Secrecy Analyses of a Full-Duplex MIMOME Network

Reza Sohrabi, *Student Member, IEEE*, Qiping Zhu, *Student Member, IEEE*, and Yingbo Hua, *Fellow, IEEE*

***Abstract*—This paper presents secrecy analyses of a full-duplex MIMOME network which consists of two full-duplex multi-antenna users (Alice and Bob) and an arbitrarily located multi-antenna eavesdropper (Eve). The paper assumes that Eve's channel state information (CSI) is completely unknown to Alice and Bob except for a small radius of secured zone. The first part of this paper aims to optimize the powers of jamming noises from both users. To handle Eve's CSI being unknown to users, the focus is placed on Eve at the most harmful location, and the large matrix theory is applied to yield a hardened secrecy rate to work on. The performance gain of the power optimization in terms of maximum tolerable number of antennas on Eve is shown to be significant. The second part of this paper shows two analyses of anti-eavesdropping channel estimation (ANECE) that can better handle Eve with any number of antennas. One analysis assumes that Eve has a prior statistical knowledge of its CSI, which yields lower and upper bounds on secure degrees of freedom of the system as functions of the number (N) of antennas on Eve and the size (K) of information packet. The second analysis assumes that Eve does not have any prior knowledge of its CSI but performs blind detection of information, which yields an approximate secrecy rate for the case of K being larger than N.**

***Index Terms*—Physical layer security, secrecy rate, full-duplex radio, MIMOME, jamming, artificial noise, anti-eavesdropping channel estimation (ANECE).**

I. Introduction

SECURITY of wireless networks is of paramount importance in today's world as billions of people around the globe are dependent upon these networks for a myriad of activities for their businesses and lives. Among several key issues in wireless security [1], confidentiality is of particular interest to many researchers in recent years and is a focus of this paper. For convenience, we will refer to confidentiality as security and vice versa.

The traditional way to keep information confidential from unauthorized persons and/or devices is via cryptography at upper layers of the network, which include the asymmetric-key method (involving a pair of public key and private key) and the symmetric-key method (involving a secret key shared between two legitimate users). As the computing capabilities of modern computers (including quantum computers) rapidly improve, the asymmetric-key method is increasingly vulnerable as this method relies on computational complexity for security. In fact, the symmetric-key method is gaining more attraction in applications [2].

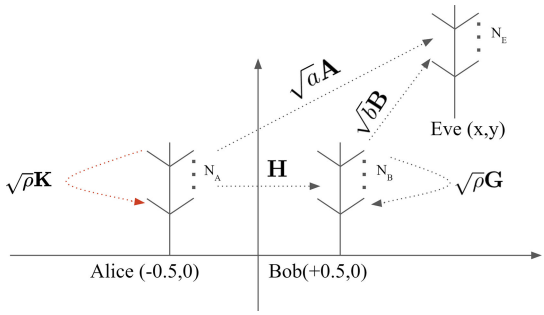


Fig. 1. A Full-Duplex MIMOME network.

However, the establishment of a secret key (or any secret) shared between two users is not trivial in itself. Even if a secret key was pre-installed in a pair of legitimate devices (during manufacturing or otherwise), the lifetime of the secret key in general shortens each time the secret key is used for encryption. For many applications such as big data streaming, such secret key must be periodically renewed or changed. To enjoy the convenience of mobility, it is highly desirable for users to be able to establish a secret key in a wireless fashion.

Establishing a secret key or directly transmitting secret information between users in a wireless fashion (without a pre-existing shared secret) is the essence of physical layer security [3]. There are two complementary approaches in physical layer security: secret-key generation and secret information transmission. The former requires users to use their (correlated) observations and an unlimited public channel to establish a secret key, and the latter requires one user to transmit secret information directly to the other. This paper is concerned with the latter, i.e., transmission of secret information (such as secret key) between users without any prior digital secret.

Specifically, this paper is focused on a network as illustrated in Fig. 1 where one legitimate user (Alice) wants to send a secret key to another legitimate user (Bob) subject to eavesdropping by an eavesdropper (Eve) anywhere. Each of the two users/devices is allowed to have multiple antennas, and both Alice and Bob are capable of full-duplex operations. Following a similar naming

Manuscript received June 8, 2019; revised September 20, 2019; accepted October 20, 2019. Date of publication October 25, 2019; date of current version November 19, 2019. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Stefano Tomasin. This work was supported in part by the Army Research Office under Grant W911NF-17-1-0581. *(Reza Sohrabi and Qiping Zhu are co-first authors.) (Corresponding author: Yingbo Hua.)*

The authors are with the Department of Electrical and Computer Engineering, University of California, Riverside, CA 92521 USA (e-mail: rsohr001@ucr.edu; qzhu005@ucr.edu; yhua@ece.ucr.edu).

Digital Object Identifier 10.1109/TSP.2019.2949501

in the literature such as [4], we call the above setup a full-duplex MIMOME network where MIMOME refers to the multi-input multi-output (MIMO) channel between Alice and Bob and the multi-antenna Eve.

The MIMOME related works in the literature include: [4]–[15] where the channel state information (CSI) at Eve is assumed to be known not only to Eve itself[1] but also to Alice and Bob; [16]–[19] where a partial knowledge of Eve's CSI is assumed to be available to Alice and Bob and an averaged secrecy or secrecy outage was considered; [13], [16], [20]–[24] where artificial noise is embedded in the signal from Alice; and [25]–[30] where Bob is treated as a full-duplex node capable to receive the signal from Alice while transmitting jamming noise.

From the literature, the idea of using jamming noise from Alice or Bob appears important. Inspired by that, this paper will first consider a case where both Alice and Bob send jamming noises while Alice transmits secret information to Bob. We will explore how to optimize the jamming powers from Alice and Bob. In [26], jamming from both users was also considered. But here for power optimization we include the effect of the residual self-interference of full-duplex radio. There are other differences in the problem formulation and objectives. We assume that Eve's CSI is completely unknown to Alice or Bob except for a radius of secured zone free of Eve around Alice. A similar idea was also applied in [31] but in a different problem setting. We will focus on Eve that is located at the most harmful position. Furthermore, to handle the small-scale fading at Eve, we apply the large matrix theory to obtain a closed-form expression of a secrecy rate, which makes the power optimization tractable. Unlike [24] where large matrix theory was also applied, we consider an arbitrary large-scale-fading at Eve among other major differences. With the optimized powers, we reveal a significant performance gain in terms of the maximum tolerable number of antennas on Eve to maintain a positive secrecy. We will also show that as the number of antennas on Eve increases, the impact of the jamming noise from either Alice or Bob on secrecy vanishes. This contribution extends a previous understanding of single-antenna users shown in [30].

Later in this paper, we will analyze a two-phase scheme for secret information transmission proposed in [30]. In the first phase, an anti-eavesdropping channel estimation (ANECE) method is applied which allows users to find their CSI but suppresses Eve's ability to obtain its CSI. In the second phase, secret information is transmitted between Alice and Bob while Eve has little or no knowledge of its CSI. We show two analyses based on two different assumptions. The first analysis assumes that Eve has a prior statistical knowledge of its CSI. With every node knowing a statistical model of CSI anywhere, we use mutual information to analyze the secret rate of the network, from which lower and upper bounds on the secure degrees of freedom are derived. These bounds are simple functions of the number of antennas on Eve. The second analysis assumes that Eve does not have any prior knowledge of its CSI. Due to ANECE in phase 1, Eve is blind to its CSI. But in phase 2, Eve performs blind detection of the information from Alice. We analyze the performance of the blind detection, from which an approximate secret rate is derived and numerically illustrated. Both of these analyses are important contributions useful for a better understanding of ANECE.

Notation: Matrices and column vectors are denoted by upper and lowercase boldface letters. The trace, Hermitian transpose, column-wise vectorization, (i,j)th element, and complex conjugate of a matrix $\mathbf{A}$ are denoted by $\mathrm{Tr}(\mathbf{A})$, $\mathbf{A}^H$, $\mathrm{vec}(\mathbf{A})$, $\mathbf{A}_{i,j}$, and $\mathbf{A}^*$, respectively. For a matrix $\mathbf{X}$ and its vectorized version $\mathbf{x}$, $\mathrm{ivec}(\mathbf{x})$ is the inverse operation of $\mathbf{x} = \mathrm{vec}(\mathbf{X})$. A diagonal matrix with elements of $\mathbf{x}$ on its diagonal is $\mathrm{diag}(\mathbf{x}^T)$. Expectation with respect to a random variable x is denoted by $\mathcal{E}_x[\cdot]$. Let the random variables X_n and X be defined on the same probability space, and we write $X_n \overset{a.s.}{\to} X$ if X_n converges to X almost surely as $n \to \infty$. The identity matrix of the size $n \times n$ is $\mathbf{I}_n$ (or $\mathbf{I}$ with n implied in the context), and $\mathbf{1}_n$ is a row vector of length n of all ones. A circularly symmetric complex Gaussian random variable x with variance σ^2 is denoted as $x \sim \mathcal{CN}(0, \sigma^2)$. The mutual information between random variables x and y is $I(x;y)$, and $h(x)$ denotes the differential entropy of x. Logarithm in base 2 is denoted by $\log(\cdot)$, and $(\cdot)^+ \triangleq \max(0, \cdot)$.

II. Optimization of Jamming Powers and Effects of Eve's Antennas

A. System Model

Our network setup is shown in Fig. 1, where Alice (with N_A antennas) intends to send secret information over a wireless channel to Bob (with N_B antennas) in the presence of possibly many passive Eves (of N_E antennas each) that may collude with each other at the network layer but not at the physical layer. We will focus on the most harmful Eve. Physical layer colluding among distributed Eves to form a large virtual antenna array is highly difficult in practice. But if a virtual antenna array from colluding Eves is likely in some applications, we could treat these colluding Eves as a single mega Eve with a large number of antennas.

The system parameters are normalized in a similar way as in [28]. In particular, the large-scale-fading factor from Alice to Eve is modeled as (when a model is needed): $a = d_A^{-\alpha} = ((x+0.5)^2 + y^2)^{-\alpha/2}$, and that from Bob to Eve is $b = d_B^{-\alpha} = ((x-0.5)^2 + y^2)^{-\alpha/2}$ where α is the path-loss exponent. We assume that no Eve is closer to Alice than a radius Δ, i.e., $d_A \geq \Delta$. The normalized large-scale-fading factor of the residual self-interference at both Alice and Bob is denoted by ρ. (In all simulations, ρ is considered to be 0.1%.) The small-scale-fading channel matrix from Alice to Eve is denoted by $\mathbf{A}$, that from Bob to Eve is $\mathbf{B}$, and that of the residual self-interference at Bob and Alice are $\mathbf{G}$ and $\mathbf{K}$, respectively.[2] The channel matrix from Alice to Bob is denoted by $\mathbf{H}$, and its SVD is denoted by

$$\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H, \tag{1}$$

where $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices, and $\mathbf{\Sigma}$ is the $N_B \times N_A$ diagonal matrix that contains the singular values of $\mathbf{H}$ (i.e., σ_i, $i = 1, \ldots, N_B$) in descending order assuming $N_A \geq N_B$.

[1] All entities are treated as "gender neutral".

[2] Up to Section III, Alice is only a transmitter, and hence it does not utilize its full-duplex capability.

All the elements in all channel matrices are modeled as i.i.d. circularly symmetric complex Gaussian random variables with zero mean and unit variance.

In this section, we assume that Alice and Bob have the knowledge of $\mathbf{H}$ but not of $\mathbf{A}$ and $\mathbf{B}$, and Eve has the knowledge of all these matrices.

Alice sends the following signal containing $r \le N_B \le N_A$ streams of secret information mixed with artificial noise:

$$\mathbf{x}_A(k) = \mathbf{V}_1\mathbf{s}(k) + \mathbf{V}_2\mathbf{w}_A(k), \tag{2}$$

where k is the index of time slot, $\mathbf{V}_1$ is the first r columns of $\mathbf{V}$, $\mathbf{V}_2$ is the last $N_A - r$ columns of $\mathbf{V}$, $\mathbf{s}(k)$ is Alice's information vector with the covariance matrix $\mathbf{Q}_r$ and $\mathrm{Tr}(\mathbf{Q}_r) = P_s$, and $\mathbf{w}_A(k)$ is an $(N_A - r) \times 1$ artificial noise vector with distribution $\mathcal{CN}(\mathbf{0}, \frac{P_n}{N_A - r}\mathbf{I})$. Here, $P_s + P_n = P_A \le P_A^{\max}$.

While Bob receives information from Alice, it also sends a jamming noise:

$$\mathbf{x}_B(k) = \mathbf{w}_B(k), \tag{3}$$

where $\mathbf{w}_B(k)$ is an $N_B \times 1$ artificial noise vector with distribution $\mathcal{CN}(\mathbf{0}, \frac{P_B}{N_B}\mathbf{I})$.

Note that both P_A and P_B are normalized powers with respect to the path loss from Alice to Bob, and with respect to the power of the background noise. So, without loss of generality, we let the power of the background noise be one.

With jamming from both Alice and Bob, the signals received by Bob and Eve are respectively:

$$\mathbf{y}_B(k) = \mathbf{H}\mathbf{V}_1\mathbf{s}(k) + \mathbf{H}\mathbf{V}_2\mathbf{w}_A(k) + \sqrt{\rho}\mathbf{G}\bar{\mathbf{w}}_B(k) + \mathbf{n}_B(k), \tag{4}$$

$$\mathbf{y}_E(k) = \sqrt{a}\mathbf{A}_1\mathbf{s}(k) + \sqrt{a}\mathbf{A}_2\mathbf{w}_A(k) + \sqrt{b}\mathbf{B}\mathbf{w}_B(k) + \mathbf{n}_E(k), \tag{5}$$

where $[\mathbf{A}_1, \mathbf{A}_2] = [\mathbf{A}\mathbf{V}_1, \mathbf{A}\mathbf{V}_2] = \mathbf{A}\mathbf{V}$. Since $\mathbf{A}_1$ and $\mathbf{A}_2$ are linear functions of the Gaussian matrix $\mathbf{A}$, they remain Gaussian. Because of the unitary nature of $\mathbf{V}$, $\mathbf{A}_1$ and $\mathbf{A}_2$ are independent of each other, and all elements in them are i.i.d. Gaussian of zero mean and unit variance. The noise vectors $\mathbf{n}_B$ and $\mathbf{n}_E$ are distributed as $\mathcal{CN}(\mathbf{0}, \mathbf{I})$. Also note that $\sqrt{\rho}\mathbf{G}\bar{\mathbf{w}}_B(k)$ is the residual self-interference originally caused by $\mathbf{w}_B(k)$ but is independent of $\mathbf{w}_B(k)$ [30].

