Title
Advanced Properties of Full-Duplex Radio for Securing Wireless Network

Journal
IEEE TRANSACTIONS ON SIGNAL PROCESSING, 67(1)

Author
Hua, Yingbo

Publication Date
2019

Data Availability
The data associated with this publication are within the manuscript.

Peer reviewed

Advanced Properties of Full-Duplex Radio for Securing Wireless Network

Yingbo Hua, *Fellow, IEEE*

***Abstract*—This paper examines the role of full-duplex radio for securing wireless network from a new perspective. It first studies the secrecy capacity of two single-antenna full-duplex users against a multi-antenna eavesdropper (Eve) who has the perfect knowledge of the channel state information (CSI) from users to Eve. It is shown that if Eve uses a basic matched-filtering, the probability of zero secrecy (or outage) can be made small by a large jamming power from both users and a small gain of residual self-interference (RSI) power. But if Eve uses the optimal matched-filtering, the probability of outage grows rapidly as either the jamming power from the users increases or the number of antennas on Eve increases, regardless of the RSI gain. To prevent any Eve from obtaining its CSI, this paper then proposes a novel anti-eavesdropping channel estimation (ANECE) method, which allows users to obtain their own CSI while keeping all Eves in handicap. This method also prevents Eves from colluding with each other at any layer. The design of ideal pilots for ANECE for multiple multi-antenna users and multiple broadband multi-antenna users is discussed. It is also shown that the capacity of Eve with any number of antennas but without its CSI can be virtually eliminated over a time window corresponding to the number of antennas at the transmitter for each realization of the CSI.**

***Index Terms*—Wireless network security, full-duplex radio, anti-eavesdropping channel estimation, mobile ad hoc network, drone network, multi-agent network.**

I. Introduction

FULL-DUPLEX radio which can receive and transmit at the same time and same frequency has received much attention in recent years. There are many works focusing on how to reduce the amount of self-interference on a full-duplex radio and how to implement prototypes of full-duplex radio, e.g., see [1]–[6]. There are also many works on how to utilize full-duplex radio for improved network capacity, e.g., see [7]–[11].

This paper is concerned with the application of full-duplex radio for securing wireless network. This line of prior works include [12]–[23]. Much of these prior works focuses on optimization of secrecy capacity over power allocation in subcarriers or in antenna beamspace. A common assumption in all these works is that the legitimate users have the full knowledge of the large-scale fading of their channels with respect to eavesdroppers (Eves). Furthermore, all these works assume that Eves have the full knowledge of both the large-scale and small-scale fading of their channels with respect to the legitimate users.

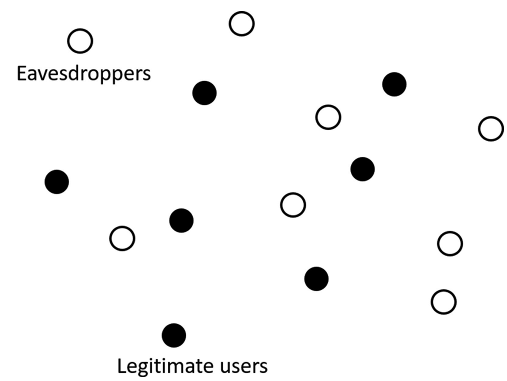

Fig. 1. An illustration of wireless network of legitimate users (such as drones) all with full-duplex capability but subject to passive eavesdropping from unknown locations.

This paper studies the application of full-duplex radio for securing wireless network from a perspective that is unique from all of the above mentioned prior works. It focuses on a wireless network of legitimate users all equipped with full-duplex radio, subject to eavesdropping from arbitrary locations as illustrated in Fig. 1. The paper first shows an analysis of the secrecy capacity of two single-antenna users against an arbitrarily located multi-antenna Eve (which could represent a network of colluding Eves) that has the full knowledge of its channel state information (CSI) with respect to the users. (In this paper, "user", "legitimate user" and "full-duplex radio" are interchangeable.) Two cases of how Eve processes their received signals are considered: one is a basic matched-filtering (BMF), and the other is the optimal matched-filtering (OMF). For BMF, it is shown that the probability of zero secrecy (outage) of the two users against the multi-antenna Eve can be always made small by a large jamming power from the two users and a small power gain of the residual self-interference (RSI) channel. But for OMF, it is shown that the probability of outage of the two users grows rapidly as either the number of the antennas on the Eve increases or the jamming power from the two users increases.

With the above important insight, a novel method for anti-eavesdropping channel estimation (ANECE) is also proposed in this paper. This method allows both users to estimate their own CSI with respect to each other but at the same time prevents any Eves from obtaining their CSI with respect to the users. It is shown that any Eve without its CSI can be made virtually blind

Manuscript received March 13, 2018; revised July 20, 2018 and September 17, 2018; accepted October 28, 2018. Date of publication November 5, 2018; date of current version November 20, 2018. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Qingjiang Shi. This work was supported in part by the Army Research Office under Grant W911NF-17-1-0581. 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.

