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.

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 CSI\(^1\).

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-case\(^2\) 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-Tucker(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 \(\mathbf{H}_E^H \in \mathbb{C}^{1 \times N_T}\) 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 \(\mathbf{g}_D \in \mathbb{C}^{N_R \times 1}\), \(\mathbf{g}_E \in \mathbb{C}^{N_E \times 1}\), and \(\mathbf{g}\), 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

\[
\mathbf{x}_b = \mathbf{v}_s + \mathbf{n}_a, \tag{1}
\]

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 \(\mathbf{n}_a \sim \mathcal{CN}(0, \mathbf{\Omega})\) denotes the AN. The total available power of the BS is denoted by \(P_{\text{tot}}\), while the design of \(\mathbf{v}\) and \(\mathbf{n}_a\) must satisfy the following power constraint:

\[
\text{tr}((\mathbf{v}\mathbf{v}^H + \mathbf{\Omega})) \leq P_{\text{tot}}. \tag{2}
\]

\(^1\)The 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.

\(^2\)In 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.

\(^3\)This 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 $z \sim CN(0,1)$ 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

$$y_B = r^H (\sqrt{P_t}g_B z + H_L x_b + n_B),$$

$$y_D = h^H_D x_b + \sqrt{P_t} g z + n_D,$$

$$y_E = H_E x_b + \sqrt{P_t} g_E z + n_E,$$

respectively, where $n_B, n_D, n_E$ denote the received noise having a zero mean and unit variance/identity covariance matrix. Since the perfect state information of $H_L$ 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 $r = \frac{g_B}{||g_B||_F}$, hence we have $||r||_F = 1$.

### 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 $h^H_D$ may be estimated at the BS and the interfering channel $g_E$ 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 $g_B$ can be acquired at the BS. By contrast, the residual self-interference channel $H_L$ 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 $H_E$ and $g_E$, and we have

$$H_E = \hat{H}_E + \Delta H_E, \quad g_E = \hat{g}_E + \Delta g_E,$$

where $\hat{H}_E$ and $\hat{g}_E$ denote the estimates of $H_E$ and $g_E$, respectively, while $\Delta H_E$ and $\Delta g_E$ 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 subcarrier, the perfect CSI of the wiretap channel will be available specifically, when the Eve is constituted by an active subcarrier, the perfect CSI of the wiretap channel will be available.

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, $\hat{H}_L, \Delta H_E, \text{and} \Delta g_E$ are bounded by the sets $\xi_{\hat{H}_L} \triangleq \{H_L : ||H_L||_F \leq \delta_{\hat{H}_L}\}$, $\xi_{\Delta H_E} \triangleq \{\Delta H_E : ||\Delta H_E||_F \leq \delta_{\Delta H_E}\}$ and $\xi_{\Delta g_E} \triangleq \{\Delta g_E : ||\Delta g_E||_F \leq \delta_{\Delta g_E}\}$ respectively, where $\delta_{\hat{H}_L}, \delta_{\Delta H_E}$ and $\delta_{\Delta g_E}$ 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 $v$ and the covariance matrix $\Omega$ of AN in order to maximize the worst-case sum secrecy rate. Defining $Q = vv^H$, we can obtain

$$\max_{v, \Omega, \delta_{\hat{H}_E} \in \xi_{\hat{H}_E}} \min_{\Delta H_E \in \xi_{\Delta H_E}, \Delta g_E \in \xi_{\Delta g_E}} \ln (1 + \eta_1) + \ln (1 + \eta_2) - \ln \det (I_{N_E} + ZN^{-1})$$

s.t. $\text{rank}(Q) = 1$, $\text{tr}(Q + \Omega) \leq P_{tot},$ \hspace{1cm} (7)

where $\eta_1 \triangleq \frac{h^H_D Q h_B}{P_t |g|^2 + h^H_D \Omega h_B + 1}$, $\eta_2 \triangleq \frac{P_t |g|^2}{r^H h_B (Q + \Omega) H^H_B r + 1}$, $Z \triangleq H_E Q H^H_E + P_t g_E g_E^H$, $N \triangleq H_E \Omega H^H_E + I_{N_E}$.

