables real-time convex optimization in signal processing. In fact, this time-scale is also similar to a single transmission time interval (TTI) in 4G LTE (one millisecond) and in 5G NR (a fraction of one millisecond).

Algorithm 1 The BCD algorithm for solving the robust joint optimization problem (22).

Set $l = 1$ and perform initialization with arbitrary feasible $\bar{a}_1(l), \bar{a}_2(l), \bar{\mathbf{W}}_E(l)$.

while the difference between the values of $F(\bar{a}_1(l), \bar{a}_2(l), \bar{\alpha}(l), \bar{\beta}(l), \bar{\gamma}(l), \bar{\mathbf{W}}_E(l), \bar{\mathbf{Q}}(l), \bar{\boldsymbol{\Omega}}(l))$ in successive iterations is larger than ϕ for some $\phi > 0$, **do**

1. Solve the problem (22) by fixing $a_1 = \bar{a}_1(l), a_2 = \bar{a}_2(l), \mathbf{W}_E = \bar{\mathbf{W}}_E(l)$, and obtain the global optimum, i.e., $\alpha^*, \beta^*, \gamma^*, \mathbf{Q}^*, \boldsymbol{\Omega}^*$.
2. Set $\bar{\alpha}(l) = \alpha^*, \bar{\beta}(l) = \beta^*, \bar{\gamma}(l) = \gamma^*, \bar{\mathbf{Q}}(l) = \mathbf{Q}^*$, and $\bar{\boldsymbol{\Omega}}(l) = \boldsymbol{\Omega}^*$.
3. Solve the problem (22) by fixing $\alpha = \bar{\alpha}(l), \beta = \bar{\beta}(l), \gamma = \bar{\gamma}(l), \mathbf{Q} = \bar{\mathbf{Q}}(l), \boldsymbol{\Omega} = \bar{\boldsymbol{\Omega}}(l)$ and obtain the global optimum, i.e., $a_1^*, a_2^*, \mathbf{W}_E^*$.
4. $l = l + 1$,
5. Set $\bar{a}_1(l) = a_1^*, \bar{a}_2(l) = a_2^*$, and $\bar{\mathbf{W}}_E(l) = \mathbf{W}_E^*$.

end while

Output: $\mathbf{Q}, \boldsymbol{\Omega}$.

Algorithm 2 The proposed BCD algorithm for solving the robust joint optimization problem (22).

1. Set $\hat{\mathbf{Q}} = \frac{\mathbf{Q} \mathbf{h}_B \mathbf{h}_B^H \mathbf{Q}}{\mathbf{h}_B^H \mathbf{Q} \mathbf{h}_B}$.

2. \mathbf{v} is generated from $\hat{\mathbf{Q}}$ by singular value decomposition.
-

Thus, $\text{tr}(\mathbf{Q}) \geq \text{tr}(\hat{\mathbf{Q}})$. Then, we can conclude that if \mathbf{Q} satisfies the power constraint of the problem (7), $\hat{\mathbf{Q}}$ should also satisfy this power constraint.

Next, let us check the objective function of the problem (7), where $\mathbf{v}\mathbf{v}^H$ appears in the numerator of η_1 , in the denominator of η_2 , as well as in \mathbf{Z} . Then, for establishing the proof, the following inequalities must be proved:

$$\mathbf{h}_B^H \mathbf{Q} \mathbf{h}_B \leq \mathbf{h}_B^H \hat{\mathbf{Q}} \mathbf{h}_B, \quad (24)$$

$$\mathbf{r}^H \mathbf{H}_L \mathbf{Q} \mathbf{H}_L^H \mathbf{r} \geq \mathbf{r}^H \mathbf{H}_L \hat{\mathbf{Q}} \mathbf{H}_L^H \mathbf{r}, \quad (25)$$

$$\mathbf{H}_E \mathbf{Q} \mathbf{H}_E^H \succeq \mathbf{H}_E \hat{\mathbf{Q}} \mathbf{H}_E^H. \quad (26)$$

Observe that (24) holds true, since $\mathbf{h}_B^H \hat{\mathbf{Q}} \mathbf{h}_B = \mathbf{h}_B^H \mathbf{Q} \mathbf{h}_B$. Additionally, since $\mathbf{Q} - \hat{\mathbf{Q}} \succeq \mathbf{0}$, (25) and (26) also hold true. \square

The following proposition shows that the limit points generated by Algorithm 1 would satisfy the KKT condition of the problem (7).