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

Digital Object Identifier 10.1109/TSP.2018.2879621

and deaf to the secret information transmitted between the users. This ANECE method is also extended in this paper to the cases of two multi-antenna users, multiple multi-antenna users as well as broadband multiple multi-antenna users. The ANECE method does not require the users to have any knowledge of the CSI at the Eves except for a minor constraint shown later. The proposed method for ANECE differs substantially from those in [41] and [42]. In [41], a recursive training scheme was proposed where each recursion involves a feedback from Bob to Alice. In [42], assuming the reciprocal property of the channel between Alice and Bob, it was suggested that Alice simply avoids transmitting any pilot but relies on Bob transmitting a pilot. None of these prior methods prevents Eve from obtaining its CSI with respect to Bob. Furthermore, the proposed method for ANECE in case of multiple users allow all users to obtain their mutual CSI while preventing Eves from obtaining their CSI with respect to any of the users.

The basic assumptions made in this paper are the following A1–A4:

A1: The circuitry for self-interference cancellation in each full-duplex radio is already built-in, and there still exists an amount of RSI, the power of which is proportional to the power of the original source of the self-interference. But the RSI waveform is virtually white and its shape is independent of the original self-interference waveform. The RSI is mainly caused by noises (such as thermal noises) in circuitry and hence tends to be white especially if the optimal cancellation scheme is used. This model is based on the prior experiences in self-interference cancellation for full-duplex radio [4]–[6] where both theory and hardware implementations are available. The above modelling of the RSI channel can also be explained as follows. Let $s_1(t) + n_1(t)$ be the signal plus noise at the input of the original self-interference channel (including all components applied for interference isolation) with the signal-to-noise ratio $\mathtt{SNR}_1 = \frac{S_1}{N_1}$, and $s_2(t) + n_2(t)$ be the signal plus noise at the output of the original channel with $\mathtt{SNR}_2 = \frac{S_2}{N_2}$. Note that $s_1(t)$ and $s_2(t)$ may be perfectly correlated but $n_1(t)$ and $n_2(t)$ are independent and white. The noise factor of the original channel is defined as $K = \frac{\mathtt{SNR}_1}{\mathtt{SNR}_2} > 1$. Assuming an optimal self-interference cancellation, the RSI would be a combination of $n_1(t)$ and $n_2(t)$, and is typically dominated by $n_2(t)$. The power of $n_2(t)$ is $N_2 = K\frac{S_2}{\mathtt{SNR}_1}$ where S_2 is typically proportional to S_1. For example, if $\mathtt{SNR}_1 = 60$ dB, $K = 5$ dB and $\frac{S_1}{S_2} = 60$ dB, then $\frac{N_2}{S_1} = -115$ dB.

A2: All Eves are passive, i.e., they do not transmit. For statistical analysis of secrecy capacity, the small-scale fading from users to Eves is assumed to have an amplitude-wise Rayleigh distribution from the users' perspective. Some results of the analysis do not either require the users to know the large-scale fading from users to Eves. For ANECE, no knowledge of the CSI from users to Eves is required by the users.

A3: All noises are white Gaussian. All users are equally capable. This assumption is made for simple exposure of key ideas although some extension from this assumption can be straightforward.

A4: All antennas are omnidirectional within angles of interest unless mentioned otherwise.

The results shown in this paper may or may not require non-collusion among Eves, which will be made clear in the context. The potential applications of these results include in particular mobile ad hoc wireless network (such as drone network) where all nodes are capable of full-duplex at a common frequency. It is also important to mention that this paper focuses on transmission of secret information between users who do not share pre-existing secret information. If there is any pre-existing secret information shared by the users, they can communicate with each other securely (without full-duplex) by following the conventional methods such as cryptography at the network layer [24]. Any existing secret shared among users could be also used by the users at the physical layer (e.g., as a seed of a pseudo-random sequence) to generate such an artificial noise that interferes Eves but is removable by legitimate receivers. The use of artificial noise in beamspace is not considered in this paper.

In Section II, the secrecy capacity of two single-antenna users against an arbitrarily located multi-antenna Eve is analysed. Some of the results are a generalization of some shown in [25] where an arbitrarily located single-antenna Eve was focused on. In Section III, the proposed ANECE method is presented for a two-user case and a multi-user case. The extensions of ANECE for two multi-antenna users, multiple multi-antenna users and multiple broadband multi-antenna users are given in the appendix.

The key novelty of this paper includes: novel insights into the performance of single-antenna full-duplex radio users against multi-antenna Eves with knowledge of their CSI, novel method for ANECE in various settings; and novel understanding of the secrecy capacity between users against Eves without knowledge of their CSI.