### 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:

$$- \ln \left( P_t |g|^2 + h^H_D \Omega h_B + 1 \right) = \max_{a_1 > 0} \left[ -a_1 \left( P_t |g|^2 + h^H_D \Omega h_B + 1 \right) + \ln a_1 + 1 \right],$$

$$- \ln \left( 1 + r^H H_L (vv^H + \Omega) H^H_L r \right) = \max_{a_2 > 0} \left[ -a_2 \left( 1 + r^H H_L (vv^H + \Omega) H^H_L r \right) + \ln a_2 + 1 \right],$$

$$- \ln \det \left( I_{N_E} + H_E (vv^H + \Omega) H^H_E + P_t g_E g_E^H \right) = N_E + \max_{\text{rank}(W) \leq 1} \ln \det (W - \Omega),$$

where $X \triangleq I_{N_E} + H_E (vv^H + \Omega) H^H_E + P_t g_E g_E^H$. Then, expanding the objective function of (7) and invoking the above relationship, we have

$$\max_{\text{rank}(Q) = 1, \text{tr}(Q + \Omega) \leq P_{tot}, \Omega > 0} \xi (a_1, a_2, W, Q, \Omega)$$

s.t. $\text{rank}(Q) = 1, \text{tr}(Q + \Omega) \leq P_{tot},$ \hspace{1cm} (11a)

where the objective function (11a) is given in the equation (12).
\[ \Xi(a_1, a_2, W_E, Q, \Omega) \triangleq \ln (h_B^T v v^H h_E + P_t |g|^2 + h_B^T \Omega h_E + 1) - a_1 (P_t |g|^2 + h_B^T \Omega h_E + 1) + \ln a_1 + \frac{1}{2} \ln (P_t |g|^2 + h_B^T H_L (v v^H + \Omega) H_L^T r) - a_2 [1 + r^H H_L (v v^H + \Omega) H_L^T r] + \ln a_2 + 1 - N_E + \ln \det (W_E) - \text{tr} (W_E X) + \text{indet} (I_{N_E} + H_E \Omega H_E^H) \]

\[ F(a_1, a_2, \alpha, \beta, \gamma, W_E, Q, \Omega) \triangleq \ln \left(1 + P_t |g|^2 + h_B^T (Q + \Omega) h_E\right) - a_1 (P_t |g|^2 + h_B^T \Omega h_E + 1) + \ln a_1 + a_2 (1 + \alpha) + \ln \left(1 + P_t |g|^2 + h_B^T \Omega h_E + 1\right) + \ln \det (W_E) - \text{tr} (W_E) - \beta - P_t \gamma + \ln \det (I_{N_E} + M) \]

\[ \lambda \left[ \begin{bmatrix} I_{N_T \times N_E} & 0 \\ 0 & -\delta_{H_E}^2 \end{bmatrix} \right] - \begin{bmatrix} (Q + \Omega)^T \otimes W_E \\ \text{vec} (H_E)^H \begin{bmatrix} (Q + \Omega)^T \otimes W_E \end{bmatrix} \text{vec} (H_E) \end{bmatrix} \begin{bmatrix} (Q + \Omega)^T \otimes W_E \end{bmatrix} \text{vec} (H_E) - \beta \right] \succeq 0, \forall \lambda \geq 0. \]

By introducing optimization variables \( \alpha, \beta, \gamma \), the problem (11) can be reformulated as

\[ \max_{\alpha > 0, \beta > 0, \gamma \geq 0} F(a_1, a_2, \alpha, \beta, \gamma, W_E, Q, \Omega) \]

s.t. \( \alpha \geq r^H H_L (Q + \Omega) H_L^T r \), \( \forall H_L \in \xi_{H_L} \); \( \beta \geq \text{tr} (W_E (H_E (Q + \Omega) H_E^H) \), \( \forall \Delta H_E \in \xi_{\Delta H_E} \); \( \gamma \geq \text{tr} (W_E (P_t g_E g_E^H) \), \( \forall \Delta g_E \in \xi_{\Delta g_E} \); \( M \leq H_E \Omega H_E^H \), \( \forall \Delta H_E \in \xi_{\Delta H_E} \); \( \text{rank} (Q) = 1 \); \( \text{tr} (Q + \Omega) \leq P_{tot}\). \]

where the objective function \( F(a_1, a_2, \alpha, \beta, \gamma, W_E, Q, \Omega) \) is defined in (17).

