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].

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
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 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.

\textbf{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 \(A\) are denoted by \(\text{Tr}(A)\), \(A^H\), \(\text{vec}(A)\), \(A_{i,j}\), and \(A^\ast\), respectively. For a matrix \(X\) and its vectorized version \(x\), \(\text{vec}(X)\) is the inverse operation of \(x = \text{vec}(X)\). A diagonal matrix with elements of \(x\) on its diagonal is \(\text{diag}(x)\). Expectation with respect to a random variable \(x\) is denoted by \(E[x]\).

\begin{equation}
H = U \Sigma V^H,
\end{equation}

where \(U\) and \(V\) are unitary matrices, and \(\Sigma\) is the \(N_B \times N_A\) diagonal matrix that contains the singular values of \(H\) (i.e., \(\sigma_i\), \(i = 1, \ldots, N_B\)) in descending order assuming \(N_A \geq N_B\).
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 $H$ but not of $A$ and $B$, and Eve has the knowledge of all these matrices.

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

$$x_A(k) = V_1 s(k) + V_2 w_A(k),$$

where $k$ is the index of time slot, $V_1$ is the first $r$ columns of $V$, $V_2$ is the last $N_A - r$ columns of $V$, $s(k)$ is Alice’s information vector with the covariance matrix $Q_r$ and $\text{Tr}(Q_r) = P_r$, and $w_A(k)$ is an $(N_A - r) \times 1$ artificial noise vector with distribution $CN(0, P_A N_A^{-1})$. Here, $P_r + P_n = P_A \leq P_A^{\max}$.

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

$$x_B(k) = w_B(k),$$

where $w_B(k)$ is an $N_B \times 1$ artificial noise vector with distribution $CN(0, P_B N_B^{-1})$.

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 at Bob are:

$$y_B(k) = H V_1 s(k) + H V_2 A w_A(k) + \sqrt{P} w_B(k) + n_B,$$

$$y_E(k) = \sqrt{a} A_1 s(k) + \sqrt{a} A_2 A w_A(k) + \sqrt{b} B w_B(k) + n_E,$$

where $[A_1, A_2] = [A V_1, A V_2] = A V$. Since $A_1$ and $A_2$ are linear functions of the Gaussian matrix $A$, they remain Gaussian. Because of the unitary nature of $V$, $A_1$ and $A_2$ are independent of each other, and all elements in them are i.i.d. Gaussian of zero mean and unit variance. The noise vectors $n_B$ and $n_E$ are distributed as $CN(0, I)$. Also note that $\sqrt{P} w_B(k)$ is the residual self-interference originally caused by $w_B(k)$ but is independent of $w_B(k)$ [30].

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

$$R_S = (R_{AB} - R_{AE})^+,$$

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

$$R_{AB} = \log |I + C_B^{-1} H V_1 Q_r V_1^H H^H|,$$

$$R_{AE} = \log |I + a C_E^{-1} A_1 Q_r A_1^H|,$$

where

$$C_B = I + \frac{P_n}{N_A - r} H V_2 V_1^H H^H + \frac{P_B}{N_B} G G^H,$$

$$C_E = I + \frac{P_n}{N_A - r} A_2 A_2^H + \frac{P_B}{N_B} B B^H.$$

Note that since $H V_1 = U_1 \Sigma_1$ and $H V_2 = U_2 \Sigma_2$ are orthogonal to each other where $U_1$ and $U_2$ are the partitions of $U$ similar to those of $V$, and $\Sigma_1$ and $\Sigma_2$ are the corresponding diagonal partitions of $\Sigma$, a sufficient statistics of $s(k)$ at Bob is $U_1^H y_B(k) = \Sigma_1 s(k) + \sqrt{P} U_1 G w_B(k) + U_1 H B(k)$ which shows that the artificial noise from Alice does not affect Bob. Consequently, an equivalent form of $R_{AB}$ is

$$R_{AB} = \log |I + C_{E,1}^{-1} Q_r \Sigma_1^2|,$$

where

$$C_{B,1} = I + \frac{P_B}{N_B} U_1^H G G^H U_1.$$

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

If we optimize $P_n, P_B$ and $Q_r$, 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:

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

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 $Q_r$, 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:

$$R_{AE} = \log |I + J_4 \Theta_3 J_3^H| - \log |I + J_4 \Theta_3 J_3^H|$$

where $J_3 = \frac{1}{\sqrt{N_A}}[A_1, A_2, B]$, $J_4 = \frac{1}{\sqrt{N_E}}[A_2, B]$, and

$$\Theta_3 = N_E \text{diag} \left[ a q_r^T, a P_n N_A - r - 1 N_A - r B N_B - 1 N_B \right],$$

$$\Theta_4 = N_E \text{diag} \left[ a P_n N_A - r - 1 N_A - r B N_B - 1 N_B \right].$$

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

**Lemma 1:** Let $J$ be an $N \times K$ matrix whose entries are i.i.d. complex random variables with variance $\frac{1}{N}$, and $\Theta$ be a
diagonal deterministic matrix. Based on Theorem (2.39) of [36], as \( N, K \to \infty \) with \( \frac{K}{N} \to \beta \), we have

\[
\frac{1}{N} \log \left| I + J \Theta \Theta^H \right| \overset{a.s.}{\to} \Omega (\beta, \Theta, \eta),
\]

where

\[
\Omega (\beta, \Theta, \eta) \triangleq \beta \mathcal{V}_{\Theta} (\eta) - \log (\eta) + (\eta - 1) \log (e),
\]

\[
\mathcal{V}_{\Theta} (\eta) \triangleq \frac{1}{L_\Theta} \sum_{j=1}^{L_\Theta} \log \left( 1 + \eta \Theta_{j,j} \right).
\]