Proposition 2. *The limit point of the sequence $\{\bar{a}_1(l), \bar{a}_2(l), \bar{\alpha}(l), \bar{\beta}(l), \bar{\gamma}(l), \bar{\mathbf{W}}_E(l), \bar{\mathbf{Q}}(l), \bar{\boldsymbol{\Omega}}(l)\}$ generated by Algorithm 1, i.e., $\{\bar{a}_1^*, \bar{a}_2^*, \bar{\alpha}^*, \bar{\beta}^*, \bar{\gamma}^*, \bar{\mathbf{W}}_E^*, \bar{\mathbf{Q}}^*, \bar{\boldsymbol{\Omega}}^*\}$ is a KKT point of the nonconvex optimization problem (7).*

Proof. The proof is given in Appendix B. \square

IV. SIMULATION RESULTS

In our simulations, we assume that all the available channel coefficients obey $\mathcal{CN}(0, 1)$. The transmit power of the UT is set to $P_t = 20$ dBm and the received noise power is normalized to 0 dBm. The simulation results were obtained by averaging over 200 independent trials. With the obtained \mathbf{v} and $\boldsymbol{\Omega}$, the achievable worst-case sum secrecy rate R_w can

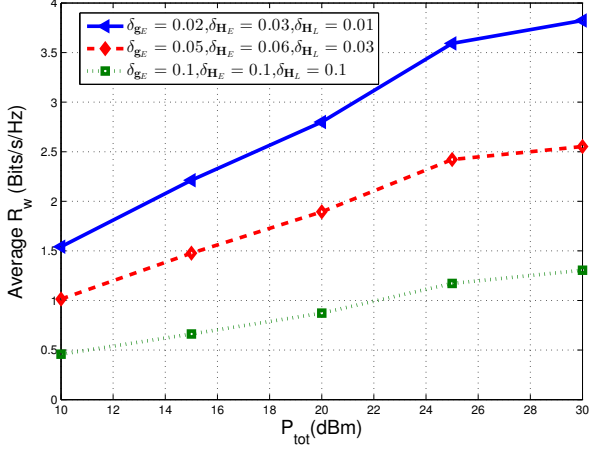


Fig. 2. The average worst-case sum secrecy rate versus P_{tot} under $N_T = 3, N_R = 2, N_E = 2$, and different values of error bounds.

be computed from (33) of Appendix A. Note that mathematically it is possible for the sum secrecy rate R_w to be negative when the transmit power P_{tot} is low, but in the real world, the achievable secrecy rate cannot be negative, since the information wiretapped by the eavesdropper cannot be higher than that transmitted by the legitimate transmitter. Hence, when we encounter negative secrecy sum rate in our simulations, we have to use $\max(R_w, 0)$ to ensure that its lowest value is zero. Assuming $N_T = 3, N_R = 2$ and $N_E = 2$, in Fig. 2 we show the achievable average R_w versus the total available transmit power P_{tot} of the FD BS, subject to the error bounds δ_{H_E} , δ_{H_L} and δ_{g_E} . When P_{tot} increases, the strength of both the confidential signals and the AN also increases. Therefore, the average R_w increases with P_{tot} , which has been validated by our simulation results in Fig. 2. On the other hand, as the error bounds increase, the CSI estimation of \mathbf{H}_E , \mathbf{H}_L and \mathbf{g}_E becomes more and more inaccurate. Therefore, both the information leakage to the Eve and the jamming signal (i.e., the AN) leakage to the BS increase with the error bounds, which results in a secrecy performance degradation. This argument has been validated by our simulation results in Fig. 2, showing that smaller error bounds result in a larger average R_w .

For illustrating the robustness of the proposed secrecy transmission scheme to imperfect CSI, we show in Fig. 3 how the achievable average R_w changes upon increasing the error bounds. For simplicity, we assume $\delta_{H_E} = \delta_{H_L} = \delta_{g_E} = \epsilon$, $P_{tot} = 10$ dBm and $N_R = 2$. We observe that the achievable average R_w decreases upon increasing ϵ , and the secrecy performance degradation becomes significant when ϵ increases from 0.01 to 0.09. Additionally, we see that a larger N_E results in a smaller average R_w . This is because when N_E becomes larger, the wiretapping capability of the Eve becomes stronger, and the achievable secrecy performance is degraded.