The notations: All vector variables are in bold lower case. All matrix variables are in bold upper case. The expectation is $\mathcal{E}\{\cdot\}$. The magnitude of a scalar or the determinant of a matrix is $|\cdot|$. The norm of a vector is $\|\cdot\|$. The complex Gaussian distribution of zero mean and covariance $\mathbf{R}$ is $\mathcal{CN}(\mathbf{0}, \mathbf{R})$. The $M \times M$ identity matrix is $\mathbf{I}_M$. The set of all $n \times m$ complex matrices is $\mathcal{C}^{n\times m}$. All other notations are defined in the context.

II. Properties of Secrecy Capacity Against Eves With CSI

In this section, the focus is on the exchange of secret information between two users (Alice and Bob) each with a single antenna and a full-duplex capability. ("secret information", "secret key" and "key" are interchangeable in this paper.) The users are subject to eavesdropping by a multi-antenna Eve at arbitrary location. Such a multi-antenna Eve could represent a network of distributed colluding Eves.

Without loss of generality, let the normalized locations of Alice and Bob be $(-0.5, 0)$ and (0.5, 0) respectively, and that of an arbitrary Eve be (x_E, y_E). (The 3-D case can be similarly treated.) Note that all location coordinates, channel gains and noise variances shown in this paper are normalized in a similar way as in [25]. The channel gain between Alice and Bob is denoted by h. The averaged power gain of the RSI at Alice or Bob is ρ, but the instantaneous RSI power gains at Alice and Bob are respectively $\rho|g_A|^2$ and $\rho|g_B|^2$ where g_A and g_B are caused by small-scale fading. (If the channel gain between Alice and Bob is direction dependent, it can be treated similarly.) Assume that h, $|g_A|$ and $|g_B|$ are known to Alice and Bob so that statistical properties conditional on these parameters are meaningful.

The channel vector from Alice to Eve (with M antennas) is denoted by $\sqrt{a}\mathbf{h}_A \in \mathcal{C}^{M\times1}$, and that from Bob to Eve is $\sqrt{b}\mathbf{h}_B \in \mathcal{C}^{M\times1}$. Here, a and b represent the large-scale fading of Eve, and could be approximated by $a = \frac{1}{d_A^\alpha}$ and $b = \frac{1}{d_B^\alpha}$ with $\alpha \geq 2$ and d_A and d_B being the distances from Alice and Bob to Eve. The small-scale fading of Eve with respect to Alice and Bob is represented by $\mathbf{h}_A$ and $\mathbf{h}_B$, respectively.

Consider an exchange of keys between Alice and Bob in two phases. In phase 1, Alice sends a key to Bob, and in phase 2, Bob sends a (different) key to Alice. The two phases can be either in two time slots or in two frequency bands. But in either case, assume that all channel amplitudes are invariant from phase 1 to phase 2.

In phase 1, Bob receives

$$y_B(k) = \sqrt{P_T} h x_A(k) + \sqrt{\rho P_J} g_B w_B(k) + n_B(k) \tag{1}$$

and an Eve receives

$$\mathbf{y}_{A,E}(k) = \sqrt{aP_T}\mathbf{h}_A x_A(k) + \sqrt{bP_J}\mathbf{h}_B v_B(k) + \mathbf{n}_{A,E}(k) \tag{2}$$

where $\mathbf{y}_{A,E}(k) \in \mathcal{C}^{M\times 1}$, k is the time index, $\sqrt{P_T}x_A(k)$ of power P_T is the signal transmitted by Alice, $\sqrt{P_J}v_B(k)$ of power P_J is the jamming noise from Bob, $\sqrt{\rho P_J} g_B w_B(k)$ of power ρP_J is the RSI noise at Bob after self-interference cancellation, n_B is the background noise at Bob, and $\mathbf{n}_{A,E}$ is the background noise vector at Eve. All noises are assumed to be white Gaussian.

Then, the SNR at Bob is

$$SNR_{A,B} = \frac{|h|^2 P_T}{1+\rho|g_B|^2 P_J}. \tag{3}$$

The effective SNR at Eve depends on how Eve processes $\mathbf{y}_{A,E}(k)$. If Eve knows $\mathbf{h}_A$, then it can perform a BMF to $\mathbf{y}_{A,E}(k)$ as follows:

$$\begin{aligned} r_{A,E,b}(k) \doteq \mathbf{h}_A^H \mathbf{y}_{A,E}(k) &= \sqrt{aP_T}\|\mathbf{h}_A\|^2 x_A(k) \\ &+ \sqrt{bP_J}\mathbf{h}_A^H\mathbf{h}_B w_B(k) + \mathbf{h}_A^H \mathbf{n}_{A,E}(k). \end{aligned} \tag{4}$$