Here, \( \Theta_{j,j} \) is the \( j \)th diagonal element of the diagonal matrix \( \Theta \), \( L_\Theta \) is the number of diagonal elements of \( \Theta \), and \( \eta \geq 0 \) is the solution to the equation

\[
1 - \eta = \frac{\beta \eta L_\Theta}{L_\Theta} \sum_{j=1}^{L_\Theta} \frac{\Theta_{j,j}}{1 + \eta \Theta_{j,j}}.
\]  

**Proof:** The proof is given in Appendix A.

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

\[
\frac{R_{AE}}{N_E} \approx \sum_{j=1}^{L_\Theta} \log \left( 1 + \bar{\eta}_3 (\Theta_{3})_{j,j} \right) - \sum_{j=1}^{L_\Theta} \log \left( 1 + \bar{\eta}_4 (\Theta_{4})_{j,j} \right) + N_E \log \left( \bar{\eta}_4 \right) + N_E (\bar{\eta}_3 - \bar{\eta}_4) \log (e)
\]

\[
\approx \mathcal{R}_{AE}
\]

where for \( i = 3, 4 \), \( \bar{\eta}_i \) is the solution of \( \eta \) to (20) with \( \beta = \bar{\beta}_i \) and \( \Theta = \bar{\Theta}_i \). 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 \( N_A = 2N_B = 8 \), \( r = N_B \), \( Q_r = \frac{P_r}{N_B} I \), \( P_r = P_A = \frac{P_A}{N_B} \). 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), \( \frac{R_{AE}}{N_E} \) becomes a good approximation of \( \frac{R_{AE}}{N_E} \).

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

With \( R_{AB} - R_{AE} \) 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 \( g(x_r) \) to 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.

\[
\min_{r} \min_{x_r} \ g(x_r)
\]

\[
s.t. \ \sum_{i=1}^{r} q_r(i) + P_n \leq P_A^{\text{max}}
\]

\[
q_r(i) \geq 0, \ \forall i = 1, \ldots, r
\]

\[
P_n \geq 0
\]

\[
0 \leq P_B \leq P_B^{\text{max}}.
\]
Algorithm 1: Algorithm for Power Optimization.

Choose proper $\epsilon, \tilde{\eta}_3, \tilde{\eta}_4$, and set $g^{\min} = 0$.

for $r = 1 : N_A$ do
  Initialize $x_r$, satisfying the constraints.
  Set $t = 0$.
  while $||x_r^t - x_r^{t-1}|| > \epsilon$ do
    Solve (25) to get $x_r^{t+1}$.
    Update $\tilde{\eta}_3, \tilde{\eta}_4$ by solving (24) using $x_r^{t+1}$.
    $t = t + 1$.
  end while
  if $g(x_r^r) < g^{\min}$ then
    $g^{\min} = g(x_r^r)$
    $x^{\min} = x_r^r$
  end if
end for

Return $x^{\min}$.

where

$$h^t(x_r) = -\log|C_B + HV_1Q_2V_2^HH^H|$$

$$= L \sum_{j=1}^{L} \log \left( 1 + \tilde{\eta}_4 \left( \Theta_3 \right)_{j,j} \right) + (x_r - x_r^t)^T \nabla_x f^t(x_r = x_r^t),$$

(26)

and $f^t(x_r) = \log|C_B| + \sum_{j=1}^{L} \log(1 + \tilde{\eta}_4 \left( \Theta_3 \right)_{j,j})$. 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 consequence, $H V_1 Q_2 V_2^HH^H = \frac{P_B}{N_B} H H^H$, and (7) becomes

$$R_{AB} =$$

$$= \log \left| I + \frac{P_B}{N_B} G G^H + \frac{P_B}{N_B} H H^H \right| - \log \left| I + \frac{P_B}{N_B} G G^H \right|$$

$$= \log \left| I + J_1 \Theta_1 H \right| - \log \left| I + J_2 \Theta_2 H \right|$$

(30)

where $J_1 = \frac{1}{\sqrt{N_B}} [H, G]$, $J_2 = \frac{1}{\sqrt{N_B}} G$, $\Theta_2 = \rho P_B I$, and $\Theta_1 = \text{diag}([P_1, P_B 1_{N_B}])$. Applying the lemma to (30) yields that for $\beta_1 = \frac{N_A + N_B}{N_B}$, $\beta_2 = 1$ and a large $N_B$,

$$\frac{R_{AB}}{N_B} \approx \Omega (\beta_1, \Theta_1, \eta_1) - \Omega (\beta_2, \Theta_2, \eta_2)$$

$$= (\beta_1 - 1) \log(1 + \eta_1 P_s) + \log \left( \frac{\eta_1}{\eta_2} \right) + (\eta_1 - \eta_2) \log(e)$$

(31)

where $\eta_1$ is the solution of $\eta$ to (20) with $\beta = \beta_1$ and $\Theta = \Theta_1$, which reduces to

$$1 - \eta_1 = \frac{(\beta_1 - 1) \eta_1 P_s}{\eta_1 \rho P_B} + \frac{\eta_1 \rho P_B}{1 + \eta_1 \rho P_B}$$

(32)

and $\eta_2$ is the solution to

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

(33)

or equivalently $\eta_2 = \frac{\eta_1 + \eta_2 \rho P_B}{2 \eta_2 \rho P_B} - 1$.