In Fig. 4 we characterize the convergence behaviour of Algorithm 1 subject to different values of P_{tot} . It is observed that the number of iterations required by Algorithm 1 increases when P_{tot} becomes larger. This is because upon increasing

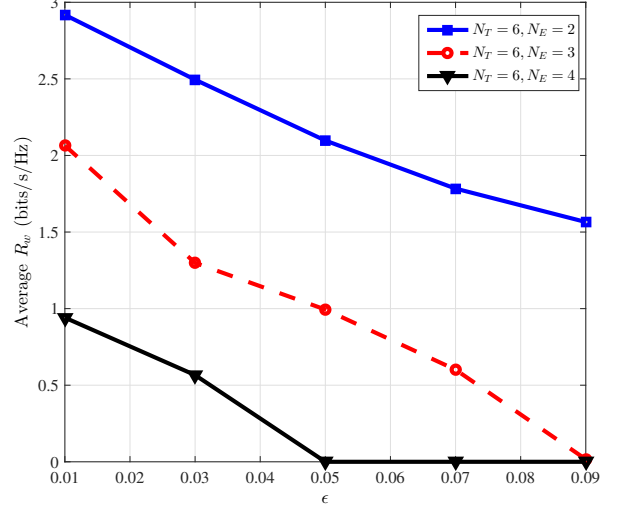


Fig. 3. The average worst-case sum secrecy rate versus the error bound ϵ under $N_R = 2, P_{tot} = 10$ dBm, and different values of N_E .

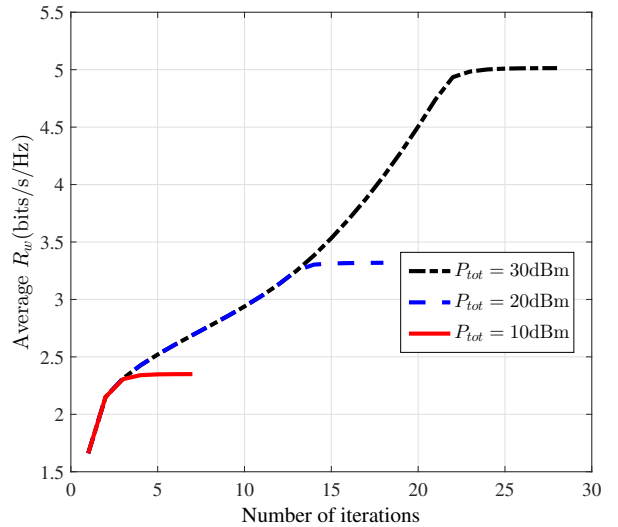


Fig. 4. The average worst-case sum secrecy rate versus P_{tot} under $N_T = 4, N_R = 2, N_E = 2, P_t = 5$ dBm.

P_{tot} , the feasible set of the problem (22) is expanded. Additionally, we see that Algorithm 1 indeed converges in all of our observations recorded.

For characterizing the performance of our proposed algorithm, we use projection matrix theory to construct a benchmarker (termed as "Projection Method"), and use the analysis results derived in Appendix A to quantify its worst-case secrecy rate. Since the CSI of the eavesdropping channel is imperfect, the precoder \mathbf{v} is constructed by projecting \mathbf{h}_B onto the null-space of the known part of the CSI of the eavesdropping channel, i.e., $\bar{\mathbf{H}}_E$. Explicitly,

$$\mathbf{v} = P_t \frac{\mathbf{U}_E \mathbf{U}_E^H \mathbf{h}_B}{\|\mathbf{U}_E \mathbf{U}_E^H \mathbf{h}_B\|_F}, \quad (27)$$

where, P_t is the transmit power of the information signal, and

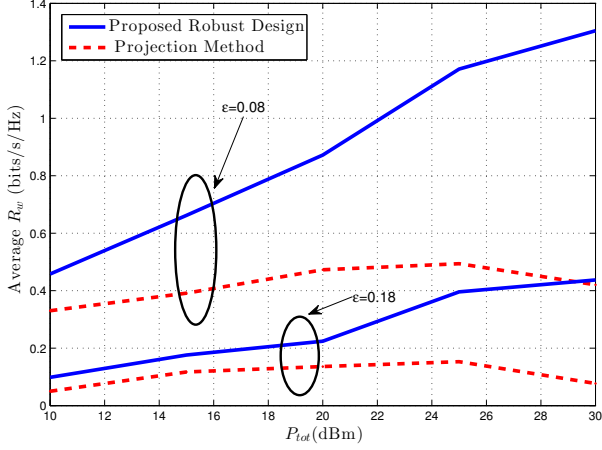


Fig. 5. Performance comparison between the proposed robust strategy and the projection method under $N_T = 3$, $N_R = 2$, $N_E = 2$ dBm.

\mathbf{U}_E is the null-space of $\bar{\mathbf{H}}_E$. For designing AN, since the self-interference channel \mathbf{H}_L and the eavesdropping channel \mathbf{H}_E are both imperfect known, we construct the artificial noise in the null-space of \mathbf{h}_B and allocate the power of the artificial noise equally. Explicitly,

$$\mathbf{\Omega} = \frac{P_N \mathbf{U}_{\mathbf{h}_B} \mathbf{U}_{\mathbf{h}_B}^H}{N_T - 1}, \quad (28)$$

where, P_N is the total power of the artificial noise and $\mathbf{U}_{\mathbf{h}_B}$ is the null-space of \mathbf{h}_B . We allocate the transmit power equally between the information signal and AN, i.e., $P_t = P_N = \frac{P_{tot}}{2}$.