In this case, the effective SNR at Eve is

$$SNR_{A,E,b} = \frac{a\|\mathbf{h}_A\|^4 P_T}{\|\mathbf{h}_A\|^2 + b|\mathbf{h}_A^H\mathbf{h}_B|^2 P_J} = \frac{a\|\mathbf{h}_A\|^2 P_T}{1 + b\frac{|\mathbf{h}_A^H\mathbf{h}_B|^2}{\|\mathbf{h}_A\|^2} P_J}. \tag{5}$$

On the other hand, if Eve knows $\mathbf{h}_A$, $\mathbf{h}_B$, b and P_J, then Eve could use the following OMF on $\mathbf{y}_{A,E}(k)$ (which is noise whitening followed by matched-filtering):

$$r_{A,E,o}(k) \doteq \mathbf{h}_A^H \mathbf{R}_{A,E}^{-1}\mathbf{y}_{A,E}(k) \tag{6}$$

where $\mathbf{R}_{A,E}$ is the covariance matrix of the noise terms in $\mathbf{y}_{A,E}(k)$ and

$$\mathbf{R}_{A,E}^{-1} = (\mathbf{I} + bP_J\mathbf{h}_B\mathbf{h}_B^H)^{-1} = \mathbf{I} - \frac{bP_J\mathbf{h}_B\mathbf{h}_B^H}{1+bP_J\|\mathbf{h}_B\|^2}. \tag{7}$$

Then, the effective SNR at Eve is

$$\begin{aligned} SNR_{A,E,o} &= aP_T\mathbf{h}_A^H\mathbf{R}_{A,E}^{-1}\mathbf{h}_A \\ &= aP_T\|\mathbf{h}_A\|^2 - \frac{abP_TP_J|\mathbf{h}_A^H\mathbf{h}_B|^2}{1+bP_J\|\mathbf{h}_B\|^2}. \end{aligned} \tag{8}$$

It is known that $SNR_{A,E,b} \leq SNR_{A,E,o}$ with equality if $\mathbf{R}_{A,E}$ is proportional to the identity matrix, and the OMF leads to the optimal detection at Eve.

In phase 2, Alice receives

$$y_A(k) = \sqrt{P_T} h x_B(k) + \sqrt{\rho P_J} g_A w_A(k) + n_A(k) \tag{9}$$

and Eve receives

$$\mathbf{y}_{B,E}(k) = \sqrt{bP_T}\mathbf{h}_B x_B(k) + \sqrt{aP_J}\mathbf{h}_A v_A(k) + \mathbf{n}_{B,E}(k) \tag{10}$$

where the notations are similarly defined as before. Note that $\mathbf{n}_{B,E}(k)$ and $\mathbf{n}_{A,E}(k)$ are independent. For convenience, a symmetry between Alice and Bob in terms of P_T, P_J and the distributions of noises has been assumed. It follows that the SNR at Alice is

$$SNR_{B,A} = \frac{|h|^2 P_T}{1+\rho|g_A|^2 P_J}. \tag{11}$$

The effective SNR at Eve based on the BMF is

$$SNR_{B,E,b} = \frac{b\|\mathbf{h}_B\|^2 P_T}{1 + a\frac{|\mathbf{h}_B^H\mathbf{h}_A|^2}{\|\mathbf{h}_B\|^2} P_J} \tag{12}$$

where Eve requires the knowledge of $\mathbf{h}_B$. And the effective SNR at Eve based on the OMF is

$$SNR_{B,E,o} = bP_T\|\mathbf{h}_B\|^2 - \frac{abP_TP_J|\mathbf{h}_A^H\mathbf{h}_B|^2}{1+aP_J\|\mathbf{h}_A\|^2} \tag{13}$$

where Eve requires the knowledge of $\mathbf{h}_A$, $\mathbf{h}_B$, a and P_J.

Assuming no collusion between Eves, the averaged secrecy capacity in bits/s/Hz against any given Eve is

$$\mathcal{S} = \frac{1}{2}(\mathcal{S}_{A,B} + \mathcal{S}_{B,A}) \tag{14}$$

$$\mathcal{S}_{A,B} = [\log_2(1+SNR_{A,B}) - \log_2(1+SNR_{A,E})]^+ \tag{15}$$

$$\mathcal{S}_{B,A} = [\log_2(1+SNR_{B,A}) - \log_2(1+SNR_{B,E})]^+ \tag{16}$$

where $SNR_{A,E}$ and $SNR_{B,E}$ should be replaced by either $SNR_{A,E,b}$ and $SNR_{B,E,b}$ or $SNR_{A,E,o}$ and $SNR_{B,E,o}$, depending on whether the BMF or the OMF is used by Eve. Here, $(x)^+ = \max(x, 0)$.