Also with $r = N_B$ and $Q_r = \frac{P_B}{N_B} I$, $R_{AE}$ in (21) reduces to

$$\frac{R_{AE}}{N_E} \approx \Omega (\beta_3, \Theta_3, \eta_3) - \Omega (\beta_4, \Theta_4, \eta_4)$$

$$= (\beta_3 - \beta_4) \log \left( 1 + \frac{a P_s \eta_3}{\beta_3 - \beta_4} \right)$$

$$+ (\beta_3 - \beta_4) \log \left( \frac{\beta_3 - \beta_4 + a P_B \eta_3}{\beta_3 - \beta_4 + a P_B \eta_4} \right)$$

$$+ (2 \beta_4 - \beta_3) \log \left( \frac{2 \beta_4 - \beta_3 + a P_n \eta_3}{2 \beta_4 - \beta_3 + a P_n \eta_4} \right)$$

$$+ \log \left( \frac{\eta_3}{\eta_4} \right) + (\eta_3 - \eta_4) \log(e)$$

(34)

where

$$\Theta_3 = N_E \text{diag} \left( \frac{a P_s}{N_B} 1_{N_A}, \frac{a P_n}{N_A - N_B} 1_{N_A - N_B}, \frac{b P_B}{N_B} 1_{N_B} \right)$$

(35)
and

\[ \Theta_4 = N_E \text{diag} \left( \frac{a_P a_P}{N_A - N_B} N_A - N_B, \frac{b_P b_P}{N_B} N_B \right), \]  

(36)

also \( \beta_3 = \frac{N_A + N_B}{N_E}, \beta_4 = \frac{N_A}{N_E}, \eta_3 \) is the solution to

\[ 1 - \eta_3 = \frac{a_P \eta_3}{1 + a_P \eta_3} + \frac{b_P \eta_3}{1 + b_P \eta_3}, \]  

\[ + \frac{1 + b_P \eta_3}{1 + b_P \eta_3 \frac{1}{N_E}}, \]  

(37)

and \( \eta_4 \) is the solution to

\[ 1 - \eta_4 = \frac{a_P \eta_4}{1 + a_P \eta_4 \frac{1}{N_E}} + \frac{b_P \eta_4}{1 + b_P \eta_4 \frac{1}{N_E}}, \]  

\[ + \frac{1 + b_P \eta_4}{1 + b_P \eta_4 \frac{1}{N_E}}, \]  

(38)

Fig. 3 shows \( \bar{N}_E \) versus \( N_A \) and \( \bar{N}_E^{\text{opt}} \) versus \( N_A \) where \( P_A^{\text{max}} = P_B^{\text{max}} = 30 \text{ dB} \) and \( N_A = 2N_B \) (\( \beta_3 = 3 \)). For \( \bar{N}_E \), we also chose \( P_s = P_n = \frac{P_A}{2} \) and \( P_A = P_B \). We see that \( \bar{N}_E^{\text{opt}} \) 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 \( P_s, P_n \) is often distributed approximately equally between different streams, and that if \( \bar{N}_E \) 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 \( \epsilon \) 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 \( Q_e = \frac{P_B}{N_B} I \). 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

\[ \lim_{N_E \to \infty} N_E (\beta_3 - \beta_4) \log \left( 1 + \frac{a_P \eta_3}{\beta_3 - \beta_4} \right) \]  

\[ = N_B \log \left( 1 + \frac{N_E a_P}{N_B} \right), \]  

(39)

\[ \lim_{N_E \to \infty} N_E (\beta_3 - \beta_4) \log \frac{\beta_3 - \beta_4 + b_P \eta_3}{\beta_3 - \beta_4 + b_P \eta_4} \]  

\[ = N_B \log 1 = 0, \]  

(40)

\[ \lim_{N_E \to \infty} N_E (2\beta_4 - 3\beta_3) \log \frac{2\beta_4 - 3\beta_3 + a_P \eta_3}{2\beta_4 - 3\beta_3 + a_P \eta_4} \]  

\[ = (N_A - N_B) \log 1 = 0, \]  

(41)

\[ \lim_{N_A \to \infty} N_E \log \left( \frac{\eta_4}{\eta_3} \right) \right) \]  

\[ \left( \eta_3 - \eta_4 \right) \log(e) \]  

\[ = \lim_{N_E \to \infty} N_E \log \left( 1 + \frac{N_B}{N_E - (N_A + N_B)} \right) - N_B \log e \]  

\[ = \lim_{N_E \to \infty} N_E \frac{N_B}{N_E - (N_A + N_B)} \log e - N_B \log e \]  

(42)

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

\[ \lim_{N_E \to \infty} R_{AE} = \lim_{N_E \to \infty} R_{AE} = N_B \log \left( 1 + \frac{N_E a_P}{N_B} \right) \]  

(43)

which is independent of \( P_n \) and \( P_B \) (and hence the optimal \( P_s \) 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 \( N_E \)).
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 \( p_A(k) \) and \( p_B(k) \) 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 \( P_i = [p_i(1), \ldots, p_i(K_i)] \) where \( i = A, B \) then the corresponding signals received by Alice, Bob and Eve can be expressed as

\[
Y_A = H^T P_B + N_A \tag{44a}
\]

\[
Y_B = H P_A + N_B \tag{44b}
\]

\[
Y_E = \sqrt{a}P_A + \sqrt{b}P_B + N_E \tag{44c}
\]

where \( H \) 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 \( H \) by Bob, \( P_A \) should be such that \( P_A P_B^H = \frac{K_1 P_A}{N_A} I_{N_A} \). Similarly, \( P_B \) should be such that \( P_A P_B^H = \frac{K_1 P_A}{N_B} I_{N_B} \).