Fig. 5 show the secrecy performance achieved both by our proposed robust strategy and by the "Projection Method". Based on our simulation results, we can find that the secrecy performance achieved by the proposed robust strategy is much better than that of the "Projection Method", and the gains become higher upon increasing P_{tot} , which has confirmed the efficiency of our proposed robust strategy.

V. CONCLUSIONS

In this paper, we investigated the worst-case physical layer security of a FD system composed of a FD multi-antenna BS, a "sophisticated/strong" multi-antenna eavesdropper, a single-antenna UT and a single-antenna DR. Assuming that only imperfect CSI of the eavesdropping and residual self-interference channels is available to the FD BS, we have proposed a robust secrecy transmission scheme, where the AN and beamforming vector are jointly designed for securing the confidential signals transmitted from the BS and UT. By employing the BCD algorithm and linear matrix inequality, we transform the worst-case non-convex problem into a sequence of convex problems in order to find its locally optimal solution. For evaluating the achievable performance, the analysis result of the worst-case secrecy rate is derived. Furthermore, by employing the projection matrix theory, we construct another secrecy transmission scheme for the performance comparison. Our simulation results have validated the effectiveness of the proposed secrecy transmission scheme and provided valuable

insights into how several relevant design parameters affect the achievable secrecy performance of the system considered.

APPENDIX A ACHIEVABLE WORST-CASE SUM SECRECY RATE CALCULATION

Invoking Algorithm 1 and Algorithm 2, we obtain the optimized rank-1 \mathbf{Q} and $\mathbf{\Omega}$. Then, the achievable worst-case sum secrecy rate R_w can be calculated from the objective function of the problem (7). Firstly, the minimum of η_2 , i.e., η_2^{\min} can be calculated as $\eta_2^{\min} = \frac{P_t \|\mathbf{g}_E\|_F^2}{\delta_{\mathbf{H}_L}^2 \lambda_{\max}((\mathbf{Q} + \mathbf{\Omega})^T \otimes \mathbf{r}\mathbf{r}^H) + 1}$.

Then, we calculate the maximum of $\ln \det(\mathbf{I}_{N_E} + \mathbf{Z}\mathbf{N}^{-1})$, which can be expanded as

$$\begin{aligned} & \ln \det(\mathbf{I}_{N_E} + \mathbf{H}_E \mathbf{Q} \mathbf{H}_E^H (\mathbf{H}_E \mathbf{\Omega} \mathbf{H}_E^H + \mathbf{I}_{N_E})^{-1}) - \\ & \ln \left(1 + \text{tr} \left(P_t \mathbf{g}_E \mathbf{g}_E^H (\mathbf{H}_E (\mathbf{\Omega} + \mathbf{Q}) \mathbf{H}_E^H + \mathbf{I}_{N_E})^{-1} \right) \right) \stackrel{(a)}{=} \\ & \ln \left(1 + \text{tr} \left(\mathbf{H}_E \mathbf{Q} \mathbf{H}_E^H (\mathbf{H}_E \mathbf{\Omega} \mathbf{H}_E^H + \mathbf{I}_{N_E})^{-1} \right) \right) - \\ & \ln \left(1 + \text{tr} \left(P_t \mathbf{g}_E \mathbf{g}_E^H (\mathbf{H}_E (\mathbf{\Omega} + \mathbf{Q}) \mathbf{H}_E^H + \mathbf{I}_{N_E})^{-1} \right) \right). \end{aligned}$$

Step (a) holds true, since $\text{rank}(\mathbf{Q}) = 1$.

Then, the upper bound of $\ln \det(\mathbf{I}_{N_E} + \mathbf{Z}\mathbf{N}^{-1})$ is given by $\ln(1 + \theta_1^*) + \ln(1 + \theta_2^*)$, where

$$\begin{aligned} \theta_1^* &= \arg \min_{\theta_1} \theta_1 \\ & \text{s.t. } \theta_1 \geq \text{tr} \left(\mathbf{H}_E \mathbf{Q} \mathbf{H}_E^H (\mathbf{H}_E \mathbf{\Omega} \mathbf{H}_E^H + \mathbf{I}_{N_E})^{-1} \right), \\ & \quad \forall \Delta \mathbf{H}_E \in \xi_{\mathbf{H}_E}; \\ \theta_2^* &= \arg \min_{\theta_2} \theta_2 \\ & \text{s.t. } \theta_2 \geq P_t \text{tr} \left(\mathbf{g}_E \mathbf{g}_E^H (\mathbf{H}_E (\mathbf{\Omega} + \mathbf{Q}) \mathbf{H}_E^H + \mathbf{I}_{N_E})^{-1} \right), \\ & \quad \forall \mathbf{g}_E \in \xi_{\mathbf{g}_E}, \forall \Delta \mathbf{H}_E \in \xi_{\mathbf{H}_E}. \end{aligned} \quad (29)$$