If CSI anywhere is known everywhere, the achievable secrecy rate of the above system is known [32] to be

$$R_S = (R_{AB} - R_{AE})^+ \tag{6}$$

where R_{AB} is the rate from Alice to Bob and R_{AE} is the rate from Alice to Eve. Namely,

$$R_{AB} = \log|\mathbf{I} + \mathbf{C}_B^{-1}\mathbf{H}\mathbf{V}_1\mathbf{Q}_r\mathbf{V}_1^H\mathbf{H}^H|, \tag{7}$$

$$R_{AE} = \log|\mathbf{I} + a\mathbf{C}_E^{-1}\mathbf{A}_1\mathbf{Q}_r\mathbf{A}_1^H|, \tag{8}$$

where

$$\mathbf{C}_B = \mathbf{I} + \frac{P_n}{N_A - r}\mathbf{H}\mathbf{V}_2\mathbf{V}_2^H\mathbf{H}^H + \frac{\rho P_B}{N_B}\mathbf{G}\mathbf{G}^H, \tag{9}$$

$$\mathbf{C}_E = \mathbf{I} + \frac{aP_n}{N_A - r}\mathbf{A}_2\mathbf{A}_2^H + \frac{bP_B}{N_B}\mathbf{B}\mathbf{B}^H. \tag{10}$$

Note that since $\mathbf{H}\mathbf{V}_1 = \mathbf{U}_1\mathbf{\Sigma}_1$ and $\mathbf{H}\mathbf{V}_2 = \mathbf{U}_2\mathbf{\Sigma}_2$ are orthogonal to each other where $\mathbf{U}_1$ and $\mathbf{U}_2$ are the partitions of $\mathbf{U}$ similar to those of $\mathbf{V}$, and $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ are the corresponding diagonal partitions of $\mathbf{\Sigma}$, a sufficient statistics of $\mathbf{s}(k)$ at Bob is $\mathbf{U}_1^H\mathbf{y}_B(k) = \mathbf{\Sigma}_1\mathbf{s}(k) + \sqrt{\rho}\mathbf{U}_1\mathbf{G}\bar{\mathbf{w}}_B(k) + \mathbf{U}_1\mathbf{n}_B(k)$ which shows that the artificial noise from Alice does not affect Bob. Consequently, an equivalent form of R_{AB} is

$$R_{AB} = \log|\mathbf{I}_r + \mathbf{C}_{B,1}^{-1}\mathbf{\Sigma}_1\mathbf{Q}_r\mathbf{\Sigma}_1|, \tag{11}$$

where

$$\mathbf{C}_{B,1} = \mathbf{I}_r + \frac{\rho P_B}{N_B}\mathbf{U}_1^H\mathbf{G}\mathbf{G}^H\mathbf{U}_1. \tag{12}$$

However, the expression shown in (7) is needed later due to its direct connection to $\mathbf{H}$ and $\mathbf{G}$.

If we optimize P_n, P_B and $\mathbf{Q}_r$ to maximize the above R_S, the solution would be a function of Eve's CSI. This does not appear to be useful in practice.

If Eve's CSI is unknown to Alice but the statistics of Eve's CSI is known to Alice, then we can consider the ergodic secrecy:

$$\bar{R}_S = (\mathcal{E}_{\mathbf{H},\mathbf{G}}[R_{AB}] - \mathcal{E}_{\mathbf{A},\mathbf{B}}[R_{AE}])^+ \tag{13}$$

which is achievable via coding over many CSI coherence periods. Closed form expression of each of the two terms in the above can be obtained using ideas in [33] and [34]. But if we use $\bar{R}_S$ as objective to optimize P_n, P_B and $\mathbf{Q}_r$, the solution would be independent of the CSI between Alice and Bob, and such a solution is not very useful either.

Because of the above reasons, we will consider the worst case of R_{AE}. The worst case is such that Eve is located at the most harmful location and has a large number of antennas.

It is shown in our earlier work [35] that the most harmful position of Eve is at $x^* = -0.5 - \Delta$ and $y^* = 0$. From now on, we will refer to a and b as corresponding to the position (x^*, y^*). In all simulations, we will use $\Delta = 0.1$ unless mentioned otherwise.

Given a large number of antennas at Eve, we can use large matrix theory to obtain a closed-form expression of R_{AE} that is no longer dependent on instantaneous CSI at Eve, which is shown next. We can rewrite (8) as follows:

$$R_{AE} = \log|\mathbf{I} + \mathbf{J}_3\bar{\mathbf{\Theta}}_3\mathbf{J}_3^H| - \log|\mathbf{I} + \mathbf{J}_4\bar{\mathbf{\Theta}}_4\mathbf{J}_4^H| \tag{14}$$

where $\mathbf{J}_3 = \frac{1}{\sqrt{N_E}}[\mathbf{A}_1, \mathbf{A}_2, \mathbf{B}]$, $\mathbf{J}_4 = \frac{1}{\sqrt{N_E}}[\mathbf{A}_2, \mathbf{B}]$, and

$$\bar{\mathbf{\Theta}}_3 = N_E\ \mathrm{diag}\left[a\mathbf{q}_r^T, \frac{aP_n}{N_A - r}\mathbf{1}_{N_A - r}, \frac{bP_B}{N_B}\mathbf{1}_{N_B}\right], \tag{15}$$

$$\bar{\mathbf{\Theta}}_4 = N_E\ \mathrm{diag}\left[\frac{aP_n}{N_A - r}\mathbf{1}_{N_A - r}, \frac{bP_B}{N_B}\mathbf{1}_{N_B}\right] \tag{16}$$

where $\mathbf{q}_r$ is the vector containing the diagonal elements of the diagonal matrix $\mathbf{Q}_r$ (assuming that Alice does not know Bob's self-interference channel). Note that the $\mathbf{J}$ matrices consist of i.i.d. random variables and the $\bar{\mathbf{\Theta}}$ matrices are diagonal. (The numbering of 3 and 4 used here is because of the numbering later.)

Lemma 1: Let $\mathbf{J}$ be an $N \times K$ matrix whose entries are i.i.d. complex random variables with variance $\frac{1}{N}$, and $\mathbf{\Theta}$ be a

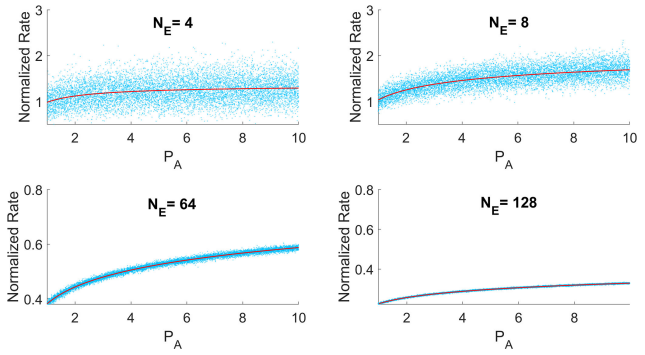


Fig. 2. Comparison of the exact random realizations of $\frac{R_{AE}}{N_E}$ with its asymptotic result (the red/solid curve).

diagonal deterministic matrix. Based on Theorem (2.39) of [36], as $N, K \to \infty$ with $\frac{K}{N} \to \beta$, we have

$$\frac{1}{N}\log\left|\mathbf{I}+\mathbf{J}\boldsymbol{\Theta}\mathbf{J}^H\right| \overset{a.s.}{\to} \Omega\left(\beta,\boldsymbol{\Theta},\eta\right), \tag{17}$$

where

$$\Omega\left(\beta,\boldsymbol{\Theta},\eta\right) \triangleq \beta\mathcal{V}_{\boldsymbol{\Theta}}\left(\eta\right)-\log\left(\eta\right)+\left(\eta-1\right)\log(e), \tag{18}$$

$$\mathcal{V}_{\boldsymbol{\Theta}}\left(\eta\right) \triangleq \frac{1}{L_{\boldsymbol{\Theta}}}\sum_{j=1}^{L_{\boldsymbol{\Theta}}}\log\left(1+\eta\boldsymbol{\Theta}_{j,j}\right). \tag{19}$$

Here, $\boldsymbol{\Theta}_{j,j}$ is the jth diagonal element of the diagonal matrix $\boldsymbol{\Theta}$, $L_{\boldsymbol{\Theta}}$ is the number of diagonal elements of $\boldsymbol{\Theta}$, and $\eta > 0$ is the solution to the equation

$$1-\eta = \frac{\beta\eta}{L_{\boldsymbol{\Theta}}}\sum_{j=1}^{L_{\boldsymbol{\Theta}}}\frac{\boldsymbol{\Theta}_{j,j}}{1+\eta\boldsymbol{\Theta}_{j,j}}. \tag{20}$$

Proof: The proof is given in Appendix A. ■

Using the lemma, it follows from (14) that for large N_E,

$$\begin{aligned} R_{AE} &\simeq \sum_{j=1}^{L_{\bar{\boldsymbol{\Theta}}_3}}\log\left(1+\bar{\eta}_3\left(\bar{\boldsymbol{\Theta}}_3\right)_{j,j}\right)-\sum_{j=1}^{L_{\bar{\boldsymbol{\Theta}}_4}}\log\left(1+\bar{\eta}_4\left(\bar{\boldsymbol{\Theta}}_4\right)_{j,j}\right) \\ &\quad + N_E\log\left(\frac{\bar{\eta}_4}{\bar{\eta}_3}\right)+N_E\left(\bar{\eta}_3-\bar{\eta}_4\right)\log(e) \\ &\triangleq \mathcal{R}_{AE} \end{aligned} \tag{21}$$

where for $i = 3, 4$, $\bar{\eta}_i$ is the solution of η to (20) with $\beta = \bar{\beta}_i$ and $\boldsymbol{\Theta} = \bar{\boldsymbol{\Theta}}_i$. Here, $\bar{\beta}_3 = \frac{N_A+N_B}{N_E}$ and $\bar{\beta}_4 = \frac{N_A-r+N_B}{N_E}$. Note that the right side of (20) is a monotonic function of $\eta \geq 0$ and hence a unique solution of $0 \leq \eta \leq 1$ can be easily found by bisection search.

It is useful to note that the asymptotic form $\mathcal{R}_{AE}$ is a good approximation of the exact form R_{AE} as long as N_E is large regardless of N_A and N_B. Shown in Fig. 2 is a comparison of the exact random realizations of $\frac{R_{AE}}{N_E}$ from (8) with its asymptotic result $\frac{\mathcal{R}_{AE}}{N_E}$ from (21) where $N_A = 2N_B = 8$, $r = N_B$, $\mathbf{Q}_r = \frac{P_s}{N_B}\mathbf{I}$, $P_s = P_n = \frac{P_A}{2}$. Note that 100 realizations of $\frac{R_{AE}}{N_E}$ corresponding to 100 random realizations of Eve's CSI for each value of P_A are shown. We see that as N_E increases (beyond 8), $\frac{\mathcal{R}_{AE}}{N_E}$ becomes a good approximation of $\frac{R_{AE}}{N_E}$.

B. Power Optimization and Maximum Tolerable Number of Antennas on Eve

With $(R_{AB}-\mathcal{R}_{AE})^+$ as the objective function, we can now develop an optimization algorithm to optimize the power distribution. Note that since the CSI required for R_{AB} is known to Alice and Bob, we do not need to replace R_{AB} by its asymptotic form.

Specifically, we can use this cost function $g(\mathbf{x}_r) \triangleq \mathcal{R}_{AE} - R_{AB}$ where $\mathbf{x}_r = [\mathbf{q}_r^T, P_n, P_B]^T$. Then, we need to solve the following problem:

$$\begin{aligned} \min_r \min_{\mathbf{x}_r} \quad & g\left(\mathbf{x}_r\right) \\ \text{s.t.} \quad & \sum_{i=1}^{r}\mathbf{q}_r(i)+P_n \leq P_A^{\max} \\ & \mathbf{q}_r(i) \geq 0,\ \forall i = 1,\ldots,r \\ & P_n \geq 0 \\ & 0 \leq P_B \leq P_B^{\max}. \end{aligned} \tag{22}$$

Here the optimization of r is simple, which can be done via sequential search. For a given r, the above problem is not convex. Although the constraints are convex, the cost $g(\mathbf{x}_r)$ is not. To see this, let us rewrite this function as follows:

$$\begin{aligned} g\left(\mathbf{x}_r\right) =& -\log\left|\mathbf{C}_B+\mathbf{H}\mathbf{V}_1\mathbf{Q}_r\mathbf{V}_1^H\mathbf{H}^H\right|+\log\left|\mathbf{C}_B\right| \\ &+\sum_{j=1}^{L_{\bar{\boldsymbol{\Theta}}_3}}\log\left(1+\bar{\eta}_3\left(\bar{\boldsymbol{\Theta}}_3\right)_{j,j}\right)-\sum_{j=1}^{L_{\bar{\boldsymbol{\Theta}}_4}}\log\left(1+\bar{\eta}_4\left(\bar{\boldsymbol{\Theta}}_4\right)_{j,j}\right) \\ &+N_E\log\left(\frac{\bar{\eta}_4}{\bar{\eta}_3}\right)+N_E\left(\bar{\eta}_3-\bar{\eta}_4\right)\log(e). \end{aligned} \tag{23}$$

The non-convex parts of $g(\mathbf{x}_r)$ are $\log|\mathbf{C}_B|$ and $\sum_{j=1}^{L_{\bar{\boldsymbol{\Theta}}_3}}\log(1+\bar{\eta}_3(\bar{\boldsymbol{\Theta}}_3)_{j,j})$, which are concave functions of $\mathbf{x}_r$. These two terms can be replaced by their upper bounds based on the first-order Taylor-series expansion around the solution of the previous iteration. Also, the dependence of $g(\mathbf{x}_r)$ on $\bar{\eta}_3$ and $\bar{\eta}_4$ can be resolved by choosing the values of $\bar{\eta}_3$ and $\bar{\eta}_4$ as follows:

$$1-\bar{\eta}_i^t = \frac{\bar{\beta}_i\bar{\eta}_i^t}{L_{\bar{\boldsymbol{\Theta}}_i}}\sum_{j=1}^{L_{\bar{\boldsymbol{\Theta}}_i}}\frac{(\bar{\boldsymbol{\Theta}}_i^t)_{j,j}}{1+\bar{\eta}_i^t(\bar{\boldsymbol{\Theta}}_i^t)_{j,j}}, i = 3,4. \tag{24}$$

where t denotes the tth iteration. In other words, at iteration t, the following convex problem is solved:

$$\begin{aligned} \mathbf{x}_r^{t+1} = \operatorname*{argmin}_{\mathbf{x}_r} \quad & h^t\left(\mathbf{x}_r\right) \\ \text{s.t.} \quad & \sum_{i=1}^{r}\mathbf{q}_r(i)+P_n \leq P_A^{\max} \\ & \mathbf{q}_r(i) \geq 0,\ \forall i = 1,\ldots,r \\ & P_n \geq 0 \\ & 0 \leq P_B \leq P_B^{\max}. \end{aligned} \tag{25}$$

Algorithm 1: Algorithm for Power Optimization.

Choose proper ϵ, $\bar{\eta}_3^0$, $\bar{\eta}_4^0$, and set $g^{\min}=0$.
for $r=1:N_A$ **do**
  Initialize $\mathbf{x}_r$ satisfying the constraints.
  Set $t=0$.
  while $\frac{\|\mathbf{x}_r^t-\mathbf{x}_r^{t-1}\|}{\|\mathbf{x}_r^{t-1}\|}>\epsilon$ **do**
    Solve (25) to get $\mathbf{x}_r^{t+1}$.
    Update $\bar{\eta}_3$, $\bar{\eta}_4$ by solving (24) using $\mathbf{x}_r^{t+1}$.
    $t=t+1$.
  end while
  if $g(\mathbf{x}_r^t)<g^{\min}$ **then**
    $g^{\min}=g(\mathbf{x}_r^t)$
    $\mathbf{x}^{\min}=\mathbf{x}_r^t$
  end if
end for
Return $\mathbf{x}^{\min}$.

where

$$h^t(\mathbf{x}_r)=-\log|\mathbf{C}_B+\mathbf{H}\mathbf{V}_1\mathbf{Q}_r\mathbf{V}_1^H\mathbf{H}^H| - \sum_{j=1}^{L_{\bar{\mathbf{\Theta}}_4}}\log\left(1+\bar{\eta}_4^t\left(\bar{\mathbf{\Theta}}_4\right)_{j,j}\right)+\left(\mathbf{x}_r-\mathbf{x}_r^t\right)^T\nabla_{\mathbf{x}_r}f^t|_{\mathbf{x}_r=\mathbf{x}_r^t}, \tag{26}$$

and $f^t(\mathbf{x}_r)=\log|\mathbf{C}_B|+\sum_{j=1}^{L_{\bar{\mathbf{\Theta}}_3}}\log(1+\bar{\eta}_3^t(\bar{\mathbf{\Theta}}_3)_{j,j})$. All constant terms in (23) are omitted in (26) as they do not affect the optimization.