In the following analysis, we assume that \( H, A \) and \( B \) 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 \( P_B \) be part of the row span of \( P_A \). More specifically, we can write \( P_A = \sqrt{\frac{K_1 P_A}{N_A}} [I_{N_A}, 0_{N_A \times (K_1 - N_A)]} \Gamma \) and \( P_B = \sqrt{\frac{K_1 P_B}{N_B}} [I_{N_B}, 0_{N_B \times (K_1 - N_B)]} \Gamma \) where \( \Gamma \) can be any \( K_1 \times K_1 \) unitary matrix. In this way, any estimates of \( A \) and \( B \) by Eve, denoted by \( \hat{A} \) and \( \hat{B} \), are ambiguous in that \( [\sqrt{\hat{A}}, \sqrt{\hat{B}}] \) can be added to \( \Theta (C_A, C_B) \) without affecting Eve’s observation \( Y_E \) where \( \Theta \in \mathbb{C}^{N_E \times N_B} \) is arbitrary and \( [C_A, C_B][P_A^T, P_B^T] = 0 \).

Let \( h = vec(H), a = vec(A), b = vec(B) \), \( y_A = vec(Y_A^T) \), \( y_B = vec(Y_B^T) \), \( n_A = vec(N_A^T) \) and \( n_B = vec(N_B^T) \). Note vec(XYZ) = (\( Z \otimes X \))vec(Y). Then (44) becomes

\[
y_A = (I_{N_A} \otimes P_B^T) h + n_A \tag{45a}
\]

\[
y_B = (P_A^T \otimes I_{N_A}) h + n_B \tag{45b}
\]

\[
y_E = \sqrt{a} (P_A^T \otimes I_{N_E}) a + \sqrt{b} (P_B^T \otimes I_{N_E}) b + n_E \tag{45c}
\]

It is known that the minimum-mean-squared-error (MMSE) estimate of a vector \( x \) from another vector \( y \) is \( \hat{x} = K_{x,y} K_y y \) with \( K_{x,y} = E[xy^H] \) and \( K_y = E[yy^H] \). And...
the error $\Delta x = x - \hat{x}$ has the covariance matrix $K_{\Delta x} = K_x - K_x\hat{y}K_y^{-1}K_y$. Let $\hat{h}_A$ be the MMSE estimate of $h$ by Alice, and $\Delta h_A = h - \hat{h}_A$ 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, $K_{\Delta h_A} = \sigma_2^2I_{N_A} + K_{\Delta A} = \sigma_2^2I_{N_A}$, $K_{\Delta h_B} = \sigma_2^2I_{N_B}$, and $K_{\Delta A} = \sigma_2^2E_BI_{N_A}E^*B$ where $\sigma_2^2 = \frac{1}{1 + K_1P_A/N_A}$, $\sigma_2^1 = \frac{1}{\frac{1}{1 + K_1P_A/N_A} + K_{\Delta A}} + 1$ and $\sigma_2^0 = \frac{1}{\frac{1}{1 + K_1P_A/N_A} + K_{\Delta A}} + 2$.

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

$$Y_B = HS_A + N_B$$
$$Y_E = \sqrt{\alpha}AS_A + N_E$$

where $S_A = [s_A(1), \ldots, s_A(K_2)]$. The corresponding vector forms of the above are

$$y_B = (I_{K_2} \otimes H)s_A + n_B$$
$$y_E = \sqrt{\alpha}(I_{K_2} \otimes A)s_A + n_E$$

where $s_A = vec(S_A)$ (which is assumed to be independent of all channel parameters). Then an achievable secrecy rate in bits/Hz in phase 2 from Alice to Bob (conditional on the MMSE channel estimation in phase 1) is

$$R_{one} = \frac{1}{K_2}(I(s_A; y_B|\hat{h}_B) - I(s_A; y_E|\hat{a})) +$$

To analyze $R_{one}$, we now assume $P_A = P_B = P$ (which holds for both phases 1 and 2) and that $s_A(k)$ are i.i.d. with $CN(0, \frac{P_A}{N_A}I_{N_A})$. We also use $H_B = \imath \{vec(h_B)\} \in \mathbb{C}^{N_B \times N_A}$ (i.e., $H_B = \imath \{vec(H_B)\}$).

We will next derive lower and upper bounds on $R_{one}$. To do that, we need to obtain lower and upper bounds on $I(s_A; y_B|\hat{h}_B)$ and those on $I(s_A; y_E|\hat{a})$.