After some further manipulations, (29) can be reformulated as the following semidefinite programming (SDP) problem:

$$\begin{aligned} & \min_{\theta_1} \theta_1 \\ & \text{s.t. } \begin{bmatrix} \bar{\mathbf{H}}_E (\theta_1 \mathbf{\Omega} - \mathbf{Q}) \bar{\mathbf{H}}_E^H & \bar{\mathbf{H}}_E (\theta_1 \mathbf{\Omega} - \mathbf{Q}) \\ (\theta_1 \mathbf{\Omega} - \mathbf{Q}) \bar{\mathbf{H}}_E^H & \theta_1 \mathbf{\Omega} - \mathbf{Q} \end{bmatrix} - t \Phi \succeq \mathbf{0}, \\ & \quad \exists t \geq 0, \end{aligned} \quad (31)$$

where we have $\Phi \triangleq \begin{bmatrix} \mathbf{I}_{N_E} & \mathbf{0} \\ \mathbf{0} & -\frac{1}{\delta_{\mathbf{H}_E}^2} \end{bmatrix}$.

Similarly, after some manipulations, (30) can be reformulated as the following SDP problem:

$$\begin{aligned} & \min_{\theta_2} \theta_2 \\ & \text{s.t. } \begin{bmatrix} \theta_2 \bar{\mathbf{H}}_E (\mathbf{\Omega} + \mathbf{Q}) \bar{\mathbf{H}}_E^H + \mathbf{I}_{N_E} - P_t \hat{\mathbf{g}}_E \hat{\mathbf{g}}_E^H & \theta_2 \bar{\mathbf{H}}_E (\mathbf{\Omega} + \mathbf{Q}) \\ \theta_2 (\mathbf{\Omega} + \mathbf{Q}) \bar{\mathbf{H}}_E^H & \theta_2 (\mathbf{\Omega} + \mathbf{Q}) \end{bmatrix} \\ & \quad - \eta \begin{bmatrix} \mathbf{I}_{N_E} & \mathbf{0} \\ \mathbf{0} & -\frac{1}{\delta_{\mathbf{H}_E}^2} \end{bmatrix} \succeq \mathbf{0}, \\ & \quad \exists \eta \geq 0, \end{aligned} \quad (32)$$

where $\hat{\mathbf{g}}_E \triangleq \bar{\mathbf{g}}_E + \frac{\bar{\mathbf{g}}_E}{\|\bar{\mathbf{g}}_E\|_F} \delta_{\mathbf{g}_E}$. R_w is then calculated by

$$R_w = \ln(1 + \eta_1) + \ln(1 + \eta_2^{\min}) - \ln[(1 + \theta_1^*)(1 + \theta_2^*)]. \quad (33)$$

APPENDIX B
PROOF OF PROPOSITION 2

For clarity, [30, Corollary 2] is introduced, which is given as follows.

Corollary 1. *Consider the problem*

$$\min_{\mathbf{X}, a} f(\mathbf{X}, a), \quad s.t. (\mathbf{X}, a) \in \mathcal{X} \times \mathcal{A}, \quad (34)$$

where $f(\mathbf{X}, a)$ is a continuously differentiable function; $\mathcal{X} \subseteq \mathbb{C}^{m \times n}$ and $\mathcal{A} \subseteq \mathbb{R}$ are closed, nonempty, and convex subsets. Suppose that the sequence $\{(\mathbf{X}^{(n)}, a^{(n)})\}$ generated by optimizing \mathbf{X} and a alternatively has limit points. Every limit point of $\{(\mathbf{X}^{(n)}, a^{(n)})\}$ is a stationary point of the problem.