Algorithm 1 details the proposed procedure for the power optimization. It is worth mentioning that a different optimization approach was explored in our previous work [35], where a stochastic optimization approach was applied to the objective function $R_{AB}-\mathcal{E}[R_{AE}]$. These two approaches more or less give the same results, but Algorithm 1 in this paper has a significantly lower complexity. To illustrate a performance gain of the optimized powers over non-optimal powers, we will not repeat similar figures as available in [35]. But next we consider the maximum tolerable number of antennas on Eve, which can be defined in several ways. One is

$$\tilde{N}_E\triangleq\max N_E,\ \text{s.t. } R_{AB}-R_{AE}>0. \tag{27}$$

which however depends on instantaneous CSI everywhere. Another is

$$\bar{N}_E\triangleq\max N_E,\ \text{s.t. } \mathcal{R}_{AB}-\mathcal{R}_{AE}>0. \tag{28}$$

where $\mathcal{R}_{AB}$ and $\mathcal{R}_{AE}$ are asymptotic forms of R_{AB} and R_{AE} respectively. Obviously, $\bar{N}_E$ is a function of $\mathbf{x}_r$. The third definition is

$$\bar{N}_E^{opt}\triangleq\max N_E,\ \text{s.t. }\left(\frac{1}{L}\sum_{l=1}^{L}(R_{AB,l}^{opt}-\mathcal{R}_{AE,l}^{opt})\right)^+>0. \tag{29}$$

where $R_{AB,l}^{opt}-\mathcal{R}_{AE,l}^{opt}$ is a value of $R_{AB}-\mathcal{R}_{AE}$ corresponding to a random realization of $\mathbf{H}$ and $\mathbf{G}$ and the corresponding optimal $\mathbf{x}_r$ and r, and L is the total number of realizations of $\mathbf{H}$ and $\mathbf{G}$ for each N_E.

To obtain $\bar{N}_E$ in (28), we will let $N_A>N_B$, $r=N_B$ and $\mathbf{Q}_r=\frac{P_s}{N_B}\mathbf{I}$. Hence, $range(\mathbf{V}_2)$ is the null-space of $\mathbf{H}$. As a consequence, $\mathbf{H}\mathbf{V}_1\mathbf{Q}_r\mathbf{V}_1^H\mathbf{H}^H=\frac{P_S}{N_B}\mathbf{H}\mathbf{H}^H$, and (7) becomes

$$\begin{aligned}R_{AB}=&\\ &\log\left|\mathbf{I}+\frac{\rho P_B}{N_B}\mathbf{G}\mathbf{G}^H+\frac{P_s}{N_B}\mathbf{H}\mathbf{H}^H\right|-\log\left|\mathbf{I}+\frac{\rho P_B}{N_B}\mathbf{G}\mathbf{G}^H\right|\\ &=\log\left|\mathbf{I}+\mathbf{J}_1\mathbf{\Theta}_1\mathbf{J}_1^H\right|-\log\left|\mathbf{I}+\mathbf{J}_2\mathbf{\Theta}_2\mathbf{J}_2^H\right|\end{aligned} \tag{30}$$

where $\mathbf{J}_1=\frac{1}{\sqrt{N_B}}[\mathbf{H},\mathbf{G}]$, $\mathbf{J}_2=\frac{1}{\sqrt{N_B}}\mathbf{G}$, $\mathbf{\Theta}_2=\rho P_B\mathbf{I}$, and $\mathbf{\Theta}_1=\mathrm{diag}([P_s\mathbf{1}_{N_A},\rho P_B\mathbf{1}_{N_B}])$. Applying the lemma to (30) yields that for $\beta_1=\frac{N_A+N_B}{N_B}$, $\beta_2=1$ and a large N_B,

$$\begin{aligned}\frac{R_{AB}}{N_B}&\simeq\Omega(\beta_1,\mathbf{\Theta}_1,\eta_1)-\Omega(\beta_2,\mathbf{\Theta}_2,\eta_2)\\ &=(\beta_1-1)\log(1+\eta_1P_s)+\log\frac{1+\eta_1\rho P_B}{1+\eta_2\rho P_B}\\ &\quad+\log\left(\frac{\eta_2}{\eta_1}\right)+(\eta_1-\eta_2)\log(e)\\ &\triangleq\frac{\mathcal{R}_{AB}}{N_B}\end{aligned} \tag{31}$$

where η_1 is the solution of η to (20) with $\beta=\beta_1$ and $\mathbf{\Theta}=\mathbf{\Theta}_1$, which reduces to

$$1-\eta_1=\frac{(\beta_1-1)\eta_1P_s}{1+\eta_1P_s}+\frac{\eta_1\rho P_B}{1+\eta_1\rho P_B}. \tag{32}$$

and η_2 is the solution to

$$1-\eta_2=\frac{\eta_2\rho P_B}{1+\eta_2\rho P_B}, \tag{33}$$

or equivalently $\eta_2=\frac{\sqrt{1+4\rho P_B}-1}{2\rho P_B}$.

Also with $r=N_B$ and $\mathbf{Q}_r=\frac{P_s}{N_B}\mathbf{I}$, $\mathcal{R}_{AE}$ in (21) reduces to

$$\begin{aligned}\frac{R_{AE}}{N_E}&\simeq\Omega(\beta_3,\mathbf{\Theta}_3,\eta_3)-\Omega(\beta_4,\mathbf{\Theta}_4,\eta_4)\\ &=(\beta_3-\beta_4)\log\left(1+\frac{aP_s\eta_3}{\beta_3-\beta_4}\right)\\ &\quad+(\beta_3-\beta_4)\log\frac{\beta_3-\beta_4+bP_B\eta_3}{\beta_3-\beta_4+bP_B\eta_4}\\ &\quad+(2\beta_4-\beta_3)\log\frac{2\beta_4-\beta_3+aP_n\eta_3}{2\beta_4-\beta_3+aP_n\eta_4}\\ &\quad+\log\left(\frac{\eta_4}{\eta_3}\right)+(\eta_3-\eta_4)\log(e)\\ &\triangleq\frac{\mathcal{R}_{AE}}{N_E}\end{aligned} \tag{34}$$

where

$$\mathbf{\Theta}_3=N_E\mathrm{diag}\left(\left[\frac{aP_s}{N_B}\mathbf{1}_{N_B},\frac{aP_n}{N_A-N_B}\mathbf{1}_{N_A-N_B},\frac{bP_B}{N_B}\mathbf{1}_{N_B}\right]\right), \tag{35}$$

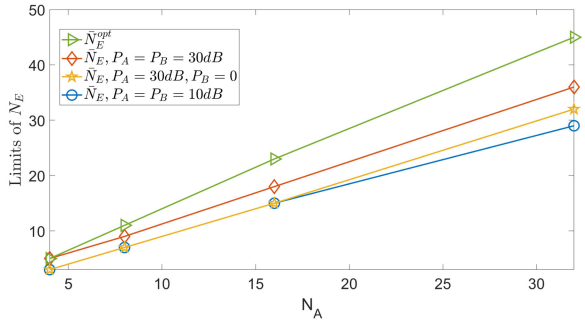


Fig. 3. Comparison of $\bar{N}_E$ and $\bar{N}_E^{opt}$ vs N_A.

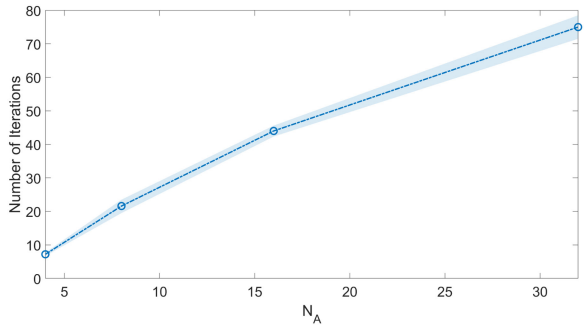


Fig. 4. 95% confidence interval of the number of iterations needed for convergence of Algorithm 1 vs. N_A. The dark line is the mean of the number of iterations.

and

$$\mathbf{\Theta}_4 = N_E \text{diag}\left(\left[\frac{aP_n}{N_A - N_B}\mathbf{1}_{N_A-N_B}, \frac{bP_B}{N_B}\mathbf{1}_{N_B}\right]\right), \quad (36)$$

also $\beta_3 = \frac{N_A+N_B}{N_E}$, $\beta_4 = \frac{N_A}{N_E}$, η_3 is the solution to

$$1-\eta_3 = \frac{aP_s\eta_3}{1+aP_s\eta_3\frac{1}{\beta_3-\beta_4}} + \frac{aP_n\eta_3}{1+aP_n\eta_3\frac{1}{2\beta_4-\beta_3}} + \frac{bP_B\eta_3}{1+bP_B\eta_3\frac{1}{\beta_3-\beta_4}}, \quad (37)$$

and η_4 is the solution to

$$1-\eta_4 = \frac{aP_n\eta_4}{1+aP_n\eta_4\frac{1}{2\beta_4-\beta_3}} + \frac{bP_B\eta_4}{1+bP_B\eta_4\frac{1}{\beta_3-\beta_4}}. \quad (38)$$

Fig. 3 shows $\bar{N}_E$ versus N_A and $\bar{N}_E^{opt}$ versus N_A where $P_A^{\max} = P_B^{\max} = 30$ dB and $N_A = 2N_B$ ($\beta_1 = 3$). For $\bar{N}_E$, we also chose $P_s = P_n = \frac{P_A}{2}$ and $P_A = P_B$. We see that $\bar{N}_E^{opt}$ is consistently larger than $\bar{N}_E$. And the gap between the two is due to the power optimization.

During simulation, we also observed that the optimal P_s is often distributed approximately equally between different streams, and that if N_E gets larger, the optimization favors smaller P_B and smaller r (the latter of which is consistent with a result in [37] which does not use full-duplex jamming at Bob).

With the same parameters as in Fig. 3, Fig. 4 illustrates a convergence property of Algorithm 1 where the mean number of iterations needed for convergence versus N_A is shown. Also shown in Fig. 4 is the 95% confidence interval of the number of iterations needed for convergence versus N_A. We used 100 random realizations of the channels for each value of N_A. The threshold ϵ used for convergence was chosen to be 0.01.

C. When the Number of Antennas on Eve is Very Large

We now consider the case where $N_E \gg N_A > N_B$, $r = N_B$ and $\mathbf{Q}_r = \frac{P_s}{N_B}\mathbf{I}$. It follows that $\beta_3 \ll 1$ and $\beta_4 \ll 1$. Hence, (37) implies $1-\eta_3 \approx \beta_3$ and (38) implies $1-\eta_4 \approx \beta_4$. Furthermore, referring to the terms in (34), we have

$$\lim_{N_E\to\infty} N_E(\beta_3-\beta_4)\log\left(1+\frac{aP_s\eta_3}{\beta_3-\beta_4}\right) = N_B\log\left(1+\frac{N_EaP_s}{N_B}\right), \quad (39)$$

$$\lim_{N_E\to\infty} N_E(\beta_3-\beta_4)\log\frac{\beta_3-\beta_4+bP_B\eta_3}{\beta_3-\beta_4+bP_B\eta_4} = N_B\log 1 = 0, \quad (40)$$

$$\lim_{N_E\to\infty} N_E(2\beta_4-\beta_3)\log\frac{2\beta_4-\beta_3+aP_n\eta_3}{2\beta_4-\beta_3+aP_n\eta_4} = (N_A-N_B)\log 1 = 0, \quad (41)$$

$$\begin{aligned}&\lim_{N_E\to\infty} N_E\left(\log\left(\frac{\eta_4}{\eta_3}\right)+(\eta_3-\eta_4)\log(e)\right)\\ &= \lim_{N_E\to\infty} N_E\left(\log\frac{1-\beta_4}{1-\beta_3}+(\beta_4-\beta_3)\log e\right)\\ &= \lim_{N_E\to\infty} N_E\log\left(1+\frac{N_B}{N_E-(N_A+N_B)}\right) - N_B\log e\\ &= \lim_{N_E\to\infty} N_E\frac{N_B}{N_E-(N_A+N_B)}\log e - N_B\log e\\ &= 0. \end{aligned} \quad (42)$$

The above equations imply that all terms, except the first, in $\mathcal{R}_{AE}$ from (34) converge to zero. Therefore,

$$\lim_{N_E\to\infty} R_{AE} = \lim_{N_E\to\infty}\mathcal{R}_{AE} = N_B\log\left(1+\frac{N_EaP_s}{N_B}\right) \triangleq \mathcal{R}_{AE}^* \quad (43)$$

which is independent of P_n and P_B (and hence the optimal P_n and P_B are now zero). This result implies that if Eve has an unlimited number of antennas then the jamming noise from either Alice or Bob has virtually no impact on Eve's capacity to receive the information from Alice. Furthermore, we see that R_{AE} increases without upper bound as N_E increases while R_{AB} stays independent of N_E (for large N_E).

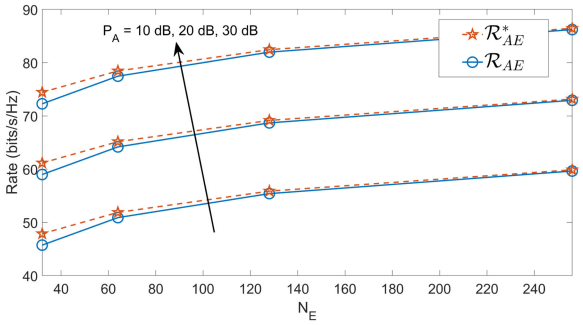


Fig. 5. The convergence of $\mathcal{R}_{AE}$ to $\mathcal{R}^*_{AE}$.

Fig. 5 compares $\mathcal{R}^*_{AE}$ from (42) with $\mathcal{R}_{AE}$ from (34) where $N_A = 2N_B = 8$ and $P_A = P_B = 2P_s = 2P_n$. We see that the two results are very close when $N_E > 40$.

III. Analysis of ANECE

A key observation from the previous section is that if Eve knows its CSI and the number of antennas on Eve is large, then neither the artificial noise from multi-antenna Alice nor the full-duplex jamming from multi-antenna Bob can rescue Alice and Bob from being totally exposed to Eve. (This observation is an extension of a previous observation for single-antenna users shown in [30].) To handle Eve with large number of antennas, there is a two-phase method involving anti-eavesdropping channel estimation (ANECE) proposed in [30]: in phase 1 the users conduct ANECE which allows users to obtain their CSI but denies Eve the same ability; and in phase 2 the users transmit information to each other with Eve not knowing its CSI. Both phases are within a common coherence period. While the earlier work has shown promising properties of ANECE, the understanding of ANECE is still incomplete. In this section, we show two new analyses of the secrecy rate of a two-user MIMOME network assisted by ANECE.

To simplify the problem, we do not consider the artificial noise from either Alice or Bob. The first analysis assumes a (globally known) statistical model for all CSI in the network. And the analysis is based on ideal full-duplex devices where there is no self-interference. When a result of this analysis is applied to practice, one must restrict the application to situations where the residual self-interference is negligible. Typically, the residual self-interference is proportional to the transmitted power which increases with the distance between devices. So, a situation where the residual self-interference is negligible corresponds generally to a short-range communication. The second analysis assumes that Eve does not know the statistical distribution of its CSI but rather assumes that Eve is able to perform blind detection of the information from Alice. These two analyses constitute an important new understanding of ANECE, which is not available elsewhere.