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

\[ r^H H_L (Q + \Omega) H_L^T r = \text{tr} (r^H H_L (Q + \Omega) H_L^H) \triangleq \left( \text{vec} (H_L)^H \begin{bmatrix} (Q + \Omega)^T \otimes r r^H \end{bmatrix} \text{vec} (H_L) \right), \]

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

\[ \text{tr} (A B C D) = \left( \text{vec} (D^T) \right)^T (C^T \otimes A) \text{vec} (B). \]

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:

\[ \max_{\alpha > 0, \beta > 0, \gamma \geq 0} F(a_1, a_2, \alpha, \beta, \gamma, W_E, Q, \Omega) \]

s.t. \( \alpha \geq \left( \text{vec} (H_L)^H \begin{bmatrix} (Q + \Omega)^T \otimes r r^H \end{bmatrix} \text{vec} (H_L) \right) \), \( \forall H_L \in \xi_{H_L} \); \( \beta \geq \left( \text{vec} (\Delta H_E)^H \begin{bmatrix} (Q + \Omega)^T \otimes W_E^H \end{bmatrix} \text{vec} (\Delta H_E) \right) \); \( \gamma \geq \left( \text{vec} (H_E)^H \begin{bmatrix} (Q + \Omega)^T \otimes W_E^H \end{bmatrix} \text{vec} (H_E) \right) \), \( \forall \Delta H_E \in \xi_{\Delta H_E} \); \( \text{rank} (Q) = 1 \); \( \text{tr} (Q + \Omega) \leq P_{tot} \). \]

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

\[ \alpha \geq \delta_{H_L} \lambda_{\text{max}} \left( (Q + \Omega)^T \otimes r r^H \right), \]

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

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:

\[ \lambda_{\gamma} \left[ \begin{bmatrix} I_{N_T \times N_E} & 0 \\ 0 & -\delta_{g_E}^2 \end{bmatrix} \right] - \begin{bmatrix} W_E & \text{vec} (g_E) \\ \text{vec} (g_E^H) & W_E g_E - \gamma \end{bmatrix} \succeq 0, \forall \lambda_{\gamma} \geq 0. \]
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
\[
\begin{bmatrix}
    \mathbf{H}_E \mathbf{Q}^H - \mathbf{M} \\
    \mathbf{H}_E \mathbf{Q}
\end{bmatrix}
\begin{bmatrix}
    \mathbf{1}_{N_E} \\
    0
\end{bmatrix}
- \frac{1}{\lambda_M}
\begin{bmatrix}
    \mathbf{I}_{N_E} \\
    0
\end{bmatrix}
\geq 0,
\]
\[\exists \lambda_M \geq 0. \tag{21}\]

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 \( \mathbf{Q} \) of the AN can be reformulated as
\[
\begin{align*}
\max_{a_1, a_2, \alpha, \beta, \gamma, \mathbf{W}_E, \mathbf{Q}, \Omega} & \quad F(a_1, a_2, \alpha, \beta, \gamma, \mathbf{W}_E, \mathbf{Q}, \Omega) \\
\text{s.t.} & \quad (18), (19), (20), (21) \text{ and } (16f). \tag{22}
\end{align*}
\]

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: \( \{\alpha, \beta, \mathbf{Q}, \Omega\} \) 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 \( \{\alpha, \beta, \mathbf{Q}, \Omega\} \), and when the block \( \{\alpha, \beta, \mathbf{Q}, \Omega\} \) 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 \( \mathbf{Q} \) and \( \hat{\mathbf{Q}} \), respectively, in the problem (7) to give the proof.