Important properties of the secrecy capacity shown in (14)–(16) will be discussed next. Some of these properties based on the BMF resemble the case of single-antenna Eves shown in [25]. It is useful to also focus on $\mathcal{S}_{A,B}$, from which one can infer the properties of $\mathcal{S}_{B,A}$ in an obvious way.

A. *Properties of $\mathcal{S}_{A,B}$ Based on BMF*

$\mathcal{S}_{A,B}$ is the secrecy capacity of the transmission from Alice to Bob against any given non-colluding Eve. The position and/or CSI of this Eve affects $\mathcal{S}_{A,B}$. It is also important to realize that the minimum of $\mathcal{S}_{A,B}$ over all possible locations and/or channel gains of Eves subject to any known constraint on Eves defines the overall secrecy capacity of the transmission from Alice to Bob against all non-colluding Eves. That minimum is also the overall secrecy capacity against all Eves that could only collude at the network layer.

For some applications such as wireless communications between drones high in air and away from buildings, there can be virtually no multi-path between them and hence virtually no small scale fading. In this case, the elements of $\mathbf{h}_A$ and $\mathbf{h}_B$ are all constants. So, the properties of $\mathcal{S}_{A,B}$ even in the absence of small-scale fading are useful.

One can verify that a generalization of Property 1 in [25] is the following:[1]

[1]The proofs of this property and some other properties that follow are similar to those of the corresponding properties shown in [25]. Due to space limitation, these proofs are omitted.

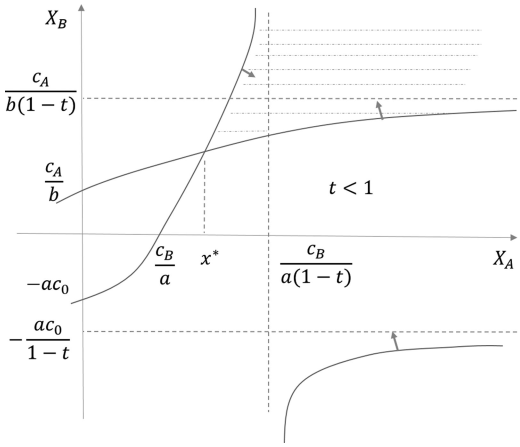

Fig. 2. Region of solutions of positive X_A and X_B to (45) for $t<1$.

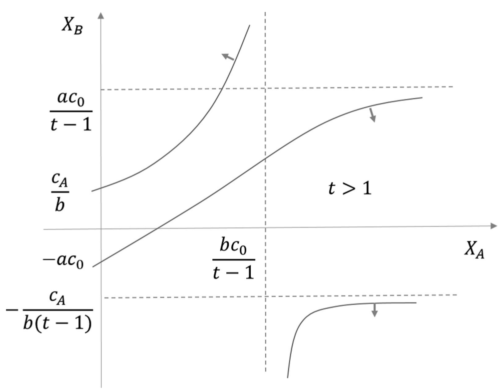

Fig. 3. Illustration of no solution of positive X_A and X_B to (45) for $t>1$ and $t^*\leq 1$.

Property 1: $\mathcal{S}_{A,B}>0$ iff

1) $\hat{b}-\hat{\rho}\hat{a}>0$ and $P_J>\hat{\gamma}$; or
2) $\hat{b}-\hat{\rho}\hat{a}<0$ and $P_J<\hat{\gamma}$; or
3) $\hat{b}-\hat{\rho}\hat{a}=0$ and $\hat{a}<1$.

where $\hat{a}=a\frac{\|\mathbf{h}_A\|^2}{|h|^2}$, $\hat{b}=b\frac{|\mathbf{h}_A^H\mathbf{h}_B|^2}{\|\mathbf{h}_A\|^2}$, $\hat{\rho}=\rho|g_B|^2$, and $\hat{\gamma}=\frac{\hat{a}-1}{\hat{b}-\hat{\rho}\hat{a}}$. ■

This property suggests that for a given pair of $\hat{a}$ and $\hat{b}$, there are typically a turning point for $\hat{\rho}$ and a turning point for P_J in order to achieve a positive secrecy. In the absence of small-scale fading, both $\hat{a}$ and $\hat{b}$ are constants.

A restatement of the above property is the following, which is also a generalization of Property 2 in [25],

Property 2:

1) For Eve in region $\mathcal{R}_1\doteq\{\hat{b}-\hat{\rho}\hat{a}>0$ and $\hat{a}<1\}$, $\mathcal{S}_{A,B}>0$ for any $P_J\geq 0$.
2) For Eve in region $\mathcal{R}_2\doteq\{\hat{b}-\hat{\rho}\hat{a}>0$ and $\hat{a}\geq 1\}$, $\mathcal{S}_{A,B}>0$ iff $P_J>\hat{\gamma}$.
3) For Eve in region $\mathcal{R}_3\doteq\{\hat{b}-\hat{\rho}\hat{a}\leq 0$ and $\hat{a}<1\}$, $\mathcal{S}_{A,B}>0$ if $P_J=0$.
4) For Eve in region $\mathcal{R}_4\doteq\{\hat{b}-\hat{\rho}\hat{a}\leq 0$ and $\hat{a}\geq 1\}$, $\mathcal{S}_{A,B}=0$ for any $P_J\geq 0$.