A theory where Eve knows the statistical distribution of its CSI can be applicable to situations where Eve's CSI is statistically stationary and experiences many cycles of coherence periods in a time window of interest. A theory where Eve does not know its CSI distribution can be applicable to situations where Eve's CSI is statistically un-stationary in a time window of interest. Both assumptions have their own merits.

A. Eve Uses a Statistical Model of Its CSI

Consider a block Rayleigh fading channel for which Alice and Bob first conduct ANECE by transmitting their pilot signals $\mathbf{p}_A(k)$ and $\mathbf{p}_B(k)$ concurrently (in full-duplex mode) where $k = 1, \ldots, K_1$ (K_1 is the length of the pilot), and then transmit information to each other (over K_2 samples). For information transmission, we will consider a one-way transmission and a two-way transmission separately.

1) Channel Estimation: Define $\mathbf{P}_i = [\mathbf{p}_i(1), \ldots, \mathbf{p}_i(K_1)]$ where $i = A, B$. then the corresponding signals received by Alice, Bob and Eve can be expressed as

$$\mathbf{Y}_A = \mathbf{H}^T\mathbf{P}_B + \mathbf{N}_A \tag{44a}$$

$$\mathbf{Y}_B = \mathbf{H}\mathbf{P}_A + \mathbf{N}_B \tag{44b}$$

$$\mathbf{Y}_E = \sqrt{a}\mathbf{A}\mathbf{P}_A + \sqrt{b}\mathbf{B}\mathbf{P}_B + \mathbf{N}_E \tag{44c}$$

where $\mathbf{H}$ is the reciprocal channel matrix between Alice and Bob, and all the noise matrices consist of i.i.d. $\mathcal{CN}(0,1)$. Here, the self-interferences at Alice and Bob are assumed to be negligible.

It is known and easy to show that for the best performance of the maximum likelihood (ML) estimation (or the MMSE estimation as shown later) of $\mathbf{H}$ by Bob, $\mathbf{P}_A$ should be such that $\mathbf{P}_A\mathbf{P}_A^H = \frac{K_1P_A}{N_A}\mathbf{I}_{N_A}$. Similarly, $\mathbf{P}_B$ should be such that $\mathbf{P}_B\mathbf{P}_B^H = \frac{K_1P_B}{N_B}\mathbf{I}_{N_B}$.

In the following analysis, we assume that $\mathbf{H}$, $\mathbf{A}$ and $\mathbf{B}$ all consist of i.i.d. $\mathcal{CN}(0,1)$ elements (from one coherence block to another). This statistical model along with the large-scale fading factors a and b is assumed to be known to everyone.

Without loss of generality, let $N_A \geq N_B$. Without affecting the channel estimation performance at Alice and Bob, but maximizing the difficulty of channel estimation for Eve, we let the row span of $\mathbf{P}_B$ be part of the row span of $\mathbf{P}_A$. More specifically, we can write $\mathbf{P}_A = \sqrt{\frac{K_1P_A}{N_A}}[\mathbf{I}_{N_A}, \mathbf{0}_{N_A\times(K_1-N_A)}]\mathbf{\Gamma}$ and $\mathbf{P}_B = \sqrt{\frac{K_1P_B}{N_B}}[\mathbf{I}_{N_B}, \mathbf{0}_{N_B\times(K_1-N_B)}]\mathbf{\Gamma}$ where $\mathbf{\Gamma}$ can be any $K_1 \times K_1$ unitary matrix. In this way, any estimates of $\mathbf{A}$ and $\mathbf{B}$ by Eve, denoted by $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$, are ambiguous in that $[\sqrt{a}\hat{\mathbf{A}}, \sqrt{b}\hat{\mathbf{B}}]$ can be added to $\mathbf{\Theta}[\mathbf{C}_A, \mathbf{C}_B]$ without affecting Eve's observation $\mathbf{Y}_E$ where $\mathbf{\Theta} \in \mathbb{C}^{N_E\times N_B}$ is arbitrary and $[\mathbf{C}_A, \mathbf{C}_B][\mathbf{P}_A^T, \mathbf{P}_B^T]^T = 0$.

Let $\mathbf{h} = vec(\mathbf{H})$, $\mathbf{a} = vec(\mathbf{A})$, $\mathbf{b} = vec(\mathbf{B})$, $\mathbf{y}_A = vec(\mathbf{Y}_A^T)$, $\mathbf{y}_B = vec(\mathbf{Y}_B)$, $\mathbf{n}_A = vec(\mathbf{N}_A^T)$ and $\mathbf{n}_B = vec(\mathbf{N}_B)$. Note $vec(\mathbf{XYZ}) = (\mathbf{Z}^T \otimes \mathbf{X})vec(\mathbf{Y})$. Then (44) becomes

$$\mathbf{y}_A = (\mathbf{I}_{N_A} \otimes \mathbf{P}_B^T)\mathbf{h} + \mathbf{n}_A \tag{45a}$$

$$\mathbf{y}_B = (\mathbf{P}_A^T \otimes \mathbf{I}_{N_B})\mathbf{h} + \mathbf{n}_B \tag{45b}$$

$$\mathbf{y}_E = \sqrt{a}(\mathbf{P}_A^T \otimes \mathbf{I}_{N_E})\mathbf{a} + \sqrt{b}(\mathbf{P}_B^T \otimes \mathbf{I}_{N_E})\mathbf{b} + \mathbf{n}_E. \tag{45c}$$

It is known that the minimum-mean-squared-error (MMSE) estimate of a vector $\mathbf{x}$ from another vector $\mathbf{y}$ is $\hat{\mathbf{x}} = \mathbf{K}_{\mathbf{x},\mathbf{y}}\mathbf{K}_{\mathbf{y}}^{-1}\mathbf{y}$ with $\mathbf{K}_{\mathbf{x},\mathbf{y}} = \mathcal{E}\{\mathbf{x}\mathbf{y}^H\}$ and $\mathbf{K}_{\mathbf{y}} = \mathcal{E}\{\mathbf{y}\mathbf{y}^H\}$. And

the error $\Delta\mathbf{x} = \mathbf{x} - \hat{\mathbf{x}}$ has the covariance matrix $\mathbf{K}_{\Delta\mathbf{x}} = \mathbf{K}_{\mathbf{x}} - \mathbf{K}_{\mathbf{x},\mathbf{y}}\mathbf{K}_{\mathbf{y}}^{-1}\mathbf{K}_{\mathbf{x},\mathbf{y}}^H$.

Let $\hat{\mathbf{h}}_A$ be the MMSE estimate of $\mathbf{h}$ by Alice, and $\Delta\mathbf{h}_A = \mathbf{h} - \hat{\mathbf{h}}_A$ be its error. Similar notations are defined for Bob and Eve. It is easy to show that the covariance matrices of the errors of these estimates are, respectively, $\mathbf{K}_{\Delta\mathbf{h}_A} = \sigma_A^2\mathbf{I}_{N_AN_B}$, $\mathbf{K}_{\Delta\mathbf{h}_B} = \sigma_B^2\mathbf{I}_{N_AN_B}$, $\mathbf{K}_{\Delta\mathbf{a}} = \sigma_{EA}^2\mathbf{I}_{N_AN_E}$ and $\mathbf{K}_{\Delta\mathbf{b}} = \sigma_{EB}^2\mathbf{I}_{N_BN_E}$ where $\sigma_A^2 = \frac{1}{1+K_1P_B/N_B}$, $\sigma_B^2 = \frac{1}{1+K_1P_A/N_A}$, $\sigma_{EA}^2 = \frac{bK_1P_B/N_B+1}{(aK_1P_A/N_A+bK_1P_B/N_B)+1}$ and $\sigma_{EB}^2 = \frac{aK_1P_A/N_A+1}{(aK_1P_A/N_A+bK_1P_B/N_B)+1}$.

2) One-Way Information Transmission: Now assume that following the pilots (over K_1 samples) transmitted by Alice and Bob in full-duplex mode, Alice transmits information (over K_2 samples) to Bob in half-duplex mode. Namely, while the first phase is in full-duplex, the second phase is in half-duplex. In the second phase, Bob and Eve receive

$$\begin{aligned}\mathbf{Y}_B &= \mathbf{H}\mathbf{S}_A + \mathbf{N}_B\\ \mathbf{Y}_E &= \sqrt{a}\mathbf{A}\mathbf{S}_A + \mathbf{N}_E\end{aligned} \tag{46}$$

where $\mathbf{S}_A = [\mathbf{s}_A(1), \ldots, \mathbf{s}_A(K_2)]$. The corresponding vector forms of the above are

$$\mathbf{y}_B = (\mathbf{I}_{K_2} \otimes \mathbf{H})\bar{\mathbf{s}}_A + \mathbf{n}_B \tag{47a}$$

$$\mathbf{y}_E = \sqrt{a}(\mathbf{I}_{K_2} \otimes \mathbf{A})\bar{\mathbf{s}}_A + \mathbf{n}_E \tag{47b}$$

where $\bar{\mathbf{s}}_A = vec(\mathbf{S}_A)$ (which is assumed to be independent of all channel parameters). Then an achievable secrecy rate in bits/s/Hz in phase 2 from Alice to Bob (conditional on the MMSE channel estimation in phase 1) is

$$\mathcal{R}_{one} = \frac{1}{K_2}\big(I(\bar{\mathbf{s}}_A;\mathbf{y}_B|\hat{\mathbf{h}}_B) - I(\bar{\mathbf{s}}_A;\mathbf{y}_E|\hat{\mathbf{a}})\big)^+ \tag{48}$$

To analyze $\mathcal{R}_{one}$, we now assume $P_A = P_B = P$ (which holds for both phases 1 and 2) and that $\mathbf{s}_A(k)$ are i.i.d. with $\mathcal{CN}(0, \frac{P_A}{N_A}\mathbf{I}_{N_A})$. We also use $\hat{\mathbf{H}}_B = ivec(\hat{\mathbf{h}}_B) \in \mathbb{C}^{N_B\times N_A}$ (i.e., $\hat{\mathbf{h}}_B = vec(\hat{\mathbf{H}}_B)$).

We will next derive lower and upper bounds on $\mathcal{R}_{one}$. To do that, we need to obtain lower and upper bounds on $I(\bar{\mathbf{s}}_A;\mathbf{y}_B|\hat{\mathbf{h}}_B)$ and those on $I(\bar{\mathbf{s}}_A;\mathbf{y}_E|\hat{\mathbf{a}})$.

First, we have

$$\begin{aligned}I(\bar{\mathbf{s}}_A;\mathbf{y}_B|\hat{\mathbf{h}}_B) &= h(\bar{\mathbf{s}}_A|\hat{\mathbf{h}}_B) - h(\bar{\mathbf{s}}_A|\mathbf{y}_B,\hat{\mathbf{h}}_B)\\ &= h(\bar{\mathbf{s}}_A) - h(\bar{\mathbf{s}}_A|\mathbf{y}_B,\hat{\mathbf{h}}_B).\end{aligned} \tag{49}$$

It is known that $h(\bar{\mathbf{s}}_A) = \log[(\pi e)^{N_AK_2}|\frac{P_A}{N_A}\mathbf{I}_{N_AK_2}|]$. It is also known [38] that for a random vector $\mathbf{s} \in \mathbb{C}^{n\times 1}$ and another random vector $\mathbf{w}$, $h(\mathbf{s}|\mathbf{w}) \le \log[(\pi e)^n|\mathbf{K}_{\mathbf{s}|\mathbf{w}}|]$ where $\mathbf{K}_{\mathbf{s}|\mathbf{w}} = \mathbf{K}_{\mathbf{s}} - \mathbf{K}_{\mathbf{s},\mathbf{w}}(\mathbf{K}_{\mathbf{w}})^{-1}\mathbf{K}_{\mathbf{s},\mathbf{w}}$ which is the covariance matrix of the MMSE estimation of $\mathbf{s}$ from $\mathbf{w}$. Note that $\mathbf{y}_B = (\mathbf{I}_{K_2}\otimes\hat{\mathbf{H}}_B)\bar{\mathbf{s}}_A + (\mathbf{I}_{K_2}\otimes\Delta\mathbf{H}_B)\bar{\mathbf{s}}_A + \mathbf{n}_B$. Then conditional on $\hat{\mathbf{H}}_B$ (which is independent of $\bar{\mathbf{s}}_A$), the covariance matrix of the MMSE estimate of $\bar{\mathbf{s}}_A$ from $\mathbf{y}_B$ is $\mathbf{K}_{\bar{\mathbf{s}}_A|\mathbf{y}_B,\hat{\mathbf{h}}_B} = \frac{P_A}{N_A}\mathbf{I}_{N_AK_2} - \frac{P_A^2}{N_A^2}(\mathbf{I}_{K_2}\otimes\hat{\mathbf{H}}_B^H)(\frac{P_A}{N_A}(\mathbf{I}_{K_2}\otimes\hat{\mathbf{H}}_B\hat{\mathbf{H}}_B^H) + \mathbf{K}_B + \mathbf{I}_{N_BK_2})^{-1}(\mathbf{I}_{K_2}\otimes\hat{\mathbf{H}}_B)$ where $\mathbf{K}_B = \mathcal{E}\{(\mathbf{I}_{K_2}\otimes\Delta\mathbf{H}_B)\bar{\mathbf{s}}_A\bar{\mathbf{s}}_A^H(\mathbf{I}\otimes\Delta\mathbf{H}_B^H)\} = \frac{P_A}{1+K_1P_A/N_A}\mathbf{I}_{N_BK_2}$. Using $|\mathbf{I}_{r_A} + \mathbf{AB}| = |\mathbf{I}_{r_B} + \mathbf{BA}|$ where r_A and r_B are the numbers of rows of $\mathbf{A}$ and $\mathbf{B}$ respectively, one can verify that $\log|\mathbf{K}_{\bar{\mathbf{s}}_A|\mathbf{y}_B,\hat{\mathbf{h}}_B}| = N_AK_2\log\frac{P_A}{N_A} + \log|\mathbf{K}_B + \mathbf{I}_{N_BK_2}| - \log|\frac{P_A}{N_A}(\mathbf{I}_{K_2}\otimes\hat{\mathbf{H}}_B\hat{\mathbf{H}}_B^H) + \mathbf{K}_B + \mathbf{I}_{N_BK_2}| = N_AK_2\log\frac{P_A}{N_A} - K_2\log|\mathbf{I}_{N_B} + \frac{P_A/N_A}{1+\frac{P_A}{1+K_1P_A/N_A}}\hat{\mathbf{H}}_B\hat{\mathbf{H}}_B^H|$. Applying the above results to (49) yields

$$\begin{aligned}&I(\bar{\mathbf{s}}_A;\mathbf{y}_B|\hat{\mathbf{h}}_B)\\ &\ge \log\left|\frac{P_A}{N_A}\mathbf{I}_{N_AK_2}\right| - \mathcal{E}\{\log|\mathbf{K}_{\bar{\mathbf{s}}_A|\mathbf{y}_B,\hat{\mathbf{h}}_B}|\}\\ &= K_2\mathcal{E}\left\{\log\left|\mathbf{I}_{N_B} + \frac{P_A/N_A}{1+\frac{P_A}{1+K_1P_A/N_A}}\hat{\mathbf{H}}_B\hat{\mathbf{H}}_B^H\right|\right\}\\ &\triangleq \mathcal{R}_B^-.\end{aligned} \tag{50}$$