We have used matrix inequality for transforming the problem (11) into the problem (22), equivalently. Therefore, if we prove that the stationary point of the problem (11) is the KKT point of the problem (7), and the same as the stationary point of the problem (22). The objective function of the problem (22) is continuously differentiable, and the feasible sets are closed and convex. Since the objective function of the problem (22) is nondecreasing in Algorithm 1, due to the power constraint, by optimizing the two variable sets: (a_1, a_2, \mathbf{W}_E) and $(\alpha, \beta, \gamma, \mathbf{Q}, \mathbf{\Omega})$ alternately, two variable sets have limit points. Using Bolzano-Weierstrass theorem, we can conclude that $\{\bar{a}_1(l), \bar{a}_2(l), \bar{\alpha}(l), \bar{\beta}(l), \bar{\gamma}(l), \bar{\mathbf{W}}_E(l), \bar{\mathbf{Q}}(l), \bar{\mathbf{\Omega}}(l)\}$ has limit points. Therefore, with [30, Corollary 2], the limit points generated by Algorithm 1, i.e., $\{\bar{a}_1^*, \bar{a}_2^*, \bar{\alpha}^*, \bar{\beta}^*, \bar{\gamma}^*, \bar{\mathbf{W}}_E^*, \bar{\mathbf{Q}}^*, \bar{\mathbf{\Omega}}^*\}$ is a stationary point of the problem (22), and $\{\bar{a}_1^*, \bar{a}_2^*, \bar{\mathbf{W}}_E^*, \bar{\mathbf{Q}}^*, \bar{\mathbf{\Omega}}^*\}$ is a stationary point of the problem (11).

Next, we will show that the stationary point of the problem (11) is a KKT point of the problem (7).

Since with Algorithm 2, we can construct \mathbf{Q} whose has one-rank without performance deterioration, the constraint rank $(\mathbf{Q}) = 1$ of the problem (7) and (11) can be omitted. For brevity, we denote the objective function of the problem (7) as $\varphi(\mathbf{Q}, \mathbf{\Omega})$. Since $\{\bar{a}_1^*, \bar{a}_2^*, \bar{\alpha}^*, \bar{\beta}^*, \bar{\gamma}^*, \bar{\mathbf{W}}_E^*, \bar{\mathbf{Q}}^*, \bar{\mathbf{\Omega}}^*\}$ is a stationary point of the problem (11), we have

$$\begin{aligned} \text{tr} \left(\nabla_{\mathbf{Q}} \Xi (a_1^*, a_2^*, \mathbf{W}_E^*, \mathbf{Q}^*, \mathbf{\Omega}^*)^H (\mathbf{Q} - \mathbf{Q}^*) \right) &\leq 0, \\ \text{tr} (\mathbf{Q} + \mathbf{\Omega}) &\leq P_{tot} \\ \text{tr} \left(\nabla_{\mathbf{\Omega}} \Xi (a_1^*, a_2^*, \mathbf{W}_E^*, \mathbf{Q}^*, \mathbf{\Omega}^*)^H (\mathbf{\Omega} - \mathbf{\Omega}^*) \right) &\leq 0, \\ \text{tr} (\mathbf{Q} + \mathbf{\Omega}) &\leq P_{tot} \\ \text{tr} \left(\nabla_{a_1} \Xi (a_1^*, a_2^*, \mathbf{W}_E^*, \mathbf{Q}^*, \mathbf{\Omega}^*)^H (a_1 - a_1^*) \right) &\leq 0, a_1 > 0 \\ \text{tr} \left(\nabla_{a_2} \Xi (a_1^*, a_2^*, \mathbf{W}_E^*, \mathbf{Q}^*, \mathbf{\Omega}^*)^H (a_2 - a_2^*) \right) &\leq 0, a_2 > 0 \end{aligned} \quad (35)$$

From the equivalent relation (8)-(10), we can obtain

$$\mathbf{W}_E^* = (\mathbf{I}_{N_E} + \mathbf{H}_E (\mathbf{v}\mathbf{v}^H + \mathbf{\Omega}) \mathbf{H}_E^H + P_t \mathbf{g}_E \mathbf{g}_E^H)^{-1} \quad (36)$$

$$a_1 = (P_t |g|^2 + \mathbf{h}_B^H \mathbf{\Omega} \mathbf{h}_B + 1)^{-1} \quad (37)$$

$$a_2 = (1 + \mathbf{r}^H \mathbf{H}_L (\mathbf{v}\mathbf{v}^H + \mathbf{\Omega}) \mathbf{H}_L^H \mathbf{r})^{-1} \quad (38)$$

Substituting (36)-(38) into (35), we can obtain

$$\begin{aligned} \nabla_{\mathbf{Q}} \Xi (a_1^*, a_2^*, \mathbf{W}_E^*, \mathbf{Q}^*, \mathbf{\Omega}^*) &= \nabla_{\mathbf{Q}} \varphi (\mathbf{Q}^*, \mathbf{\Omega}^*) \\ \nabla_{\mathbf{\Omega}} \Xi (a_1^*, a_2^*, \mathbf{W}_E^*, \mathbf{Q}^*, \mathbf{\Omega}^*) &= \nabla_{\mathbf{\Omega}} \varphi (\mathbf{Q}^*, \mathbf{\Omega}^*). \end{aligned} \quad (39)$$