Firstly, we have the following linear matrix inequality
\[
\mathbf{Q} - \hat{\mathbf{Q}} = \mathbf{Q}^{1/2} \left( \mathbf{I} - \frac{\mathbf{Q}^{1/2} \mathbf{H}_B \mathbf{h}_B^H \mathbf{Q}^{1/2}}{\mathbf{h}_B^H \mathbf{Q} \mathbf{h}_B} \right) \mathbf{Q}^{1/2} \geq 0. \tag{23}
\]

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

1. Set \( l = 1 \) and perform initialization with arbitrary feasible \( \hat{a}_1(l), \hat{a}_2(l), \mathbf{W}_E(l) \).
2. while the difference between the values of \( F(\hat{a}_1(l), \hat{a}_2(l), \alpha(l), \beta(l), \gamma(l), \mathbf{W}_E(l), \mathbf{Q}(l), \Omega(l)) \) in successive iterations is larger than \( \phi \) for some \( \phi > 0 \), do
   1. Solve the problem (22) by fixing \( a_1 = \hat{a}_1(l), a_2 = \hat{a}_2(l), \mathbf{W}_E = \mathbf{W}_E(l), \) and obtain the global optimum, i.e., \( \alpha^*, \beta^*, \gamma^*, \mathbf{Q}^*, \Omega^* \).
   2. Set \( \hat{\alpha}(l) = \alpha^*, \hat{\beta}(l) = \beta^*, \hat{\gamma}(l) = \gamma^*, \hat{\mathbf{Q}}(l) = \mathbf{Q}^*, \) and \( \hat{\Omega}(l) = \Omega^* \).
   3. Solve the problem (22) by fixing \( \mathbf{a} = \hat{\mathbf{a}}(l), \beta = \hat{\beta}(l), \gamma = \hat{\gamma}(l), \mathbf{Q} = \hat{\mathbf{Q}}(l), \Omega = \hat{\Omega}(l) \) and obtain the global optimum, i.e., \( a_1^*, a_2^*, \mathbf{W}_E^* \).
   4. \( l = l + 1 \).
5. Set \( \hat{a}_1(l) = a_1^*, \hat{a}_2(l) = a_2^*, \) and \( \mathbf{W}_E(l) = \mathbf{W}_E^* \).
end while
Output: \( \mathbf{Q}, \Omega \).

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

1. Set \( \hat{\mathbf{Q}} = \mathbf{Q}_0 h_B^H h_B \).
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 \( \hat{\mathbf{Q}} \) 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 \( |h_l|_2 \), in the denominator of \( \eta_2 \), as well as in \( \mathbf{Z} \). Then, for establishing the proof, the following inequalities must be proved:
\[
\mathbf{h}_B^H \mathbf{Q} \mathbf{h}_B \leq \mathbf{h}_B^H \hat{\mathbf{Q}} \mathbf{h}_B, \tag{24}
\]
\[
\mathbf{r}^H \mathbf{H}_L \mathbf{Q} \mathbf{h}_B \mathbf{r} \geq \mathbf{r}^H \mathbf{H}_L \hat{\mathbf{Q}} \mathbf{h}_B \mathbf{r}, \tag{25}
\]
\[
\mathbf{H}_E \mathbf{Q} \mathbf{h}_B \geq \mathbf{H}_E \hat{\mathbf{Q}} \mathbf{h}_B. \tag{26}
\]

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 \( \mathbf{Q} - \hat{\mathbf{Q}} \succeq 0 \), (25) and (26) also hold true.

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 \( \{\hat{a}_1(l), \hat{a}_2(l), \hat{\alpha}(l), \hat{\beta}(l), \hat{\gamma}(l), \mathbf{W}_E(l), \mathbf{Q}(l), \hat{\Omega}(l)\} \) generated by Algorithm 1, i.e., \( \{\hat{a}_1^*, \hat{a}_2^*, \hat{\alpha}^*, \hat{\beta}^*, \hat{\gamma}^*, \mathbf{W}_E^*, \mathbf{Q}^*, \hat{\Omega}^*\} \) 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 \( \Omega \), the achievable worst-case sum secrecy rate \( R_w \) can
performance degradation becomes significant when $\epsilon$ is small. 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 $\delta_{HE} = \delta_{HL} = \delta_{gE} = \epsilon$, $P_{tot} = 10$ dBm and $N_R = 2$. We observe that the achievable average $R_w$ decreases upon increasing $\epsilon$, and the secrecy performance degradation becomes significant when $\epsilon$ 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 $P_{tot}$, 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 $v$ is constructed by projecting $h_B$ onto the null-space of the known part of the CSI of the eavesdropping channel, i.e., $H_E$. Explicitly,