First, we have

$$I(s_A; y_B|\hat{h}_B) = h(s_A|\hat{h}_B) - h(s_A|y_E|\hat{a}) = h(s_A) - h(s_A|y_B, \hat{h}_B).$$

(49)

It is known that $h(s_A) = \log([\pi e]^{N_AK_2p_{A_N}I_{N_A}K_2}]$. It is also known [38] that for a random vector $x \in \mathbb{C}^{n \times 1}$ and another random vector $w$, $h(x|w) \leq \log((\pi e)^{\frac{p_{A_N}I_{N_A}K_2}})$ where $K_{sym} = K_x - K_{sym}(K_{sym})^{-1}K_{sym}$ which is the covariance matrix of the MMSE estimation of $s$ from $w$. Note that $y_B = (I_{K_2} \otimes H_B)s_A + (I_{K_2} \otimes \Delta H_B)\bar{s}_A + n_B$. Then conditional on $H_B$ (which is independent of $s_A$), the covariance matrix of the MMSE estimate of $\bar{s}_A$ from $y_B$ is $K_{sym}[y_B, h_B] = \frac{P_A}{N_A}I_{N_A}K_2 - \frac{P_A}{N_A}I_{K_2} \otimes H_B^{H} + K_2 + I_{N_B}K_2) \otimes (I_{K_2} \otimes \Delta H_B)s_A$. Using $\|r_{A} + \bar{A} B\| = |r_{B} + BA|$ where $r_A$ and $r_B$ are the numbers of rows of $A$ and $B$ respectively, one can verify that $\log \left| K_{sym[y_B, h_B]} \right| = N_AK_2 \log \left| \frac{P_A}{N_A}I_{N_A}K_2 \right| - \log \left| \frac{P_A}{N_A}I_{K_2} \otimes H_B^{H} + K_2 + I_{N_B}K_2 \right| = N_AK_2 \log \frac{P_A}{N_A}I_{N_A}K_2 - K_2 \log |N_B| + \frac{P_A}{N_A}I_{N_B}K_2 \otimes \Delta H_B^{H}$. Applying the above results to (49) yields

$$\begin{align*}
I(s_A; y_B|\hat{h}_B) &\geq \log \left| \frac{P_A}{N_A}I_{N_A}K_2 \right| - \mathcal{E} \left\{ \log \left| K_{sym[y_B, h_B]} \right| \right\} \\
&= K_2\mathcal{E} \left\{ \log \left| \frac{P_A}{1 + K_1P_A/N_A} \right| \
&\quad + \frac{P_A}{N_A}I_{N_B}K_2 \otimes \Delta H_B^{H} \right\} \\
&\geq \mathcal{R}_B.
\end{align*}$$

(50)

To derive an upper bound on $I(s_A; y_B|\hat{h}_B)$, we now write

$$I(s_A; y_B|\hat{h}_B) = h(y_B|\hat{h}_B) - h(y_B|\hat{h}_B, s_A).$$

(51)

Here, $h(y_B|\hat{h}_B) \leq \mathcal{E} \left\{ \log \| (\pi e)^{\frac{p_{A_N}I_{N_A}K_2}} \| \right\}$, and $h(y_B|\hat{h}_B, s_A) = \mathcal{E} \left\{ \log \| (\pi e)^{\frac{P_A}{N_A}I_{N_B}K_2} \| \right\}$. Note that conditional on $\hat{h}_B$ and $s_A$, the covariance matrix of $y_B$ is invariant to $\hat{h}_B$. Now define

$$M_A = \begin{cases}
\frac{N_A}{P_A}S_A^TS_A, & K_2 < N_A \\
\frac{N_A}{P_A}S_A^TS_A - 1, & K_2 \geq N_A
\end{cases}$$

(52)

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

$$\begin{align*}
\mathcal{E} \left\{ \log \left| 1 + \frac{P_A}{1 + K_1P_A/N_A} \right| \left( M_A + I_{t_A} \right) \right\} \\
&\geq t_A \mathcal{E} \left\{ \log \left| 1 + \frac{P_A}{1 + K_1P_A/N_A} \right| \left( M_A + I_{t_A} \right) \right\} \\
&= t_A \mathcal{E} \left\{ \log \left| 1 + \frac{P_A}{1 + K_1P_A/N_A} \exp \left( \frac{1}{t_A} \right) \left( M_A \right) \right| \right\} \\
&\geq t_A \log \left| 1 + \frac{P_A}{1 + K_1P_A/N_A} \exp \left( \frac{1}{t_A} \right) \left( M_A \right) \right| \\
&= t_A \log \left| 1 + \frac{P_A}{1 + K_1P_A/N_A} \exp \left( \frac{1}{t_A} \right) \left( M_A \right) \right| \\
&= t_A \log \left| 1 + \frac{P_A}{1 + K_1P_A/N_A} \exp \left( \frac{1}{t_A} \right) \right| \\
&= t_A \log \left| 1 + \frac{P_A}{1 + K_1P_A/N_A} \exp \left( \frac{1}{t_A} \right) \right|
\end{align*}$$

(53a)