■

As illustrated by Fig. 2 and 3 in [25] for $M=1$, each of the above regions corresponds to a geometric region in space. Without the small-scale fading, all the regions are separated from each other by circular boundaries. It is seen here that $\mathcal{R}_2$ is the most important region of Eve that Alice and Bob need to pay special attention to especially if Eves could collude at the network layer. This is because if there is a colluding Eve nearby Alice, then typically $\hat{a}>1$, and hence one must keep $\hat{\rho}$ small such that $\hat{b}-\hat{\rho}\hat{a}>0$ in order to keep $\mathcal{S}_{A,B}>0$ under some $P_J>\hat{\gamma}$.

Let $P_{J,opt}=\arg\max_{P_J}\mathcal{S}_{A,B}$. A generalization of Property 6 in [25] is the following:

Property 3:

1) For Eve in region $\mathcal{R}_1$,

$$P_{J,opt}=\left[\hat{\gamma}+\sqrt{\hat{\gamma}^2+\hat{\beta}}\right]^+\geq 0 \tag{17}$$

with $\hat{\beta}=\frac{\hat{a}\hat{b}-\hat{\rho}+\hat{a}\hat{P}_T(\hat{b}-\hat{\rho})}{\hat{\rho}\hat{b}(\hat{b}-\hat{\rho}\hat{a})}$ and $\hat{P}_T=P_T|h|^2$.
2) For Eve in region $\mathcal{R}_2$, $P_{J,opt}$ is given by (17) but is strictly positive.
3) For Eve in region $\mathcal{R}_3$, $P_{J,opt}=0$. ■

A useful insight from this property is that for any given $\hat{a}$ and $\hat{b}$, as $\hat{\rho}$ decreases and P_T increases, $P_{J,opt}$ becomes $\sqrt{\frac{\hat{a}(1+\hat{P}_T)}{\hat{\rho}\hat{b}}}=\mathcal{O}(\sqrt{\frac{\hat{P}_T}{\hat{\rho}}})$. It is also useful to know that for $\mathcal{R}_2$, $\arg\max_{\hat{a},\hat{b}}P_{J,opt}=\arg\min_{\hat{a},\hat{b}}\mathcal{S}_{A,B}$.

Now consider the impact of small-scale fading on the secrecy capacity $\mathcal{S}_{A,B}$. Both $\hat{a}$ and $\hat{b}$ depend not only on a and b but also on the small-scale fading of the channels of Eve. For multi-path rich environment, it is a good assumption that $\mathbf{h}_A\in\mathcal{C}^{M\times 1}$ and $\mathbf{h}_B\in\mathcal{C}^{M\times 1}$ are independent and each has the distribution $\mathcal{CN}(\mathbf{0},\mathbf{I})$. Then, it follows that that $2X_A\doteq 2\|\mathbf{h}_A\|^2$ and $2X_B\doteq 2\|\mathbf{h}_B\|^2$ are independent and each has the probability density function (PDF) of the standard Chi-squared random variable of $2M$ degrees of freedom, i.e., $f_{X_A}(u)=f_{X_B}(u)=\frac{1}{\Gamma(M)}u^{M-1}e^{-u}$.

It is obvious that X_A and X_B are independent of $\mathbf{u}_A\doteq\frac{1}{\|\mathbf{h}_A\|}\mathbf{h}_A$ and $\mathbf{u}_B\doteq\frac{1}{\|\mathbf{h}_B\|}\mathbf{h}_B$ that represent the directions of $\mathbf{h}_A$ and $\mathbf{h}_B$. Then, it follows that $\mathbf{u}_A^H\mathbf{h}_B$ and $\mathbf{u}_B^H\mathbf{h}_A$ are independent of X_A and X_B, respectively. Furthermore, conditional on $\mathbf{u}_A$, $\mathbf{u}_A^H\mathbf{h}_B$ is complex Gaussian with zero mean and unit variance, which is invariant to $\mathbf{u}_A$. Hence, either conditional or unconditional upon $\mathbf{u}_A$, $\mathbf{u}_A^H\mathbf{h}_B$ remains to be $\mathcal{CN}(0,1)$. Similarly, $\mathbf{u}_B^H\mathbf{h}_A$ is $\mathcal{CN}(0,1)$. Therefore, $Y_B\doteq\frac{|\mathbf{h}_A^H\mathbf{h}_B|^2}{\|\mathbf{h}_A\|^2}$ and $Y_A\doteq\frac{|\mathbf{h}_B^H\mathbf{h}_A|^2}{\|\mathbf{h}_B\|^2}$ are independent of X_A and X_B, respectively, and have the PDF $f_{Y_A}(x)=f_{Y_B}(x)=e^{-x}$ for $x\geq 0$.