$$v = P_t \frac{U_E U_E^H h_B}{\|U_E U_E^H h_B\|_F},$$

where, $P_t$ is the transmit power of the information signal, and $P_{tot}$, 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 $v$ is constructed by projecting $h_B$ onto the null-space of the known part of the CSI of the eavesdropping channel, i.e., $H_E$. Explicitly,
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 Q and Ω. 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 $\eta_2$, i.e., $\eta_2^{\text{min}}$ can be calculated as $\eta_2^{\text{min}} = \frac{P_t||g_E||_F^2}{\delta_{H_E} \lambda_{\text{max}}((Q+\Omega)H_E^H)^{-1}}$.

Then, we calculate the maximum of $\ln \det (I_{N_E} + ZN^{-1})$, which can be expanded as

$$
\ln \det \left( I_{N_E} + H_E Q H_E^H (H_E \Omega H_E^H + I_{N_E})^{-1} \right) - \ln \left( 1 + \ln \left( 1 + \ln \left( 1 + \ln \left( 1 + \ln (1 + \theta_1) + \ln (1 + \theta_2) \right) \right) \right) \right)
$$

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

Then, the upper bound of $\ln \det (I_{N_E} + ZN^{-1})$ is given by $\ln (1 + \theta_1) + \ln (1 + \theta_2)$, where

$$
\theta_1 = \arq \min \theta_1,
$$

$$
\theta_2 = \arq \min \theta_2,
$$

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

$$
\begin{align*}
\min_{\theta_1} & \quad 0 \quad 0 \\
\text{s.t.} & \quad \hat{H}_E (\theta_1 \Omega - Q) H_E^H \hat{H}_E (\theta_1 \Omega - Q) - t \Phi \succeq 0, \\
& \quad \exists t \geq 0,
\end{align*}
$$

where we have $\Phi \triangleq \begin{bmatrix} I_{N_E} & 0 \\ 0 & -H_E \end{bmatrix}$.

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

$$
\begin{align*}
\min_{\theta_2} & \quad 0 \quad 0 \\
\text{s.t.} & \quad \hat{H}_E (\theta_2 \Omega - Q) H_E^H \hat{H}_E (\theta_2 \Omega - Q) - \eta \left[ \begin{array}{c} I_{N_E} \\ 0 \end{array} \right] - \frac{P_t g_E g_E^H}{||g_E||_F^2} \succeq 0, \\
& \quad \exists \eta \geq 0,
\end{align*}
$$

where $g_E \triangleq g_E + \frac{\bar{g}_E}{||g_E||_F^2} \bar{g}_E$. $R_w$ is then calculated by

$$
R_w = \ln (1 + \eta) + \ln \left( 1 + \eta_2^{\text{min}} \right) - \ln \left( (1 + \theta_1) (1 + \theta_2) \right).
$$

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 B
PROOF OF PROPOSITION 2

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

**Corollary 1. Consider the problem**

\[
\min_{X, \alpha} f(X, \alpha), \quad s.t. (X, \alpha) \in \mathcal{X} \times \mathcal{A},
\]  
(34)

where \( f(X, \alpha) \) 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 \( \{(X^{(n)}, \alpha^{(n)})\} \) generated by optimizing \( X \) and \( \alpha \) alternatively has limit points. Every limit point of \( \{(X^{(n)}, \alpha^{(n)})\} \) is a stationary point of the problem.

We have used matrix inequality for transforming the problem (11) into the problem (22), equivalently, Therefore, 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: \( (a_1, a_2, W_E) \) and \( (\alpha, \beta, Q, \Omega) \) alternately, two variable sets have limit points. Using Bolzano-Weierstrass theorem, we can conclude that \( \{a_1(l), a_2(l), \alpha(l), \beta(l), \gamma(l), W_E(l), Q(l), \Omega(l)\} \) has limit points. Therefore, with [30, Corollary 2], the limit points generated by Algorithm 1, i.e., \( \{a_1^*, a_2^*, \alpha^*, \beta^*, \gamma^*, W_E^*, Q^*, \Omega^*\} \) is a stationary point of the problem (22), and \( \{a_1^*, a_2^*, W_E^*, Q^*, \Omega^*\} \) 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 \( Q \) whose has one-rank without performance deterioration, the constraint rank \( (Q) = 1 \) of the problem (7) and (11) can be omitted. For brevity, we denote the objective function of the problem (7) as \( \varphi(Q, \Omega) \). Since \( \{a_1^*, a_2^*, \alpha^*, \beta^*, \gamma^*, W_E^*, Q^*, \Omega^*\} \) is a stationary point of the problem (11), we have