Authorized licensed use limited to: Univ of Calif Riverside. Downloaded on April 16, 2020 at 16:44:24 UTC from IEEE Xplore. Restrictions apply.
where (53a) is due to the matrix Minkowski inequality $|X + Y|^{1/n} \geq |X|^{1/n} + |Y|^{1/n}$ where $X$ and $Y$ are $n \times n$ positive definite matrices [39], (53b) is due to the Jensen’s inequality and that log$(1 + a e^x)$ is a convex function of $x$ when $a > 0$, and (53c) is based on [40, Th. 1] where $\gamma \approx 0.57721566$ is Euler’s constant. Defining $e_A = \exp \left( \frac{1}{N_A} \sum_{t=1}^{T} \sum_{k=1}^{N} \frac{r_{t,k}}{1 - \frac{1}{N_A} (1 - \gamma)} \right)$ and applying the above results since (51), we have from (51) that

$$I(\hat{s}_A; Y_B|\hat{h}_B) \leq K_2 E \left( \log \left( \frac{I_{N_B} + \frac{P_A/N_A \hat{H}_B \hat{H}_B^H}{1 + K_1 P_A/N_A} \right) \right) + N_B \log \left( \frac{1 + \frac{P_A}{1 + K_1 P_A/N_A} \frac{K_2}{1 + \frac{P_A}{1 + K_1 P_A/N_A} \frac{N_A}{\sigma_A^2}} \right) \right) \triangleq R_{E}^+ (54)$$

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

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

$$I(\hat{s}_A; Y_E|\hat{a}) \geq K_2 E \left( \log \left( I_{N_E} + \frac{P_A/N_A \hat{A} \hat{A}^H}{1 + P_A \sigma_{E A}^2} \right) \right) \triangleq R_{E}^- (55)$$

and

$$I(\hat{s}_A; Y_E|\hat{a}) \leq R_{E}^- + N_E \log \left( \frac{(1 + P_A \sigma_{E A}^2) K_2}{(1 + P_A \sigma_{E A}^2) N_A} \right) \triangleq R_{E}^- (56)$$

When $P_A = P_B = P \to \infty$, we have $\sigma_{E A}^2 \to 0$, $\frac{b}{a N_A}$, $\frac{a N_A}{a N_B + b N_A}$, $\frac{a N_B}{a N_B + b N_A}$ and $E\{a_\hat{t} a_\hat{t}^H\} \to 1$. From [41, Th. 2], we know that $E\{\log [1 + \frac{1}{N} XX^H]\} \to \min(r, t) \log P + o(\log P)$ as $P \to \infty$ where the entries of $X \in C^{r \times t}$ are i.i.d. $CN(0, 1)$. Therefore, from (50) and (54),

$$\lim_{P \to \infty} \frac{R_{E}^-}{\log P} = \lim_{P \to \infty} \frac{R_{E}^+}{\log P} = K_2 \min\{N_A, N_B\} \triangleq R_{E}^- (57)$$

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

$$\lim_{P \to \infty} \frac{R_{E}^-}{\log P} = 0 \triangleq R_{E}^- (58)$$

and

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

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

$$\lim_{P \to \infty} \frac{R_{E}^-}{\log P} = \begin{cases} \min\{N_A, N_B\}, & K_2 \leq N_A \\ \left( \min\{N_A, N_B\} - \frac{N_E}{K_2} (K_2 - N_A) \right)^+, & K_2 > N_A \end{cases} \triangleq R_{E}^- (60)$$

and

$$\lim_{P \to \infty} \frac{R_{E}^-}{\log P} = \min\{N_A, N_B\}. \triangleq R_{E}^- (61)$$

Note that $\lim_{P \to \infty} \frac{R_{E}^-}{\log P}$ 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 $\lim_{P \to \infty} \frac{R_{E}^-}{\log P} = \min\{N_A, N_B\}$ 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 $H, A$ and $B$. 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

$$Y_A = H^T S_B + N_A$$
$$Y_B = H S_A + N_B$$
$$Y_E = \sqrt{a} S_A + \sqrt{b} S_B + N_E$$

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

$$y_A = (I_{K_2} \otimes H^T) s_B + n_A$$
$$y_B = (I_{K_2} \otimes H) s_A + n_B$$
$$y_E = \sqrt{a} (I_{K_2} \otimes A) s_A + \sqrt{b} (I_{K_2} \otimes B) s_B + n_E$$

where both $\hat{s}_A$ and $\hat{s}_B$ 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]):

$$R_{two} = \frac{1}{K_2} \left( I(\hat{s}_B; y_A | \hat{h}_A) + I(\hat{s}_A; y_B | \hat{h}_B) \right)$$
$$- I(\hat{s}_A, \hat{s}_B; y_E | \hat{a}, \hat{b})$$

(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 $I(\hat{s}_A; Y_B|\hat{h}_B)$. To show a similar pair of lower and upper bounds on $I(\hat{s}_B; Y_A|\hat{h}_A)$, we let $\hat{H}_A = ivec(\hat{h}_A)$. One
where $e_{E1} = \exp\left(\frac{1}{2}\sum_{k=1}^{K_2} \sum_{i=1}^{N_A+N_B-j} \frac{1}{k} - \gamma\right)$. The second inequality in (68) is from the fact (e.g., see [43, Th. 3]) that $|S_{AB}TS_{AB}^H| \geq \sigma_{2\min}^2 |S_{AB}S_{AB}^H|$ where $\sigma_{2\min}^2 = \min(\sigma_{EA}^2, \sigma_{EB}^2)$. Similarly, for $K_2 \geq N_A+N_B$, we have

$$
\mathcal{E}\left\{\log |I + S_{AB}TS_{AB}^H|\right\} \\
= \mathcal{E}\left\{\log |I + TS_{AB}S_{AB}^H|\right\} \\
\geq (N_A+N_B) \mathcal{E}\left\{\log \left(1 + |T| \frac{\pi_{A+N\min}}{e} e_{E2}\right)\right\} \\
\geq (N_A+N_B) \log \left(1 + |T| \frac{\pi_{A+N\min}}{e} e_{E2}\right)
$$
(69)