To derive an upper bound on $I(\bar{\mathbf{s}}_A;\mathbf{y}_B|\hat{\mathbf{h}}_B)$, we now write

$$I(\bar{\mathbf{s}}_A;\mathbf{y}_B|\hat{\mathbf{h}}_B) = h(\mathbf{y}_B|\hat{\mathbf{h}}_B) - h(\mathbf{y}_B|\hat{\mathbf{h}}_B,\bar{\mathbf{s}}_A). \tag{51}$$

Here, $h(\mathbf{y}_B|\hat{\mathbf{h}}_B) \le \mathcal{E}\{\log[(\pi e)^{N_BK_2}|\frac{P_A}{N_A}(\mathbf{I}_{K_2}\otimes\hat{\mathbf{H}}_B\hat{\mathbf{H}}_B^H) + \mathbf{K}_B + \mathbf{I}_{N_BK_2}|]\} = K_2\mathcal{E}\{\log[(\pi e)^{N_B}|\frac{P_A}{N_A}(\hat{\mathbf{H}}_B\hat{\mathbf{H}}_B^H) + (1+\frac{P_A}{1+K_1P_A/N_A})\mathbf{I}_{N_B}|]\}$, and $h(\mathbf{y}_B|\hat{\mathbf{h}}_B,\bar{\mathbf{s}}_A) = \mathcal{E}\{\log[(\pi e)^{N_BK_2}|\frac{1}{1+K_1P_A/N_A}(\mathbf{S}_A^T\mathbf{S}_A^*\otimes\mathbf{I}_{N_B}) + \mathbf{I}_{N_BK_2}|]\} = N_B\mathcal{E}\{\log[(\pi e)^{K_2}|\frac{1}{1+K_1P_A/N_A}(\mathbf{S}_A^T\mathbf{S}_A^*) + \mathbf{I}_{K_2}|]\}$. Note that conditional on $\hat{\mathbf{h}}_B$ and $\bar{\mathbf{s}}_A$ the covariance matrix of $\mathbf{y}_B$ is invariant to $\hat{\mathbf{h}}_B$. Now define

$$\mathbf{M}_A = \begin{cases}\frac{N_A}{P_A}\mathbf{S}_A^T\mathbf{S}_A^*, & K_2 < N_A\\ \frac{N_A}{P_A}\mathbf{S}_A^*\mathbf{S}_A^T, & K_2 \ge N_A\end{cases} \tag{52}$$

which is a full rank matrix for any N_A and K_2 and a self-product of $\sqrt{\frac{N_A}{P_A}}\mathbf{S}_A$ with i.i.d. $\mathcal{CN}(0,1)$ entries. Also define $t_A = \min\{N_A, K_2\}$ and $r_A = \max\{N_A, K_2\}$. It follows that (as part of $h(\mathbf{y}_B|\hat{\mathbf{h}}_B,\bar{\mathbf{s}}_A)$)

$$\begin{aligned}&\mathcal{E}\left\{\log\left|\frac{1}{1+K_1P_A/N_A}\left(\mathbf{S}_A^T\mathbf{S}_A^*\right) + \mathbf{I}_{K_2}\right|\right\}\\ &= \mathcal{E}\left\{\log\left|\frac{P_A/N_A}{1+K_1P_A/N_A}\mathbf{M}_A + \mathbf{I}_{t_A}\right|\right\}\\ &\ge t_A\mathcal{E}\left\{\log\left(1 + \left|\frac{P_A/N_A}{1+K_1P_A/N_A}\mathbf{M}_A\right|^{\frac{1}{t_A}}\right)\right\} \qquad (53a)\\ &= t_A\mathcal{E}\left\{\log\left(1 + \frac{P_A/N_A}{1+K_1P_A/N_A}\exp\left(\frac{1}{t_A}\ln|\mathbf{M}_A|\right)\right)\right\}\\ &\ge t_A\log\left(1 + \frac{P_A/N_A}{1+K_1P_A/N_A}\exp\left(\frac{1}{t_A}\mathcal{E}\left\{\ln|\mathbf{M}_A|\right\}\right)\right) \qquad (53b)\\ &= t_A\log\left(1 + \frac{P_A/N_A}{1+K_1P_A/N_A}\exp\left(\frac{1}{t_A}\sum_{j=1}^{t_A}\sum_{k=1}^{r_A-j}\frac{1}{k} - \gamma\right)\right) \qquad (53c)\end{aligned}$$

where (53a) is due to the matrix Minkowskis inequality $|\mathbf{X}+\mathbf{Y}|^{1/n} \geq |\mathbf{X}|^{1/n}+|\mathbf{Y}|^{1/n}$ where $\mathbf{X}$ and $\mathbf{Y}$ are $n \times n$ positive definite matrices [39], (53b) is due to the Jensen's inequality and that $\log(1+ae^x)$ is a convex function of x when $a>0$, and (53c) is based on [40, Th. 1] where $\gamma \cong 0.57721566$ is Euler's constant. Defining $e_A = \exp(\frac{1}{t_A}\sum_{j=1}^{t_A}\sum_{k=1}^{r_A-j}\frac{1}{k}-\gamma)$ and applying the above results since (51), we have from (51) that

$$\begin{aligned} & I(\bar{\mathbf{s}}_A;\mathbf{y}_B|\hat{\mathbf{h}}_B) \\ & \leq K_2\mathcal{E}\left\{\log\left|\mathbf{I}_{N_B}+\frac{P_A/N_A\hat{\mathbf{H}}_B\hat{\mathbf{H}}_B^H}{1+\frac{P_A}{1+K_1P_A/N_A}}\right|\right\} \\ & \quad + N_B\log\left(\frac{\left(1+\frac{P_A}{1+K_1P_A/N_A}\right)^{K_2}}{\left(1+\frac{P_A/N_A}{1+K_1P_A/N_A}e_A\right)^{t_A}}\right) \\ & \triangleq \mathcal{R}_B^+ \end{aligned} \tag{54}$$

From (50) and (54) we see that the difference between the upper and lower bounds on $I(\bar{\mathbf{s}}_A;\mathbf{y}_B|\hat{\mathbf{h}}_B)$ is the second term in (54).

To consider $I(\bar{\mathbf{s}}_A;\mathbf{y}_E|\hat{\mathbf{a}})$ in (48), we let $\hat{\mathbf{A}} = ivec(\hat{\mathbf{a}})$. Similar to the discussions leading to (50) and (54), one can verify that

$$I(\bar{\mathbf{s}}_A;\mathbf{y}_E|\hat{\mathbf{a}}) \geq K_2\mathcal{E}\left\{\log\left|\mathbf{I}_{N_E}+\frac{P_A/N_A\hat{\mathbf{A}}\hat{\mathbf{A}}^H}{1+P_A\sigma_{EA}^2}\right|\right\} \triangleq \mathcal{R}_E^- \tag{55}$$

and

$$\begin{aligned} & I(\bar{\mathbf{s}}_A;\mathbf{y}_E|\hat{\mathbf{a}}) \\ & \leq \mathcal{R}_E^- + N_E\log\left(\frac{(1+P_A\sigma_{EA}^2)^{K_2}}{\left(1+(P_A\sigma_{EA}^2/N_A)e_A\right)^{t_A}}\right) \\ & \triangleq \mathcal{R}_E^+ \end{aligned} \tag{56}$$

When $P_A=P_B=P\to\infty$, we have $\sigma_{EA}^2 \to \frac{bN_A}{aN_B+bN_A}$, $\sigma_B^2\to 0$, $\mathcal{E}\{\hat{a}_i\hat{a}_i^*\}\to\frac{aN_B}{aN_B+bN_A}$ and $\mathcal{E}\{\hat{h}_{B,i}\hat{h}_{B,i}^*\}\to 1$. From [41, Th. 2], we know that $\mathcal{E}\{\log|\mathbf{I}_r+\frac{P}{t}\mathbf{X}\mathbf{X}^H|\}\to \min(r,t)\log P + o(\log P)$ as $P\to\infty$ where the entries of $\mathbf{X}\in\mathbb{C}^{r\times t}$ are i.i.d. $\mathcal{CN}(0,1)$. Therefore, from (50) and (54),

$$\lim_{P\to\infty}\frac{\mathcal{R}_B^-}{\log P} = \lim_{P\to\infty}\frac{\mathcal{R}_B^+}{\log P} = K_2\min\{N_A,N_B\} \tag{57}$$

And from (55) and (56), we have

$$\lim_{P\to\infty}\frac{\mathcal{R}_E^-}{\log P} = 0 \tag{58}$$

and

$$\lim_{P\to\infty}\frac{\mathcal{R}_E^+}{\log P} = \begin{cases} 0, & K_2\leq N_A \\ N_E(K_2-N_A), & K_2>N_A \end{cases} \tag{59}$$

Combining (57), (58) and (59) and using $\mathcal{R}_{one}^+ \triangleq \frac{1}{K_2}[\mathcal{R}_B^+ - \mathcal{R}_E^-]^+$ and $\mathcal{R}_{one}^- \triangleq \frac{1}{K_2}[\mathcal{R}_B^- - \mathcal{R}_E^+]^+$ (i.e., $\mathcal{R}_{one}^- \leq \mathcal{R}_{one} \leq \mathcal{R}_{one}^+$), we have

$$\begin{aligned} & \lim_{P\to\infty}\frac{\mathcal{R}_{one}^-}{\log P} \\ & = \begin{cases} \min\{N_A,N_B\}, & K_2\leq N_A \\ \left(\min\{N_A,N_B\}-\frac{N_E}{K_2}(K_2-N_A)\right)^+, & K_2>N_A \end{cases} \end{aligned} \tag{60}$$

and

$$\lim_{P\to\infty}\frac{\mathcal{R}_{one}^+}{\log P} = \min\{N_A,N_B\}. \tag{61}$$

Note that $\lim_{P\to\infty}\frac{\mathcal{R}_{one}}{\log P}$ is called the secure degrees of freedom of the one-way information transmission. From (60) and (61), we see that when $K_2\leq N_A$, we have $\lim_{P\to\infty}\frac{\mathcal{R}_{one}}{\log P} = \min\{N_A,N_B\}$ which equals the degrees of freedom of the main channel capacity from Alice to Bob. This supports and complements a conclusion from [30] where the analysis did not use the complete statistical model of $\mathbf{H}$, $\mathbf{A}$ and $\mathbf{B}$. We also see from (60) that if $K_2>N_A$, the above lower bound on secure degrees of freedom decreases linearly as N_E increases.

3) Two-Way Information Transmission: Now we consider a two-way (full-duplex) communication in the second phase where the signals received by Alice, Bob and Eve in a coherence period are

$$\begin{aligned} \mathbf{Y}_A &= \mathbf{H}^T\mathbf{S}_B + \mathbf{N}_A \\ \mathbf{Y}_B &= \mathbf{H}\mathbf{S}_A + \mathbf{N}_B \\ \mathbf{Y}_E &= \sqrt{a}\mathbf{A}\mathbf{S}_A + \sqrt{b}\mathbf{B}\mathbf{S}_B + \mathbf{N}_E \end{aligned} \tag{62}$$

where $\mathbf{S}_A=[\mathbf{s}_A(1),\ldots,\mathbf{s}_A(K_2)]$ and $\mathbf{s}_A(t)\sim\mathcal{CN}(0,\frac{P_A}{N_A}\mathbf{I})$. Similarly $\mathbf{S}_B=[\mathbf{s}_B(1),\ldots,\mathbf{s}_B(K_2)]$ and $\mathbf{s}_B(t)\sim\mathcal{CN}(0,\frac{P_B}{N_B}\mathbf{I})$. Note that all information symbols from Alice and Bob are i.i.d. The vectorized forms of (62) are

$$\begin{aligned} \mathbf{y}_A &= (\mathbf{I}_{K_2}\otimes\mathbf{H}^T)\bar{\mathbf{s}}_B + \mathbf{n}_A \\ \mathbf{y}_B &= (\mathbf{I}_{K_2}\otimes\mathbf{H})\bar{\mathbf{s}}_A + \mathbf{n}_B \\ \mathbf{y}_E &= \sqrt{a}(\mathbf{I}_{K_2}\otimes\mathbf{A})\bar{\mathbf{s}}_A + \sqrt{b}(\mathbf{I}_{K_2}\otimes\mathbf{B})\bar{\mathbf{s}}_B + \mathbf{n}_E \end{aligned} \tag{63}$$

where both $\bar{\mathbf{s}}_A$ and $\bar{\mathbf{s}}_B$ are assumed to be independent of all channel parameters. Conditional on the MMSE channel estimation in phase 1, an achievable secrecy rate in phase 2 by the two-way wiretap channel is (e.g., see [42]):

$$\begin{aligned} \mathcal{R}_{two} = \frac{1}{K_2}\big(& I(\bar{\mathbf{s}}_B;\mathbf{y}_A|\hat{\mathbf{h}}_A) + I(\bar{\mathbf{s}}_A;\mathbf{y}_B|\hat{\mathbf{h}}_B) \\ & - I(\bar{\mathbf{s}}_A,\bar{\mathbf{s}}_B;\mathbf{y}_E|\hat{\mathbf{a}},\hat{\mathbf{b}})\big)^+ \end{aligned} \tag{64}$$

The following analysis is similar to the previous section, for which we will only provide the key steps and results.

From (50) and (54), we already know a pair of lower and upper bounds on $I(\bar{\mathbf{s}}_A;\mathbf{y}_B|\hat{\mathbf{h}}_B)$. To show a similar pair of lower and upper bounds on $I(\bar{\mathbf{s}}_B;\mathbf{y}_A|\hat{\mathbf{h}}_A)$, we let $\hat{\mathbf{H}}_A = ivec(\hat{\mathbf{h}}_A)$. One

can verify that

$$I(\bar{\mathbf{s}}_B;\mathbf{y}_A|\hat{\mathbf{h}}_A) \geq K_2\mathcal{E}\left\{\log\left|\mathbf{I}_{N_A}+\frac{P_B/N_B}{1+\frac{\sigma^2 P_B}{1+\sigma^2 T_1 P_B/N_B}}\hat{\mathbf{H}}_A^T\hat{\mathbf{H}}_A^*\right|\right\} \triangleq \mathcal{R}_A^- \tag{65}$$

and

$$I(\bar{\mathbf{s}}_B;\mathbf{y}_A|\hat{\mathbf{h}}_A) \leq \mathcal{R}_A^- + N_A\log\left(\frac{\left(1+\frac{P_B}{1+K_1P_B/N_B}\right)^{K_2}}{\left(1+\frac{P_B/N_B}{1+K_1P_B/N_B}e_B\right)^{t_B}}\right) \triangleq \mathcal{R}_A^+ \tag{66}$$

where $e_B=\exp(\frac{1}{t_B}\sum_{j=1}^{t_B}\sum_{k=1}^{r_B-j}\frac{1}{k}-\gamma)$, $t_B=\min\{N_B, K_2\}$ and $r_B=\max\{N_B,K_2\}$.