Referring to (3) and (5), it follows that $\mathcal{S}_{A,B}=0$ iff $SNR_{A,B}\leq SNR_{A,E,b}$ or equivalently,

$$X_A-v_1Y_B-v_2>0 \tag{18}$$

where $v_1=\frac{b|h|^2P_J}{a(1+\rho|g_B|^2P_J)}$, $v_2=\frac{|h|^2}{a(1+\rho|g_B|^2P_J)}$. Both v_1 and v_2 are invariant to the small-scale fading of Eves. But they are functions of a, b and the CSI of Alice and Bob. It follows that the probability of $\mathcal{S}_{A,B}=0$, conditional on v_1 and v_2, is

$$\mathcal{P}(\mathcal{S}_{A,B}=0|v_1,v_2)=\int_0^\infty e^{-y}dy\int_{v_1y+v_2}^\infty\frac{1}{\Gamma(M)}x^{M-1}e^{-x}dx. \tag{19}$$

Since v_1 is proportional to $\frac{b}{a}$ and v_2 is proportional to $\frac{1}{a}$, one can verify that $\arg\max_{x_E,y_E}\mathcal{P}(\mathcal{S}_{A,B}=0|v_1,v_2)$, subject to $a=\frac{1}{d_A^\alpha}$, $b=\frac{1}{d_B^\alpha}$ and $d_A\geq\Delta$, equals to $(x_E,y_E)=(-0.5-\Delta,0)$, i.e., the secrecy is the worst against the Eve that is the closest to Alice on her opposite side from Bob. Also, as a increases, both v_1 and v_2 approach zero and hence $\mathcal{P}(\mathcal{S}_{A,B}=0|v_1,v_2)$ approaches one.

An alternative form of $\mathcal{P}(\mathcal{S}_{A,B}=0|v_1,v_2)$ can be shown by switching the order of the two integrals in (19) to be

$$\mathcal{P}(\mathcal{S}_{A,B}=0|v_1,v_2)=\frac{1}{\Gamma(M)}\left[\Gamma(M,v_2)-e^{\frac{v_2}{v_1}}\right. \\ \left.\times\left(1+\frac{1}{v_1}\right)^{-M}\Gamma\left(M,v_2\left(1+\frac{1}{v_1}\right)\right)\right] \tag{20}$$

where $\Gamma(p,z)=\int_z^\infty x^{p-1}e^{-x}dx$. It is known [28] that $\lim_{p\to\infty}\frac{\Gamma(p,z)}{\Gamma(p)}=1$ for any given z. Therefore, one has:

Property 4:

$$\lim_{M\to\infty}\mathcal{P}(\mathcal{S}_{A,B}=0|v_1,v_2)=1 \tag{21}$$

for any given v_1 and v_2. ■

To see an asymptotical dependence of $\mathcal{P}(\mathcal{S}_{A,B}=0|v_1,v_2)$ on M, one can verify the following:

Property 5: For a large P_J such that $\rho|g_B|^2P_J\gg 1$, $v_1\approx\frac{b|h|^2}{a\rho|g_B|^2}$ and $v_2\approx\frac{|h|^2}{a\rho|g_B|^2P_J}\to 0$, it follows from (20) that

$$\mathcal{P}(\mathcal{S}_{A,B}=0|v_1,v_2)\to 1-\frac{1}{(1+\frac{1}{v_1})^M}. \tag{22}$$

■

The above two properties show how detrimental it is to Alice and Bob if Eve has multiple antennas.

One could also discuss the properties of $\mathcal{S}_{A,B}$ based on the OMF. But it is obvious that the use of OMF at Eve must make $\mathcal{S}_{A,B}$ even worse than those shown above. So, those properties are omitted. The next section discusses the properties of $\mathcal{S}$ that results from dual transmissions.

B. Properties of $\mathcal{S}$ Based on BMF

It is obvious that $\mathcal{S}=0$ iff $\mathcal{S}_{A,B}\le 0$ and $\mathcal{S}_{B,A}\le 0$, or equivalently iff $SNR_{A,B}\le SNR_{A,E}$ and $SNR_{B,A}\le SNR_{B,E}$. Related to Part 3) of Property 10 in [25] is:

Property 6: There is a $P_J\ge 0$ such that $\mathcal{S}>0$ if

$$\rho<|h|^2|\mathbf{h}_A^H\mathbf{h}_B|^2\max\left\{\frac{b}{a}\frac{1}{|g_B|^2\|\mathbf{h}_A\|^4},\frac{a}{b}\frac{1}{|g_A|^2\|\mathbf{h}_B\|^4}\right\}. \tag{23}$$

■

Proof: The above follows from the application of Parts 1 and 2 of Property 2 to the two components in $\mathcal{S}$. ■

This property covers the important and typical situation where Eve is not far away from Alice and Bob. And it shows how in this situation the required ρ is affected by all the parameters of the CSI in a deterministic fashion. For example, one can see that the required ρ is small if the two vectors $\mathbf{h}_A$ and $\mathbf{h}_B$ are nearly orthogonal, and/or the norms of the two vectors are large (all of which tend to be true when M is large).