where $e_{E2} = \exp\left(\frac{1}{2}\sum_{j=1}^{N_A+N_B} \sum_{k=1}^{K_2-j} \frac{1}{k} - \gamma\right)$. Therefore, using (68) and (69), we have from (67) that

$$
I(\hat{s}_A,\hat{s}_B;\hat{y}_E|\hat{a},\hat{b}) \\
\leq K_2 \mathcal{E}\left\{\log \left(\frac{P_A^2 \AA^H + P_B^2 \BB^H}{1 + P_A^2 \sigma_{2\min}^2} + |I|\right)\right\} \\
\leq K_2 N_E \log \left(\frac{1 + P_A^2 \sigma_{2\min}^2}{1 + P_A^2 \sigma_{2\min}^2} + |I|\right) \leq K_2 \leq N_A+N_B \\

\leq \mathcal{R}_{E,t}^+ \tag{70}
$$

One can also verify $I(\hat{s}_A,\hat{s}_B;\hat{y}_E|\hat{a},\hat{b}) \geq K_2 \mathcal{E}\left\{\log \left(\frac{P_A^2 \AA^H + P_B^2 \BB^H}{1 + P_A^2 \sigma_{2\min}^2} + |I|\right)\right\} \geq \mathcal{R}_{E,t}^-$ which is the first term in (70).

When $P_A = P_B = P \to \infty$, we have $\sigma_{2\min}^2 = \frac{b_{N_A}}{a_{N_A}+b_{N_A}}, \sigma_{2\min}^2 = \frac{a_{N_B}}{a_{N_B}+b_{N_B}}, \sigma_A^* = 0, \sigma_B^* = 0, E[\tilde{a}_i \tilde{a}_i^*] \to \frac{a_{N_A}}{a_{N_A}+b_{N_A}}, E[\tilde{b}_i \tilde{b}_i^*] \to \frac{a_{N_B}}{a_{N_B}+b_{N_B}}, E[\tilde{h}_A \tilde{h}_A^*] \to 1, E[\tilde{h}_B \tilde{h}_B^*] \to 1, \frac{\pi_{A+N\min}}{e} e_{E2} = P \left((\frac{\pi_{A+N\min}}{e})^{N_A(N_B)}\right) \to 1/(N_A+N_B)$.

Then, similar to (57), we have

$$
\lim_{P \to \infty} \mathcal{R}_{E,t}^- = \lim_{P \to \infty} \mathcal{R}_{E,t}^+ = K_2 \min\{N_A,N_B\} \tag{71}
$$

One can also verify that
as upper and lower bounds on $R_{\text{two}}$, we have

$$\lim_{p \to \infty} \frac{R_{\text{two}}}{\log P} = \begin{cases} 2 \min \{N_A, N_B\}, & K_2 \leq N_A + N_B \\ 2 \min \{N_A, N_B\} - \frac{N_E}{K_2} (K_2 - N_A - N_B), & K_2 > N_A + N_B \end{cases},$$

(74)

and

$$\lim_{p \to \infty} \frac{R_{\text{two}}^+}{\log P} = 2 \min \{N_A, N_B\}$$

(75)

We see that if $K_2 \leq N_A + N_B$, $\lim_{p \to \infty} \frac{R_{\text{two}}}{\log P} = 2 \min \{N_A, N_B\}$ 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 $\lim_{p \to \infty} \frac{R_{\text{two}}}{\log P}$ 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 $\sqrt{a}A \in \mathbb{C}^{N_E \times K_2}$ 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 $K_2 \leq N_A$, 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

$$Y_E = \sqrt{a}AS_A + N_E$$

(66)

where the elements in $S_A \in \mathbb{C}^{N_A \times K_2}$ 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:

$$\hat{S}, \hat{A} = \arg \min_{S \in \mathbb{S}_N^{N_A \times K_2}, \sqrt{a} \in \mathbb{C}^{N_E \times K_2}} \|Y_E - \sqrt{a}AS\|_F^2.$$

(77)

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

$$\hat{S} = \arg \min_{S \in \mathbb{S}_N^{N_A \times K_2}} \|Y_E - Y_E S (SS^H)^{-1} S\|_F^2$$

(78)

or equivalently $\hat{S} = \arg \max_{S \in \mathbb{S}_N^{N_A \times K_2}} f(s)$, where $f(s) = \text{Tr}(S^H (SS^H)^{-1} S Z)$, $s = \text{vec}(S)$ and $Z = Y_E^H Y_E$. The above problem is computationally expensive. But we assume that Eve is able to afford it.