Therefore, with the equation above, we can further obtain

$$\begin{aligned} \text{tr} \left(\nabla_{\mathbf{Q}} \varphi (\mathbf{Q}^*, \mathbf{\Omega}^*)^H (\mathbf{Q} - \mathbf{Q}^*) \right) &\leq 0, \text{tr} (\mathbf{Q} + \mathbf{\Omega}) \leq P_{tot} \\ \text{tr} \left(\nabla_{\mathbf{\Omega}} \varphi (\mathbf{Q}^*, \mathbf{\Omega}^*)^H (\mathbf{\Omega} - \mathbf{\Omega}^*) \right) &\leq 0, \text{tr} (\mathbf{Q} + \mathbf{\Omega}) \leq P_{tot}. \end{aligned} \quad (40)$$

Then, we can conclude that $\mathbf{Q}^*, \mathbf{\Omega}^*$ is the optimal solution of the following problem:

$$\max_{\mathbf{Q}, \mathbf{\Omega}} \text{tr} \left(\nabla_{\mathbf{Q}} \varphi (\mathbf{Q}^*, \mathbf{\Omega}^*)^H (\mathbf{Q} - \mathbf{Q}^*) \right), \quad s.t. \text{tr} (\mathbf{Q} + \mathbf{\Omega}) \leq P_{tot} \quad (41)$$

Hence, $\mathbf{Q}^*, \mathbf{\Omega}^*$ should satisfy the KKT conditions of the problem (41), i.e.,

$$\begin{aligned} \nabla_{\mathbf{Q}} \varphi (\mathbf{Q}^*, \mathbf{\Omega}^*) - \lambda_{\mathbf{Q}} \mathbf{I} + \mathbf{D}_{\mathbf{Q}} &= \mathbf{0} \\ \mathbf{Q}^* \mathbf{D}_{\mathbf{Q}} &= \mathbf{0}, \lambda_{\mathbf{Q}} \geq 0, \mathbf{D}_{\mathbf{Q}} \succeq \mathbf{0} \\ \nabla_{\mathbf{\Omega}} \varphi (\mathbf{Q}^*, \mathbf{\Omega}^*) - \lambda_{\mathbf{\Omega}} \mathbf{I} + \mathbf{D}_{\mathbf{\Omega}} &= \mathbf{0} \\ \mathbf{Q}^* \mathbf{D}_{\mathbf{\Omega}} &= \mathbf{0}, \lambda_{\mathbf{\Omega}} \geq 0, \mathbf{D}_{\mathbf{\Omega}} \succeq \mathbf{0}, \end{aligned} \quad (42)$$

where $\mathbf{D}_{\mathbf{\Omega}}$, $\lambda_{\mathbf{\Omega}}$, $\mathbf{D}_{\mathbf{Q}}$, and $\lambda_{\mathbf{Q}}$ are Lagrangian multipliers. The condition (42) is exactly the KKT condition of the problem (7).

REFERENCES

- [1] D. Bharadia, E. McMillin and S. Katti, "Full duplex radios," in *Proc. ACM SIGCOMM'13*, Hong Kong, China, Aug. 2013, pp. 375-386.
- [2] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan and R. Wichman, "In-band FD wireless: Challenges and opportunities," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1637-1652, Sep. 2014.
- [3] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2180-2189, Jun. 2008.
- [4] T. Lv, Y. Yin, Y. Lu, S. Yang, E. Liu and G. Clapworthy, "Physical detection of misbehavior in relay systems with unreliable channel state information," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 7, pp 1517 - 1530 Apr. 2018.
- [5] T. Lv, H. Gao and S. Yang, "Secrecy transmit beamforming for heterogeneous networks," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 6, pp. 1154-1170, Jun. 2015.
- [6] Z. Kong, S. Yang, F. Wu, S. Peng, L. Zhong and L. Hanzo, "Iterative distributed minimum total MSE approach for secure communications in MIMO interference channels," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 3, pp. 594-608, Mar. 2016.
- [7] R. Cao, T. F. Wong, T. Lv, H. Gao and S. Yang, "Detecting Byzantine attacks without clean reference," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 12, pp. 2717-2731, Dec. 2016.
- [8] Y. Zou, X. Wang, W. Shen and L. Hanzo, "Security versus reliability analysis of opportunistic relaying," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 6, pp. 2653-2661, Jul. 2014.
- [9] X. Ding, T. Song, Y. Zou, X. Chen and L. Hanzo, "Security-reliability tradeoff analysis of artificial noise aided two-way opportunistic relay selection," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 5, pp. 3930-3941, May 2017.
- [10] H. Jin, W. Y. Shin and B. C. Jung, "On the multi-user diversity with secrecy in uplink wiretap networks," *IEEE Commun. Lett.*, vol. 17, no. 9, pp. 1778-1781, Sep. 2013.
- [11] H. Jin, B. C. Jung and W. Y. Shin, "On the secrecy capacity of multi-cell uplink networks with opportunistic scheduling," in *Proc. IEEE International Conference on Communications (ICC'06)*, Malaysia, May 2016, pp. 1-5.
- [12] V.D. Nguyen, H. V. Nguyen, O. A. Dobre, and O.-S. Shin, "A new design paradigm for secure full-duplex multiuser systems," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 7, pp. 1480-1498, Jul. 2018.