For $I(\bar{\mathbf{s}}_A,\bar{\mathbf{s}}_B;\mathbf{y}_E|\hat{\mathbf{a}},\hat{\mathbf{b}})$, we use $\hat{\mathbf{B}}=ivec(\hat{\mathbf{b}})$ (similar to $\hat{\mathbf{A}}$). One can verify that $\mathbf{K}_{\mathbf{y}_E|\hat{\mathbf{a}},\hat{\mathbf{b}}}=\frac{P_A}{N_A}(\mathbf{I}_{K_2}\otimes\hat{\mathbf{A}}\hat{\mathbf{A}}^H)+\frac{P_B}{N_B}(\mathbf{I}_{K_2}\otimes\hat{\mathbf{B}}\hat{\mathbf{B}}^H)+\mathbf{K}_{EA}+\mathbf{K}_{EB}+\mathbf{I}_{N_EK_2}$ where $\mathbf{K}_{EA}=\mathcal{E}\{(\mathbf{I}_{K_2}\otimes\Delta\mathbf{A})\bar{\mathbf{s}}_A\bar{\mathbf{s}}_A^H(\mathbf{I}_{K_2}\otimes\Delta\mathbf{A})^H\}=\sigma_{EA}^2P_A\mathbf{I}_{N_EK_2}$ and $\mathbf{K}_{EB}=\mathcal{E}\{(\mathbf{I}_{K_2}\otimes\Delta\mathbf{B})\bar{\mathbf{s}}_B\bar{\mathbf{s}}_B^H(\mathbf{I}_{K_2}\otimes\Delta\mathbf{B})^H\}=\sigma_{EB}^2P_B\mathbf{I}_{N_EK_2}$. Also note that $\mathbf{y}_E=(\mathbf{S}_A^T\otimes\mathbf{I}_{N_E})\mathbf{h}_{EA}+(\mathbf{S}_B^T\otimes\mathbf{I}_{N_E})\mathbf{h}_{EB}+\mathbf{n}_E$. Then,

$$\begin{aligned}
&I(\bar{\mathbf{s}}_A,\bar{\mathbf{s}}_B;\mathbf{y}_E|\hat{\mathbf{a}},\hat{\mathbf{b}})\\
&=h(\mathbf{y}_E|\hat{\mathbf{a}},\hat{\mathbf{b}})-h(\mathbf{y}_E|\hat{\mathbf{a}},\hat{\mathbf{b}},\bar{\mathbf{s}}_A,\bar{\mathbf{s}}_B)\\
&\leq\mathcal{E}\{\log[(\pi e)^{K_2N_E}|\mathbf{K}_{\mathbf{y}_E|\hat{\mathbf{a}},\hat{\mathbf{b}}}|]\}-h(\mathbf{y}_E|\hat{\mathbf{a}},\hat{\mathbf{b}},\bar{\mathbf{s}}_A,\bar{\mathbf{s}}_B)\\
&=\mathcal{E}\{\log|\mathbf{K}_{\mathbf{y}_E|\hat{\mathbf{a}},\hat{\mathbf{b}}}|\}-\mathcal{E}\{\log|\sigma_{EA}^2(\mathbf{S}_A^T\mathbf{S}_A^*\otimes\mathbf{I}_{N_E})\\
&\quad+\sigma_{EB}^2(\mathbf{S}_B^T\mathbf{S}_B^*\otimes\mathbf{I}_{N_E})+\mathbf{I}_{N_EK_2}|\}\\
&=K_2\mathcal{E}\left\{\log\left|\frac{P_A}{N_A}\hat{\mathbf{A}}\hat{\mathbf{A}}^H+\frac{P_B}{N_B}\hat{\mathbf{B}}\hat{\mathbf{B}}^H+(1+P_A\sigma_{EA}^2\right.\right.\\
&\qquad\left.\left.+\ P_B\sigma_{EB}^2)\mathbf{I}_{N_E}\right|\right\}\\
&\quad-N_E\mathcal{E}\{\log|\sigma_{EA}^2\mathbf{S}_A^T\mathbf{S}_A^*+\sigma_{EB}^2\mathbf{S}_B^T\mathbf{S}_B^*+\mathbf{I}_{K_2}|\}
\end{aligned}\tag{67}$$

Define $\mathbf{S}_{AB}=[\check{\mathbf{S}}_A^T,\check{\mathbf{S}}_B^T]\in\mathbb{C}^{K_2\times(N_A+N_B)}$ where $\mathbf{S}_A=\frac{P_A}{N_A}\check{\mathbf{S}}_A$ and $\mathbf{S}_B=\frac{P_B}{N_B}\check{\mathbf{S}}_B$. Define $\mathbf{T}=diag\{\sigma_{EA}^2\frac{P_A}{N_A}\mathbf{I}_{N_A},\sigma_{EB}^2\frac{P_B}{N_B}\mathbf{I}_{N_B}\}$. Then we can rewrite the last term from (67) as $\mathcal{E}\{\log|\sigma_{EA}^2\mathbf{S}_A^T\mathbf{S}_A^*+\sigma_{EB}^2\mathbf{S}_B^T\mathbf{S}_B^*+\mathbf{I}_{K_2}|\}=\mathcal{E}\{\log|\mathbf{I}_{K_2}+\mathbf{S}_{AB}\mathbf{T}\mathbf{S}_{AB}^H|\}$.

For $K_2<N_A+N_B$, we have

$$\begin{aligned}
&\mathcal{E}\{\log|\mathbf{I}_{K_2}+\mathbf{S}_{AB}\mathbf{T}\mathbf{S}_{AB}^H|\}\\
&\geq K_2\mathcal{E}\left\{\log\left(1+|\mathbf{S}_{AB}\mathbf{T}\mathbf{S}_{AB}^H|^{\frac{1}{K_2}}\right)\right\}\\
&=K_2\mathcal{E}\left\{\log\left(1+\exp\left(\frac{1}{K_2}\ln|\mathbf{S}_{AB}\mathbf{T}\mathbf{S}_{AB}^H|\right)\right)\right\}\\
&\geq K_2\mathcal{E}\left\{\log\left(1+\exp\left(\frac{1}{K_2}\ln\sigma_{\min}^{2K_2}|\mathbf{S}_{AB}\mathbf{S}_{AB}^H|\right)\right)\right\}\\
&\geq K_2\log\left(1+\sigma_{\min}^2e_{E1}\right)
\end{aligned}\tag{68}$$

where $e_{E1}=\exp(\frac{1}{K_2}\sum_{j=1}^{K_2}\sum_{k=1}^{N_A+N_B-j}\frac{1}{k}-\gamma)$. The second inequality in (68) is from the fact (e.g., see [43, Th. 3]) that $|\mathbf{S}_{AB}\mathbf{T}\mathbf{S}_{AB}^H|\geq\sigma_{\min}^{2K_2}|\mathbf{S}_{AB}\mathbf{S}_{AB}^H|$ where $\sigma_{\min}^2=\min\{\sigma_{EA}^2\frac{P_A}{N_A},\sigma_{EB}^2\frac{P_B}{N_B}\}$. Similarly, for $K_2\geq N_A+N_B$, we have

$$\begin{aligned}
&\mathcal{E}\{\log|\mathbf{I}+\mathbf{S}_{AB}\mathbf{T}\mathbf{S}_{AB}^H|\}\\
&=\mathcal{E}\{\log|\mathbf{I}+\mathbf{T}\mathbf{S}_{AB}^H\mathbf{S}_{AB}|\}\\
&\geq(N_A+N_B)\mathcal{E}\left\{\log\left(1+|\mathbf{T}|^{\frac{1}{N_A+N_B}}\right.\right.\\
&\qquad\left.\left.\times\exp\left(\frac{1}{N_A+N_B}\ln|\mathbf{S}_{AB}^H\mathbf{S}_{AB}|\right)\right)\right\}\\
&\geq(N_A+N_B)\log\left(1+|\mathbf{T}|^{\frac{1}{N_A+N_B}}e_{E2}\right)
\end{aligned}\tag{69}$$

where $e_{E2}=\exp(\frac{1}{N_A+N_B}\sum_{j=1}^{N_A+N_B}\sum_{k=1}^{K_2-j}\frac{1}{k}-\gamma)$ Therefore, using (68) and (69), we have from (67) that

$$\begin{aligned}
&I(\bar{\mathbf{s}}_A,\bar{\mathbf{s}}_B;\mathbf{y}_E|\hat{\mathbf{a}},\hat{\mathbf{b}})\\
&\leq K_2\mathcal{E}\left\{\log\left|\frac{\frac{P_A}{N_A}\hat{\mathbf{A}}\hat{\mathbf{A}}^H+\frac{P_B}{N_B}\hat{\mathbf{B}}\hat{\mathbf{B}}^H}{1+P_A\sigma_{EA}^2+P_B\sigma_{EB}^2}+\mathbf{I}\right|\right\}\\
&+\begin{cases}K_2N_E\log\left(\dfrac{1+P_A\sigma_{EA}^2+P_B\sigma_{EB}^2}{1+\sigma_{\min}^2e_{E1}}\right), & K_2\leq N_A+N_B\\ N_E\log\left(\dfrac{\left(1+P_A\sigma_{EA}^2+P_B\sigma_{EB}^2\right)^{K_2}}{\left(1+|\mathbf{T}|^{\frac{1}{N_A+N_B}}e_{E2}\right)^{N_A+N_B}}\right), & K_2>N_A+N_B\end{cases}\\
&\triangleq\mathcal{R}_{E,t}^+
\end{aligned}\tag{70}$$

One can also verify $I(\bar{\mathbf{s}}_A,\bar{\mathbf{s}}_B;\mathbf{y}_E|\hat{\mathbf{a}},\hat{\mathbf{b}})\geq K_2\mathcal{E}\{\log|\frac{\frac{P_A}{N_A}\hat{\mathbf{A}}\hat{\mathbf{A}}^H+\frac{P_B}{N_B}\hat{\mathbf{B}}\hat{\mathbf{B}}^H}{1+P_A\sigma_{EA}^2+P_B\sigma_{EB}^2}+\mathbf{I}|\}\triangleq\mathcal{R}_{E,t}^-$ which is the first term in (70).

When $P_A=P_B=P\to\infty$, we have $\sigma_{EA}^2\to\frac{bN_A}{aN_B+bN_A}$, $\sigma_{EB}^2\to\frac{aN_B}{aN_B+bN_A}$, $\sigma_A^2\to0$, $\sigma_B^2\to0$, $\mathcal{E}\{\hat{a}_i\hat{a}_i^*\}\to\frac{aN_B}{aN_B+bN_A}$, $\mathcal{E}\{\hat{b}_i\hat{b}_i^*\}\to\frac{bN_A}{aN_B+bN_A}$, $\mathcal{E}\{\hat{h}_{A,i}\hat{h}_{A,i}^*\}\to1$, $\mathcal{E}\{\hat{h}_{B,i}\hat{h}_{B,i}^*\}\to1$, $\sigma_{\min}^2=P\min\{\frac{\sigma_{EA}^2}{N_A},\frac{\sigma_{EB}^2}{N_B}\}$ and $|\mathbf{T}|^{\frac{1}{N_A+N_B}}=P((\frac{\sigma_{EA}^2}{N_A})^{N_A}(\frac{\sigma_{EB}^2}{N_B})^{N_B})^{1/(N_A+N_B)}$.

Then, similar to (57), we have

$$\lim_{P\to\infty}\frac{\mathcal{R}_A^-}{\log P}=\lim_{P\to\infty}\frac{\mathcal{R}_A^+}{\log P}=K_2\min\{N_A,N_B\} \tag{71}$$

One can also verify that

$$\lim_{P\to\infty}\frac{\mathcal{R}_{E,t}^-}{\log P}=0 \tag{72}$$

and

$$\lim_{P\to\infty}\frac{\mathcal{R}_{E,t}^+}{\log P}=\begin{cases}0, & K_2\leq N_A+N_B\\ N_E(K_2-N_A-N_B), & K_2>N_A+N_B\end{cases} \tag{73}$$

Now applying (57), (71), (72) and (73), and using $\mathcal{R}_{two}^+\triangleq\frac{1}{K_2}[\mathcal{R}_A^++\mathcal{R}_B^+-\mathcal{R}_{E,t}^-]^+$ and $\mathcal{R}_{two}^-\triangleq\frac{1}{K_2}[\mathcal{R}_A^-+$

$\mathcal{R}_B^- - \mathcal{R}_{E,t}^+]^+$ as upper and lower bounds on $\mathcal{R}_{two}$, we have

$$\lim_{P\to\infty}\frac{\mathcal{R}_{two}^-}{\log P} = \begin{cases} 2\min\{N_A, N_B\}, & K_2 \le N_A + N_B \\ \left(2\min\{N_A, N_B\} - \dfrac{N_E}{K_2}(K_2 - N_A - N_B)\right)^+, & K_2 > N_A + N_B \end{cases} \quad (74)$$

and

$$\lim_{P\to\infty}\frac{\mathcal{R}_{two}^+}{\log P} = 2\min\{N_A, N_B\} \quad (75)$$

We see that if $K_2 \le N_A + N_B$, $\lim_{P\to\infty}\frac{\mathcal{R}_{two}}{\log P} = 2\min\{N_A, N_B\}$ which equals the degrees of freedom of the full-duplex channel between Alice and Bob. And if $K_2 > N_A + N_B$, the above lower bound on $\lim_{P\to\infty}\frac{\mathcal{R}_{two}}{\log P}$ decreases linearly as N_E increases. We see an advantage of two-way information transmission over one-way information transmission.

B. Eve Uses Blind Detection With Zero Knowledge of Its CSI

Now we reconsider the case of one-way information transmission from Alice to Bob in the second phase but assume that Eve performs a blind detection of the information transmitted from Alice. For the blind detection shown next, we also assume that $K_2 > N_A$ and Eve's knowledge of its CSI matrix $\sqrt{a}\mathbf{A} \in \mathbb{C}^{N_E\times K_2}$ is zero. (The two-way information transmission between Alice and Bob in either half-duplex or ideal full-duplex can be treated similarly. For the case of $K_2 \le N_A$, Eve cannot receive any information from the users due to its unknown CSI.)

The signal received by Eve during information transmission from Alice over K_2 sampling intervals is

$$\mathbf{Y}_E = \sqrt{a}\mathbf{A}\mathbf{S}_A + \mathbf{N_E} \quad (76)$$

where the elements in $\mathbf{S}_A \in \mathbb{C}^{N_A\times K_2}$ are assumed to be independently chosen from a known constellation $\mathbb{S}_N$ with size N. Assume that Eve performs the blind detection as follows:

$$\left(\hat{\mathbf{S}}, \hat{\mathbf{A}}\right) = \underset{\mathbf{S}\in\mathbb{S}_N^{N_A\times K_2}, \sqrt{a}\mathbf{A}\in\mathbb{C}^{N_E\times K_2}}{\operatorname{argmin}} \|\mathbf{Y}_E - \sqrt{a}\mathbf{A}\mathbf{S}\|_F^2. \quad (77)$$

Given any $\mathbf{S}$, the optimal $\sqrt{a}\mathbf{A}$ is $\mathbf{Y}_E\mathbf{S}^H(\mathbf{S}\mathbf{S}^H)^{-1}$. Then, the above problem reduces to the following (an issue of uniqueness will be addressed later)

$$\hat{\mathbf{S}} = \underset{\mathbf{S}\in\mathbb{S}_N^{N_A\times K_2}}{\operatorname{argmin}} \|\mathbf{Y}_E - \mathbf{Y}_E\mathbf{S}^H\left(\mathbf{S}\mathbf{S}^H\right)^{-1}\mathbf{S}\|_F^2, \quad (78)$$

or equivalently $\hat{\mathbf{S}} = \operatorname{argmax}_{\mathbf{S}\in\mathbb{S}_N^{N_A\times K_2}} f(\mathbf{s})$, where $f(\mathbf{s}) = \mathrm{Tr}(\mathbf{S}^H(\mathbf{S}\mathbf{S}^H)^{-1}\mathbf{S}\mathbf{Z})$, $\mathbf{s} = \mathrm{vec}(\mathbf{S})$ and $\mathbf{Z} = \mathbf{Y}_E^H\mathbf{Y}_E$. The above problem is computationally expensive. But we assume that Eve is able to afford it.