Assume that the solution $S$ of the above problem is so close to the actual information matrix $S_0$ that $f(s)$ can be replaced by its 2nd-order Taylor’s series expansion (which is conservative for Alice and Bob or equivalently optimistic for Eve). Then $\hat{s} = \text{vec}(\hat{S})$ has the following properties

$$\nabla_s \hat{f}(s) |_{s = \hat{s}} = 0,$$

(79)

$$\nabla_{ss} \hat{f}(s) |_{s = \hat{s}} = 0,$$

(80)

where $\hat{f}(s)$ is the second-order Taylor series expansion [44] of $f(s)$ around $s_0 = \text{vec}(S_0)$, i.e.,

$$\hat{f}(s) = f(s_0) + \nabla_s^T f(s) |_{s = s_0} (s - s_0) + \frac{1}{2} \left( \nabla_{ss} f(s) |_{s = s_0} (s - s_0)^T (s - s_0) \right).$$

(81)

which involves the Hessian matrices: $H_{ss} = \frac{\partial}{\partial s} \nabla_s f(s) |_{s = s_0}$, $H_{ss'} = \frac{\partial}{\partial s} \nabla_s f(s) |_{s = s_0}$, $H_{ss'} = \frac{\partial}{\partial s} \nabla_s f(s) |_{s = s_0}$. Subject to uniqueness of solution, solving (79) and (80) results in the following [44]

$$\hat{s} - s_0 = (H_{ss} - H_{ss'} H_{ss'}^{-1} (H_{ss'} H_{ss'}^{-1} \nabla_s f) - \nabla_s f)^{-1}.$$

(82)

Furthermore,

$$\nabla_s f = (Z (I - S^H (SS^H)^{-1} S) S^H)^{-1} S,$$

(83)

$$\nabla_{ss} f = (\nabla_s f)^T,$$

(84)

$$H_{ss} = \left[ Z - \frac{\partial (SS^H)^{-1} S - S^H (SS^H)^{-1} S}{\partial S} \right] \times \left[ \frac{\partial (SS^H)^{-1} S}{\partial S} \right]^T$$

(85)

$$\times \left[ \frac{\partial (SS^H)^{-1} S}{\partial S} \right],$$

$$H_{ss'} = \left[ \left( \frac{\partial (SS^H)^{-1} S}{\partial S} \right) \left( \frac{\partial (SS^H)^{-1} S - I}{\partial S} \right) \right]^T$$

(86)

$$\times \left( \frac{\partial (SS^H)^{-1} S - I}{\partial S} \right),$$

where $\Pi$ is a permutation matrix with

$$\Pi_{i,j} = \begin{cases} 1 & j = ((i - 1) \mod N_A) K_2 + \lfloor (i - 1) / N_A \rfloor \\ 0 & \text{else} \end{cases}$$

(87)

where $a \mod b$ 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, $H_{ss}$ 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 $H_{ss}$ and $H_{ss'}$, and removing $N_A^2$ corresponding elements from each of $\nabla_s f$ and $\nabla_{ss} f$. This results in
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.

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 $1 \leq K_2 \leq N_A$ in one-way information transmission or $1 \leq K_2 \leq N_A + N_B$ 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 $\gamma$ is defined as

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

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

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

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

$$ F^n_A(x) = \frac{1}{n} \sum_{i=1}^{n} 1\{\lambda_i(A) \leq x\} $$

(93)

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

It is obvious that

$$ \frac{1}{n} \log |I + \gamma A| = \frac{1}{n} \sum_{i=1}^{n} \log(1 + \gamma \lambda_i(A)) $$

$$ = \int_0^\infty \log(1 + \gamma x) dF_A(x), $$

(94)

and if $n \to \infty$ then

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

(95)

which is the Shannon transform of the eigenvalue distribution of the matrix $A$ when $n$ is large.

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

$$ \eta_\Theta(x) = \frac{1}{L_\Theta} \sum_{j=1}^{L_\Theta} \frac{1}{1 + x \Theta_{jj}}. $$

(96)

From Theorem 2.39 in [36], the $\eta-$transform of the eigenvalue distribution of $\Theta$ with parameter $\gamma$, denoted by $\eta_\Theta$ here, satisfies

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

(97)

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

$$ 1 - \eta = \beta \left( 1 - \frac{1}{L_\Theta} \sum_{j=1}^{L_\Theta} \frac{1}{1 + \eta \Theta_{jj}} \right) $$

(98)

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

$$ \mathcal{V}_{\Theta}(\gamma) = \beta \mathcal{V}_\Theta(\gamma) - \log(\eta) + (\eta - 1) \log(e) $$

(99)

which is $\Omega(\beta, \Theta, \eta)$ 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 $Y_E$ as in (44c) the following:

$$ \hat{A} = A \Theta C_A $$

(100)

$$ \hat{B} = B \Theta C_B $$

(101)

where $[C_A, C_B] \in \mathbb{C}^{\min\{N_A, N_B\} \times (N_A + N_B)}$ is a known matrix satisfying $[C_A, C_B][P_A, P_B]^T = 0$. 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:

$$ \min_{S \in S_{N_A \times K_2}} \|Y_E - AS\|^2 $$

(102)

or equivalently

$$ \min_{S \in S_{N_A \times K_2}} \|Y_E - (\hat{A} - \Theta C_A) S\|^2. $$

(103)

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

$$ \Theta = - (Y_E - \hat{A} S)(C_A S(C_A S)^H)^{-1}. $$

(104)

Then, the problem of (103) reduces to

$$ \min_{S \in S_{N_A \times K_2}} \|Y_E - (\hat{A} S)(I_{K_2} - P_{C_A S})\|^2 $$

(105)

where $P_{C_A S} = (C_A S)^H(C_A S(C_A S)^H)^{-1}C_A S$. The problem of (105) is more complex than (78) due to higher order of the cost function in terms of $S$. 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.
SOHRABI et al.: SECRECY ANALYSES OF A FULL-DUPLEX MIMO NETWORK

REFERENCES


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.