\[
\begin{align*}
\text{tr} \left( \nabla_Q \Omega \Xi (a^*_1, a^*_2, W_E^*, Q^*, \Omega^*)^H (Q - Q^*) \right) & \leq 0, \\
\text{tr} (Q + \Omega) & \leq P_{tot}, \\
\text{tr} \left( \nabla_Q \Omega \Xi (a^*_1, a^*_2, W_E^*, Q^*, \Omega^*)^H (Q - Q^*) \right) & \leq 0, \\
\text{tr} (Q + \Omega) & \leq P_{tot}, \\
\text{tr} \left( \nabla_{a_1, a_2} (a^*_1, a^*_2, W_E^*, Q^*, \Omega^*)^H (a_1 - a_1^*) \right) & \leq 0, a_1 > 0, \\
\text{tr} \left( \nabla_{a_1, a_2} (a^*_1, a^*_2, W_E^*, Q^*, \Omega^*)^H (a_2 - a_2^*) \right) & \leq 0, a_2 > 0.
\end{align*}
\]  
(35)

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

\[
W_E = (I_{N_E} + H_E (vv^H + \Omega) H_E^H + P_{ge} g_E g_E^H)^{-1}
\]  
(36)
\[
a_1 = (P_{ge} |g|^2 + h_E^H H_B h_B + 1)^{-1}
\]  
(37)
\[
a_2 = (1 + r^H H_L (vv^H + \Omega) H_L^H r)^{-1}
\]  
(38)

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

\[
\nabla_Q \Xi (a^*_1, a^*_2, W_E^*, Q^*, \Omega^*) = \nabla_Q \varphi(Q^*, \Omega^*)
\]  
(39)
\[
\nabla_{a_1, a_2} (a^*_1, a^*_2, W_E^*, Q^*, \Omega^*) = \nabla_{a_1, a_2} \varphi(Q^*, \Omega^*)
\]  
(39)

Therefore, with the equation above, we can further obtain

\[
\begin{align*}
\text{tr} \left( \nabla_Q \varphi(Q^*, \Omega^*)^H (Q - Q^*) \right) & \leq 0, \text{tr} (Q + \Omega) \leq P_{tot}, \\
\text{tr} \left( \nabla_{a_1, a_2} \varphi(Q^*, \Omega^*)^H (Q - Q^*) \right) & \leq 0, \text{tr} (Q + \Omega) \leq P_{tot}.
\end{align*}
\]  
(40)

Then, we can conclude that \( Q^*, \Omega^* \) is the optimal solution of the following problem:

\[
\max_{Q, \Omega} \left( \nabla_Q \varphi(Q^*, \Omega^*)^H (Q - Q^*) \right), \text{ s.t. } \text{tr} (Q + \Omega) \leq P_{tot}
\]  
(41)

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

\[
\nabla_Q \varphi(Q^*, \Omega^*) - \lambda_Q Q + D_Q = 0
\]  
(42)
\[
Q^* D_Q = 0, \lambda_Q \geq 0, D_Q \geq 0
\]  
(42)
\[
\nabla_{a_1, a_2} \varphi(Q^*, \Omega^*) - \lambda_{a_1} a_1 + D_{a_1} = 0
\]  
(42)
\[
Q^* D_{a_1} = 0, \lambda_{a_1} \geq 0, D_{a_1} \geq 0
\]  
(42)

where \( D_Q, \lambda_Q, D_{a_1}, \lambda_{a_1} \) are Lagrangian multipliers. The condition (42) is exactly the KKT codition of the problem (7).

**REFERENCES**


9