Assume that the solution $\hat{\mathbf{S}}$ of the above problem is so close to the actual information matrix $\mathbf{S}_0$ that $f(\mathbf{s})$ can be replaced by its 2nd-order Taylor's series expansion (which is conservative for Alice and Bob or equivalently optimistic for Eve). Then $\hat{\mathbf{s}} = \mathrm{vec}(\hat{\mathbf{S}})$ has the following properties

$$\nabla_{\mathbf{s}}\tilde{f}(\mathbf{s})\,|_{\mathbf{s}=\hat{\mathbf{s}}} = \mathbf{0}, \quad (79)$$

$$\nabla_{\mathbf{s}^*}\tilde{f}(\mathbf{s})\,|_{\mathbf{s}=\hat{\mathbf{s}}} = \mathbf{0}, \quad (80)$$

where $\tilde{f}(\mathbf{s})$ is the second-order Taylor series expansion [44] of $f(\mathbf{s})$ around $\mathbf{s}_0 = \mathrm{vec}(\mathbf{S}_0)$, i.e.,

$$\begin{aligned}\tilde{f}(\mathbf{s}) = {} & f(\mathbf{s}_0) + \nabla_{\mathbf{s}}^T f(\mathbf{s})\,|_{\mathbf{s}=\mathbf{s}_0}(\mathbf{s}-\mathbf{s}_0) + \nabla_{\mathbf{s}^*}^T f(\mathbf{s})\,|_{\mathbf{s}=\mathbf{s}_0}(\mathbf{s}-\mathbf{s}_0)^* \\ & + \frac{1}{2}\Big[(\mathbf{s}-\mathbf{s}_0)^H\mathbf{H}_{ss}(\mathbf{s}-\mathbf{s}_0) + (\mathbf{s}-\mathbf{s}_0)^H\mathbf{H}_{s^*s}(\mathbf{s}-\mathbf{s}_0)^* \\ & + (\mathbf{s}-\mathbf{s}_0)^T\mathbf{H}_{ss^*}(\mathbf{s}-\mathbf{s}_0) + (\mathbf{s}-\mathbf{s}_0)^T\mathbf{H}_{s^*s^*}(\mathbf{s}-\mathbf{s}_0)^*\Big].\end{aligned} \quad (81)$$

which involves the Hessian matrices: $\mathbf{H}_{ss} = \frac{\partial}{\partial\mathbf{s}}(\nabla_{\mathbf{s}^*}f)|_{\mathbf{s}=\mathbf{s}_0}$, $\mathbf{H}_{s^*s} = \frac{\partial}{\partial\mathbf{s}^*}(\nabla_{\mathbf{s}^*}f)|_{\mathbf{s}=\mathbf{s}_0}$, $\mathbf{H}_{ss^*} = \frac{\partial}{\partial\mathbf{s}}(\nabla_{\mathbf{s}}f)|_{\mathbf{s}=\mathbf{s}_0}$, $\mathbf{H}_{s^*s^*} = \frac{\partial}{\partial\mathbf{s}^*}(\nabla_{\mathbf{s}}f)|_{\mathbf{s}=\mathbf{s}_0}$. Subject to uniqueness of solution, solving (79) and (80) results in the following [44]

$$\hat{\mathbf{s}} - \mathbf{s}_0 = \left(\mathbf{H}_{ss} - \mathbf{H}_{s^*s}\mathbf{H}_{ss}^{-T}\mathbf{H}_{s^*s}^H\right)^{-1}\left(\mathbf{H}_{s^*s}\mathbf{H}_{ss}^{-T}\nabla_{\mathbf{s}}f - \nabla_{\mathbf{s}^*}f\right). \quad (82)$$

Furthermore,

$$\nabla_{\mathbf{s}^*}f = \left(\mathbf{Z}\left(\mathbf{I} - \mathbf{S}^H\left(\mathbf{S}\mathbf{S}^H\right)^{-1}\mathbf{S}\right)\right)^T \otimes \left(\mathbf{S}\mathbf{S}^H\right)^{-1}\mathbf{s}, \quad (83)$$

$$\nabla_{\mathbf{s}}f = (\nabla_{\mathbf{s}}f)^*, \quad (84)$$

$$\begin{aligned}\mathbf{H}_{ss} = {} & \Big[\mathbf{Z} - \mathbf{Z}\mathbf{S}^H\left(\mathbf{S}\mathbf{S}^H\right)^{-1}\mathbf{S} - \mathbf{S}^H\left(\mathbf{S}\mathbf{S}^H\right)^{-1}\mathbf{S}\mathbf{Z} \\ & + \mathbf{S}^H\left(\mathbf{S}\mathbf{S}^H\right)^{-1}\mathbf{S}\mathbf{Z}\mathbf{S}^H\left(\mathbf{S}\mathbf{S}^H\right)^{-1}\mathbf{S}\Big]^T \otimes \left(\mathbf{S}\mathbf{S}^H\right)^{-1} \\ & + \left(\mathbf{S}^H\left(\mathbf{S}\mathbf{S}^H\right)^{-1}\mathbf{S} - \mathbf{I}\right)^T \\ & \otimes \mathbf{S}^H\left(\mathbf{S}\mathbf{S}^H\right)^{-1}\mathbf{S}\mathbf{Z}\mathbf{S}^H\left(\mathbf{S}\mathbf{S}^H\right)^{-1}\mathbf{S},\end{aligned} \quad (85)$$

$$\begin{aligned}\mathbf{H}_{s^*s} = {} & \Big[\Big(\left(\left(\mathbf{S}\mathbf{S}^H\right)^{-1}\mathbf{S}\mathbf{Z}\right)\left(\mathbf{S}^H\left(\mathbf{S}\mathbf{S}^H\right)^{-1}\mathbf{S} - \mathbf{I}\right)\Big)^T \\ & \otimes \left(\mathbf{S}\mathbf{S}^H\right)^{-1}\mathbf{S} + \Big[\left(\mathbf{S}\mathbf{S}^H\right)^{-1}\mathbf{S}\Big]^T \otimes \left(\left(\mathbf{S}\mathbf{S}^H\right)^{-1}\mathbf{S}\mathbf{Z}\right) \\ & \times \left(\mathbf{S}^H\left(\mathbf{S}\mathbf{S}^H\right)^{-1}\mathbf{S} - \mathbf{I}\right)\Big]\boldsymbol{\Pi}\end{aligned} \quad (86)$$

where $\boldsymbol{\Pi}$ is a permutation matrix with

$$\boldsymbol{\Pi}_{i,j} = \begin{cases} 1 & j = \left((i-1)_{\bmod N_A}\right)K_2 + \lfloor (i-1)/N_A \rfloor \\ 0 & \text{else} \end{cases} \quad (87)$$

where $a_{\bmod b}$ denotes the remainder of the division of a by b. For more details about complex derivatives, please refer to [44].

Because of the blind nature, $\mathbf{H}_{ss}$ is always rank deficient by N_A^2. To remove the ambiguity, we can treat the first N_A of the transmitted vectors from Alice as known, which is equivalent to removing N_A^2 corresponding rows and N_A^2 corresponding columns from each of $\mathbf{H}_{ss}$ and $\mathbf{H}_{s^*s}$, and removing N_A^2 corresponding elements from each of $\nabla_{\mathbf{s}}f$ and $\nabla_{\mathbf{s}^*}f$. This results in

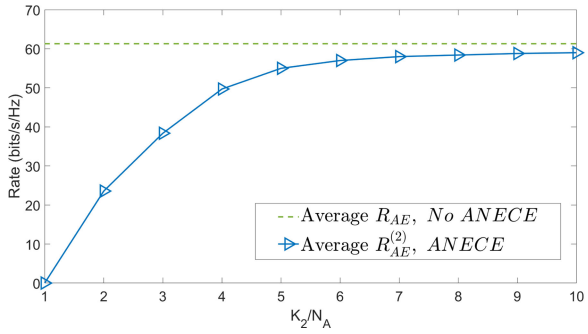


Fig. 6. Eve's rates vs K_2/N_A for known or unknown CSI at Eve. With ANECE, Eve does not know its CSI. Otherwise, Eve does.

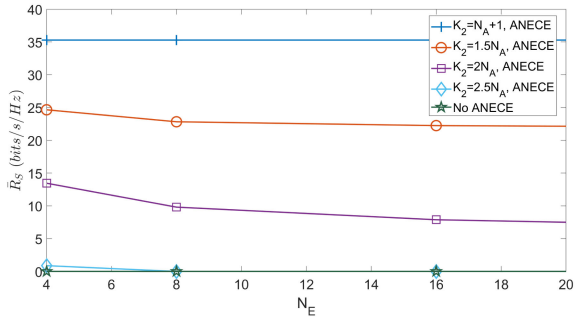


Fig. 7. $\bar{R}_S$ vs. N_E for known or unknown CSI at Eve. With ANECE, Eve does not know its CSI. Otherwise, Eve does.

$\bar{\mathbf{H}}_{ss}, \bar{\mathbf{H}}_{s^*s}, \bar{\nabla}_{\mathbf{s}} f$ and $\bar{\nabla}_{\mathbf{s}^*} f$, respectively. Hence the MSE matrix $\bar{\mathbf{M}}$ of the remaining unknown parameters can be formed as

$$\bar{\mathbf{M}} = \mathcal{E}\left[\left(\overline{\hat{\mathbf{s}} - \mathbf{s}_0}\right)\left(\overline{\hat{\mathbf{s}} - \mathbf{s}_0}\right)^H\right], \tag{88}$$

where $\overline{\hat{\mathbf{s}} - \mathbf{s}_0}$ is the approximation of errors in the vector of all $N_A(K_2 - N_A)$ remaining symbols and

$$\overline{\hat{\mathbf{s}} - \mathbf{s}_0} = \left(\bar{\mathbf{H}}_{ss} - \bar{\mathbf{H}}_{s^*s}\bar{\mathbf{H}}_{ss}^{-T}\bar{\mathbf{H}}_{s^*s}^{H}\right)^{-1}\left(\bar{\mathbf{H}}_{s^*s}\bar{\mathbf{H}}_{ss}^{-T}\bar{\nabla}_{\mathbf{s}} f - \bar{\nabla}_{\mathbf{s}^*} f\right) \tag{89}$$

Finally, for $K_2 > N_A$, Eve's effective rate (with the information in the first N_A vectors of $\mathbf{s}(k)$ removed) can be approximated as

$$R_{AE}^{(2)} = \frac{1}{K_2}\left(\log|\bar{\mathbf{Q}}| - \log|\bar{\mathbf{M}}|\right), \tag{90}$$

where $\bar{\mathbf{Q}}$ is the covariance matrix of the vector of all remaining symbols.

To evaluate $R_{AE}^{(2)}$, one has to specify the actual constellation $\mathbb{S}_N$ of each symbol in $\mathbf{S}$, compute $\overline{\hat{\mathbf{s}} - \mathbf{s}_0}$ for each actual realization of $\mathbf{S}_0$ according to (89), and then obtain a sample averaged version of $\bar{\mathbf{M}}$ in (88). Each of the realizations of $\mathbf{S}_0$ should be coupled with an independent realization of the channel matrix $\mathbf{A}$ and the noise matrix $\mathbf{N}_E$. With the final sample-averaged versions of $\bar{\mathbf{Q}}$ and $\bar{\mathbf{M}}$, $R_{AE}^{(2)}$ in (90) can be obtained.

For the next two plots, we assume that $\mathbb{S}_N$ is 4-QAM,[3] 100 random realizations of $\mathbf{S}_0$, $\mathbf{A}$ and $\mathbf{N}_E$ are used in computing $R_{AE}^{(2)}$. Also, during information transmission from Alice to Bob, $P_A = 30$ dB (and $P_B = 0$). In this case, due to high power, we expect the Taylor's series expansion applied in our derivation is accurate.

Fig. 6 shows $R_{AE}^{(2)}$ versus K_2/N_A where $N_A = N_B = 4$ and $N_E = 8$. We see that only when K_2 becomes much larger than N_A, $R_{AE}^{(2)}$ approaches R_{AE}. Note that $R_{AE}^{(2)}$ is based on unknown CSI at Eve and blind detection at Eve while R_{AE} is based on the assumption that Eve knows its CSI perfectly.

Fig. 7 shows the averaged secret rate $\bar{R}_S = (\mathcal{E}[R_{AB} - R_{AE}^{(2)}])^+$ versus N_E where $N_A = N_B = 4$. (The curves in this figure were zoomed in for the range of N_E from 4 to 20. The actually computed points were at $N_E = 4, 8, 16, 32$.) In this case, $(\mathcal{E}[R_{AB} - R_{AE}])^+$ is zero for all values of N_E. But when Eve is blind to its CSI (caused by ANECE), the secrecy rates become substantial. In this case, we also see that for given $K_2 > N_A$ the secrecy rate decreases as the number of antennas on Eve increases.

The above results in this subsection complement the analytical insights shown in the previous subsection (e.g., see (60)). Due to different assumptions, we cannot make a precise comparison between (60) and Figs. 6 and 7 while the general trends predicted in both cases are somewhat consistent. An additional discussion of the blind detection where Eve uses a partial knowledge of its CSI from phase 1 is shown in Appendix B.

[3]For higher order constellations, the simulation became too slow and consuming.

IV. Conclusion

In this paper, we have investigated the secrecy performance of a full-duplex MIMOME network in some important scenarios. In the first part of this paper, we studied how to optimize the jamming powers from both Alice and Bob when Eve's CSI is unknown to Alice and Bob but Eve knows all CSI. To handle Eve's CSI being unknown to Alice and Bob, we focused on Eve at the most harmful location and adopted the large matrix theory that yields a hardened secret rate for any large number of antennas on Eve. With the optimized powers, we revealed a significant improvement in terms of the maximum tolerable number of antennas on Eve. In the second part of this paper, we analyzed the full-duplex MIMOME network subject to the application of anti-eavesdropping channel estimation (ANECE) in a two-phase scheme. Assuming that a statistical model of CSI anywhere is known everywhere, we derived lower and upper bounds on the secure degrees of freedom of the network, which reveal clearly how the number of antennas on Eve affect these bounds. In particular, for $1 \le K_2 \le N_A$ in one-way information transmission or $1 \le K_2 \le N_A + N_B$ in two-way information transmission, the lower and upper bounds coincide and equal to those of the channel capacity between Alice and Bob. Furthermore, assuming that Eve does not have any prior knowledge of its CSI but uses blind detection in phase 2 of the two-phase scheme, we provided and illustrated an approximate secrecy rate for $K_2 > N_A$ in one-way information transmission. But the exact secrecy rate of the full-duplex MIMOME network with

ANECE for K_2 larger than the total number of transmitting antennas still remains elusive. Nevertheless, the contributions shown in this paper are significant additions to our previous works shown in [30] and [35], which expands the understanding of full-duplex radio for wireless network security.

Appendix A
Proof of Lemma 1

The following proof is a simple digest from [36] that is useful to help readers to understand Lemma 1 more easily. The Shannon transform of the distribution of a random variable X with parameter γ is defined as

$$\mathcal{V}_X(\gamma) = \mathcal{E}_X[\log(1+\gamma X)], \tag{91}$$

and the $\eta-$transform of the distribution of X with parameter γ is defined as

$$\eta_X(\gamma) = \mathcal{E}_X\left[\frac{1}{1+\gamma X}\right], \tag{92}$$

where $\gamma \geq 0$. The empirical cumulative distribution function of the eigenvalues of an $n \times n$ random non-negative-definite Hermitian matrix $\mathbf{A}$ is defined as

$$F^n_{\mathbf{A}}(x) = \frac{1}{n}\sum_{i=1}^{n} 1\{\lambda_i(\mathbf{A}) \leq x\} \tag{93}$$

where $\lambda_1(\mathbf{A}), \ldots, \lambda_n(\mathbf{A})$ are the eigenvalues of $\mathbf{A}$, and $1\{.\}$ is the indicator function. When $F^n_{\mathbf{A}}(x)$ converges as $n \to \infty$, the corresponding limit is denoted by $F_{\mathbf{A}}(x)$.