- [13] F. Zhu, F. Gao, M. Yao and H. Zou, "Joint information- and jamming-beamforming for physical layer security with full duplex base station," *IEEE Trans. Signal Process.*, vol. 62, no. 24, pp. 6391-6401, Dec. 2014.
- [14] F. Zhu, F. Gao, T. Zhang, K. Sun and M. Yao, "Physical-layer security for full duplex communications with self-interference mitigation," *IEEE Trans. Wireless Commun.* vol. 15, no. 1, pp. 329-340, Jan. 2016.
- [15] S. Parsaefard and T. Le-Ngoc, "Improving wireless secrecy rate via full-duplex relay-assisted protocols," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 10, pp. 2095-2107, Oct. 2015.
- [16] Z. Mobini, M. Mohammadi, and C. Tellambura, "Wireless-Powered Full-Duplex Relay and Friendly Jamming for Secure Cooperative Communications," *IEEE Trans. Inf. Forensics Security*, vol. 14, pp. 621-634, Mar. 2019.
- [17] S. Li, Q. Li, and S. Shao, "Robust secrecy beamforming for full-duplex two-way relay networks under imperfect channel state information," *Sci. China Inf. Sci.*, vol. 61, no. 2, Feb. 2018.
- [18] R. Feng, Q. Li, Q. Zhang and J. Qin, "Robust secure beamforming in MISO full-duplex two-way secure communications," *IEEE Trans. Wireless Commun.*, vol. 65, no. 1, pp. 408-414, Jan. 2016.
- [19] J. Huang and A. L. Swindlehurst, "Robust secure transmission in MISO channels based on worst-case optimization," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 1696-1707, Apr. 2012.
- [20] D. P. Bertsekas, *Nonlinear Programming*. Athena Scientific, 1999.
- [21] J. Huang and A. L. Swindlehurst, "Robust Secure Transmission in MISO Channels Based on Worst-Case Optimization," *IEEE Trans. Signal Process.*, vol. 60, no.4, pp. 1696-1707, Apr. 2012.
- [22] D. W. K. Ng, E. S. Lo and R. Schober, "Robust Beamforming for Secure Communication in Systems with Wireless Information and Power Transfer," *IEEE Trans. Wireless Commun.*, vol.13, no. 8, pp. 4599-4615, Aug. 2014.
- [23] Q. Li, L. Yang, Q. Zhang and J. Qin, "Robust AN-Aided Secure Precoding for an AF MIMO Untrusted Relay System," *IEEE Trans. Veh. Technol.*, vol. 66, no. 11, pp. 10572-10576, Nov. 2017.
- [24] M. Tian, X. Huang, Q. Zhang and J. Qin, "Robust AN-Aided Secure Transmission Scheme in MISO Channels with Simultaneous Wireless Information and Power Transfer," *IEEE Signal Process. Lett.*, vol. 22, no. 6, pp. 723-727, Jun. 2015.
- [25] C. Wang and H. M. Wang, "Robust Joint Beamforming and Jamming for Secure AF Networks: Low-Complexity Design," *IEEE Trans. Veh. Technol.*, vol. 64, no. 5, pp. 2192-2198, May 2015.
- [26] 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, Jan. 2015.
- [27] S. Yang, T. Lv and L. Hanzo, "Semidefinite programming relaxation based virtually antipodal detection for MIMO systems using Gray-coded high-order QAM," *IEEE Trans. Veh. Technol.*, vol. 62, no. 4, pp. 1667-1677, May 2013.
- [28] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [29] Z. Luo, J. F. Sturm and S. Zhang, "Multivariate nonnegative quadratic mappings," *SIAM J. Optim.*, vol. 14, no. 4, pp. 1140-1162, Jul. 2004.
- [30] L. Grippo and M. Sciandrone, "On the convergence of the block nonlinear Gauss-Seidel method under convex constraints," *Operation Research Letter*, vol. 26, pp. 127-136, 2000.
- [31] J. Mattingley and S. Boyd, "Real-time convex optimization in signal processing," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 50-61, May 2010.