It is obvious that

$$\begin{aligned}\frac{1}{n}\log|\mathbf{I}+\gamma\mathbf{A}| &= \frac{1}{n}\sum_{i=1}^{n}\log(1+\gamma\lambda_i(\mathbf{A})) \\ &= \int_0^\infty \log(1+\gamma x) dF^n_{\mathbf{A}}(x),\end{aligned} \tag{94}$$

and if $n \to \infty$ then

$$\frac{1}{n}\log|\mathbf{I}+\gamma\mathbf{A}| \to \int_0^\infty \log(1+\gamma x) dF_{\mathbf{A}}(x) \tag{95}$$

which is the Shannon transform of the eigenvalue distribution of the matrix $\mathbf{A}$ when n is large.

The Shannon transform of the eigenvalue distribution of $\mathbf{\Theta}$ with parameter η is obviously given by (19). And the η-transform of the eigenvalue distribution of $\mathbf{\Theta}$ with parameter x is obviously given by

$$\eta_{\mathbf{\Theta}}(x) = \frac{1}{L_{\mathbf{\Theta}}}\sum_{j=1}^{L_{\mathbf{\Theta}}}\frac{1}{1+x\mathbf{\Theta}_{j,j}}. \tag{96}$$

From Theorem 2.39 in [36], the η-transform of the eigenvalue distribution of $\mathbf{J\Theta J}^H$ with parameter γ, denoted by η here, satisfies

$$\beta = \frac{1-\eta}{1-\eta_{\mathbf{\Theta}}(\gamma\eta)}. \tag{97}$$

Applying (96) to the above equation with $\gamma = 1$ yields

$$1-\eta = \beta\left(1 - \frac{1}{L_{\mathbf{\Theta}}}\sum_{j=1}^{L_{\mathbf{\Theta}}}\frac{1}{1+\eta\mathbf{\Theta}_{j,j}}\right) \tag{98}$$

which reduces to (20). Also from Theorem 2.39 in [36], the Shannon transform of the eigenvalue distribution of $\mathbf{J\Theta J}^H$ with parameter $\gamma = 1$ is

$$\mathcal{V}_{\mathbf{J\Theta J}^H}(1) = \beta\mathcal{V}_{\mathbf{\Theta}}(\eta) - \log(\eta) + (\eta-1)\log(e) \tag{99}$$

which is $\Omega(\beta, \mathbf{\Theta}, \eta)$ in (18). ■

Appendix B
Eve Uses Blind Detection With Partial Knowledge of Its CSI

Now we consider the case where Eve can use its signal in phase 1 to obtain its CSI up to a subspace ambiguity, i.e., in the absence of noise, Eve can obtain from $\mathbf{Y}_E$ as in (44c) the following:

$$\hat{\mathbf{A}} = \mathbf{A} + \mathbf{\Theta}\mathbf{C}_A \tag{100}$$

$$\hat{\mathbf{B}} = \mathbf{B} + \mathbf{\Theta}\mathbf{C}_B \tag{101}$$

where $[\mathbf{C}_A, \mathbf{C}_B] \in \mathbb{C}^{\min\{N_A, N_B\}\times(N_A+N_B)}$ is a known matrix satisfying $[\mathbf{C}_A, \mathbf{C}_B][\mathbf{P}_A^T, \mathbf{P}_B^T]^T = 0$. For convenience and without loss of generality, we assume here $a = b = 1$.

With one-way information transmission from Alice to Bob in phase 2, Eve can now perform a constrained blind detection as follows:

$$\min_{\mathbf{S}\in\mathbb{S}_N^{N_A\times K_2}, \mathbf{A}|\hat{\mathbf{A}}=\mathbf{A}+\mathbf{\Theta}\mathbf{C}_A} \|\mathbf{Y}_E - \mathbf{A}\mathbf{S}\|^2 \tag{102}$$

or equivalently

$$\min_{\mathbf{S}\in\mathbb{S}_N^{N_A\times K_2}, \mathbf{\Theta}} \|\mathbf{Y}_E - (\hat{\mathbf{A}} - \mathbf{\Theta}\mathbf{C}_A)\mathbf{S}\|^2. \tag{103}$$

For any given $\mathbf{S}$, the solution for $\mathbf{\Theta}$ is

$$\mathbf{\Theta} = -(\mathbf{Y}_E - \hat{\mathbf{A}}\mathbf{S})(\mathbf{C}_A\mathbf{S})^H(\mathbf{C}_A\mathbf{S}(\mathbf{C}_A\mathbf{S})^H)^{-1}. \tag{104}$$

Then, the problem of (103) reduces to

$$\min_{\mathbf{S}\in\mathbb{S}_N^{N_A\times K_2}} \|(\mathbf{Y}_E - \hat{\mathbf{A}}\mathbf{S})(\mathbf{I}_{K_2} - \mathbf{P}_{\mathbf{C}_A\mathbf{S}})\|^2 \tag{105}$$

where $\mathbf{P}_{\mathbf{C}_A\mathbf{S}} = (\mathbf{C}_A\mathbf{S})^H(\mathbf{C}_A\mathbf{S}(\mathbf{C}_A\mathbf{S})^H)^{-1}\mathbf{C}_A\mathbf{S}$. The problem of (105) is more complex than (78) due to higher order of the cost function in terms of $\mathbf{S}$. A performance analysis of (105) can be done in a similar way as for (78) but is omitted.

Acknowledgment

The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Office or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

REFERENCES

[1] Y. Zou, J. Zhu, X. Wang, and L. Hanzo, "A survey on wireless security: Technical challenges, recent advances, and future trends," *Proc. IEEE*, vol. 104, no. 9, pp. 1727–1765, Sep. 2016.

[2] M. Koziol, "Wi-Fi gets more secure: Everything you need to know about WPA3," *IEEE Spectrum*, Sep. 2018. [Online]. Available: https://spectrum.ieee.org/tech-talk/telecom/security/everything-you-need-to-know-about-wpa3

[3] M. Bloch and J. Barros, *Physical-Layer Security*. Cambridge, U.K.: Cambridge Press, 2011.

[4] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas Part II: The MIMOME wiretap channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5515–5532, Nov. 2010.

[5] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 4961–4972, Aug. 2011.

[6] A. Khisti, G. Wornell, A. Wiesel, and Y. Eldar, "On the Gaussian MIMO wiretap channel," in *Proc. Inf. Theory, Int. Symp.*, 2007, pp. 2471–2475.

[7] P. K. Gopala, L. Lai, and H. E. Gamal, "On the secrecy capacity of fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4687–4698, Oct. 2008.

[8] Z. Chu, K. Cumanan, Z. Ding, M. Johnston, and S. Y. L. Goff, "Secrecy rate optimizations for a MIMO secrecy channel with a cooperative jammer," *IEEE Trans. Veh. Technol.*, vol. 64, no. 5, pp. 1833–1847, May 2015.

[9] E. Tekin and A. Yener, "The general Gaussian multiple-access and two-way wiretap channels: Achievable rates and cooperative jamming," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2735–2751, Jun. 2008.

[10] T. Liu and S. Shamai, "A note on the secrecy capacity of the multiple-antenna wiretap channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2547–2553, Jun. 2009.

[11] S. A. A. Fakoorian and A. L. Swindlehurst, "Solutions for the MIMO Gaussian wiretap channel with a cooperative jammer," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 5013–5022, Oct. 2011.

[12] Z. Chu, H. Xing, M. Johnston, and S. L. Goff, "Secrecy rate optimizations for a MISO secrecy channel with multiple multiantenna eavesdroppers," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 283–297, Jan. 2016.

[13] X. Zhou and M. R. McKay, "Secure transmission with artificial noise over fading channels: Achievable rate and optimal power allocation," *IEEE Trans. Veh. Technol.*, vol. 59, no. 8, pp. 3831–3842, Oct. 2010.

[14] O. Cepheli, G. Dartmann, G. K. Kurt, and G. Ascheid, "A joint optimization scheme for artificial noise and transmit filter for half and full duplex wireless cyber physical systems," *IEEE Trans. Sustain. Comput.*, vol. 3, no. 2, pp. 126–136, Apr.–Jun. 2018.

[15] M. Masood, A. Ghrayeb, P. Babu, I. Khalil, and M. Hasna, "A minorization–maximization algorithm for maximizing the secrecy rate of the MIMOME wiretap channel," *IEEE Commun. Lett.*, vol. 21, no. 3, pp. 520–523, Mar. 2017.

[16] R. Negi and S. Goel, "Secret communication using artificial noise," in *Proc. IEEE 62nd Veh. Technol. Conf.*, vol. 3, Sep. 2005, pp. 1906–1910.

[17] Z. Li, R. Yates, and W. Trappe, "Achieving secret communication for fast Rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 9, no. 9, pp. 2792–2799, Sep. 2010.

[18] J. Li and A. P. Petropulu, "On ergodic secrecy rate for Gaussian MISO wiretap channels," *IEEE Trans. Wireless Commun.*, vol. 10, no. 4, pp. 1176–1187, Apr. 2011.

[19] A. Hyadi, Z. Rezki, A. Khisti, and M. S. Alouini, "Secure broadcasting with imperfect channel state information at the transmitter," *IEEE Trans. Wireless Commun.*, vol. 15, no. 3, pp. 2215–2230, Mar. 2016.

[20] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2180–2189, Jun. 2008.

[21] W.-C. Liao, T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "QoS-based transmit beamforming in the presence of eavesdroppers: An optimized artificial-noise-aided approach," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1202–1216, Mar. 2011.

[22] A. L. Swindlehurst, "Fixed SINR solutions for the MIMO wiretap channel," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2009, pp. 2437–2440.

[23] A. Mukherjee and A. L. Swindlehurst, "Fixed-rate power allocation strategies for enhanced secrecy in MIMO wiretap channels," in *Proc. IEEE 10th Workshop Signal Process. Adv. Wireless Commun.*, Jun. 2009, pp. 344–348.

[24] S. Liu, Y. Hong, and E. Viterbo, "Artificial noise revisited," *IEEE Trans. Inf. Theory*, vol. 61, no. 7, pp. 3901–3911, Jul. 2015.

[25] G. Zheng, I. Krikidis, J. Li, A. P. Petropulu, and B. Ottersten, "Improving physical layer secrecy using full-duplex jamming receivers," *IEEE Trans. Signal Process.*, vol. 61, no. 20, pp. 4962–4974, Oct. 15, 2013.

[26] Y. Zhou, Z. Z. Xiang, Y. Zhu, and Z. Xue, "Application of full-duplex wireless technique into secure MIMO communication: Achievable secrecy rate based optimization," *IEEE Signal Process. Lett.*, vol. 21, no. 7, pp. 804–808, Jul. 2014.

[27] L. Li, Z. Chen, D. Zhang, and J. Fang, "A full-duplex Bob in the MIMO Gaussian wiretap channel: Scheme and performance," *IEEE Signal Process. Lett.*, vol. 23, no. 1, pp. 107–111, Jan. 2016.

[28] Y. Hua, Q. Zhu, and R. Sohrabi, "Fundamental properties of full-duplex radio for secure wireless communications," 2017, *arXiv:1711.10001*.

[29] L. Chen, Q. Zhu, W. Meng, and Y. Hua, "Fast power allocation for secure communication with full-duplex radio," *IEEE Trans. Signal Process.*, vol. 65, no. 14, pp. 3846–3861, Jul. 15, 2017.

[30] Y. Hua, "Advanced properties of full-duplex radio for securing wireless network," *IEEE Trans. Signal Process.*, vol. 67, no. 1, pp. 120–135, Jan. 2019.

[31] W. Liu, Z. Ding, T. Ratnarajah, and J. Xue, "On ergodic secrecy capacity of random wireless networks with protected zones," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6146–6158, Aug. 2016.

[32] I. Csiszar and J. Korner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 3, pp. 339–348, May 1978.

[33] M. Chiani, M. Z. Win, and H. Shin, "MIMO networks: The effects of interference," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 336–349, Jan. 2010.

[34] H. Shin and J. H. Lee, "Closed-form formulas for ergodic capacity of MIMO Rayleigh fading channels," in *Proc. IEEE Int. Conf. Commun.*, vol. 5, May 2003, pp. 2996–3000.

[35] R. Sohrabi and Y. Hua, "A new look at secrecy capacity of MIMOME using artificial noise from Alice and Bob without knowledge of Eves CSI," in *Proc. IEEE Global Conf. Signal Inf. Process.*, 2018, pp. 1291–1295.

[36] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Found. Trends Commun. Inf. Theory*, vol. 1, no. 1, pp. 1–182, 2004.

[37] Y. Liu, H. Chen, and L. Wang, "Secrecy capacity analysis of artificial noisy MIMO channels—An approach based on ordered eigenvalues of Wishart matrices," *IEEE Trans. Inf. Forensics Secur.*, vol. 12, no. 3, pp. 617–630, Mar. 2017.

[38] A. A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[39] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 2012.

[40] O. Oyman, R. Nabar, H. Bolcskei, and A. Paulraj, "Characterizing the statistical properties of mutual information in MIMO channels," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2784–2795, Nov. 2003.

[41] A. Grant, "Rayleigh fading multi-antenna channels," *EURASIP J. Adv. Signal Process.*, vol. 2002, no. 3, Dec. 2002, Art. no. 260208.

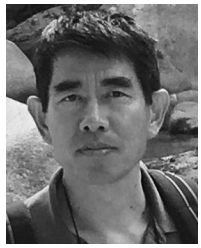
[42] E. Tekin and A. Yener, "The general Gaussian multiple-access and two-way wiretap channels: Achievable rates and cooperative jamming," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2735–2751, Jun. 2008.

[43] S. Jin, X. Gao, and X. You, "On the ergodic capacity of rank-1 Ricean-fading MIMO channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 2, pp. 502–517, Feb. 2007.

[44] K. Kreutz-Delgado, "The complex gradient operator and the CR-calculus," 2009, *arXiv:0906.4835*.

Reza Sohrabi (S'18) received the B.S. degree in electrical engineering from the University of Tehran, Tehran, Iran, in 2014, and the master's and Ph.D. degrees in electrical engineering from the University of California at Riverside, Riverside, CA, USA, in 2015 and 2018, respectively. Since 2019, he has been a Data Scientist with Stitch Fix, Inc. His research interests include wireless communications, data science, and machine learning.

Qiping Zhu (S'15) received the B.E. degree in telecommunications engineering from the Beijing Institute of Petrochemical Technology, Beijing, China, in 2013, and the Ph.D. degree in electrical engineering from the University of California at Riverside, Riverside, CA, USA, in 2019. His research interests include physical layer security, full-duplex radio, MIMO communications, and resource allocation.



Yingbo Hua (S'86–M'88–SM'92–F'02) was born in China. He received the B.S. degree from Southeast University, Nanjing, China, in 1982, and the Ph.D. degree from Syracuse University, Syracuse, NY, USA, in 1988.

He was on the Faculty of the University of Melbourne, Melbourne, VIC, Australia, before he took a Full Professor position in 2001 with the University of California at Riverside, where he has advanced to Professor IX. He has authored and coauthored hundreds of articles in signal processing of wireless networks with a recent focus on wireless network security. He is currently serving the second terms as a Senior Area Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, Associate Editor for the IEEE TRANSACTIONS ON SIGNAL AND INFORMATION PROCESSING OVER NETWORKS, and a member of Steering Committee for the IEEE WIRELESS COMMUNICATION LETTERS. He is a Fellow of AAAS.