Next, the properties of $\mathcal{S}$ subject to random $\mathbf{h}_A$ and $\mathbf{h}_B$ are discussed. One can verify that $\mathcal{S}=0$ iff

$$\begin{cases}|h|^2+b|h|^2\Phi X_BP_J\le a(1+\rho|g_B|^2P_J)X_A\\ |h|^2+a|h|^2\Phi X_AP_J\le b(1+\rho|g_A|^2P_J)X_B\end{cases} \tag{24}$$

where $\Phi=\frac{|\mathbf{h}_A^H\mathbf{h}_B|^2}{\|\mathbf{h}_A\|^2\|\mathbf{h}_B\|^2}$, $X_A=\|\mathbf{h}_A\|^2$ and $X_B=\|\mathbf{h}_B\|^2$. One can rewrite (24) as $\Phi\le T_b$ where

$$T_b=\min\left\{\frac{c_1X_A-c_2}{X_B},\frac{c_3X_B-c_4}{X_A}\right\} \tag{25}$$

with $c_1=\frac{a}{b}\frac{1+\rho|g_B|^2P_J}{|h|^2P_J}$, $c_2=\frac{1}{bP_J}$, $c_3=\frac{b}{a}\frac{1+\rho|g_A|^2P_J}{|h|^2P_J}$ and $c_4=\frac{1}{aP_J}$. Therefore,

$$p_0\doteq\mathcal{P}\{\mathcal{S}=0\}=\mathcal{P}\{\Phi\le T_b\}. \tag{26}$$

It is useful to examine p_0 subject to fixed ρ, P_J, $|h|^2$, $|g_A|^2$ and $|g_B|^2$ but random X_A, X_B and Φ. Note that the first set of parameters are known to Alice and Bob while the second set of parameters are random unknowns.

Under the assumption that $\mathbf{h}_A$ and $\mathbf{h}_B$ are independent $\mathcal{CN}(\mathbf{0},\mathbf{I})$, it is known that X_A, X_B and Φ are independent. It is also known that $2X_A$ and $2X_B$ are the standard Chi-squared random variables each with degree $2M$, which can be used to derive the PDF of T_b. But the distribution of Φ does not seem readily available in the literature except for the case of two real-valued vectors [26]. To see useful properties of $\mathcal{S}$, one needs to understand the PDFs of T_b and Φ as discussed next.

Property 7: Let

$$t^*=\sqrt{c_1c_3}=\frac{\sqrt{(1+\rho|g_B|^2P_J)(1+\rho|g_A|^2P_J)}}{|h|^2P_J}. \tag{27}$$

For $0<t<t^*$, the PDF of T_b is

$$f_{T_b}(t)=\int_{\frac{c_2c_3+c_4t}{c_1c_3-t^2}}^\infty f_{X_A}(x)\left[\frac{x}{c_3}f_{X_B}\left(\frac{c_4+tx}{c_3}\right)\right. \\ \left.+\frac{c_1x-c_2}{t^2}f_{X_B}\left(\frac{c_1x-c_2}{t}\right)\right]dx \tag{28}$$

where $f_{X_A}(x)=f_{X_B}(x)=\frac{1}{\Gamma(M)}x^{M-1}e^{-x}u(x)$ and $u(x)$ is the unit step function. (The step function will not be used when it is obvious that the variable must be positive.) For $t>t^*$, $f_{T_b}(t)=0$. For $t<0$, $f_{T_b}(t)$ is generally nonzero, but becomes negligible if P_J becomes large. The explicit expression of $f_{T_b}(t)$ for $t<0$ is omitted due to its unimportance. ■

Proof: The CDF (cumulative distribution function) of T_b is

$$F_{T_b}(t)=\mathcal{P}\{T_b\le t\}=1-\mathcal{P}\{T_b>t\} \\ =1-\mathcal{P}\{c_1X_A-tX_B>c_2,c_3X_B-tX_A>c_4\}. \tag{29}$$

One can verify that the two linear inequalities in the last expression can possibly hold simultaneously iff $t<t^*$. In other words, if $t\ge t^*$, $F_{T_b}(t)=1$. Subject to $0<t<t^*$, one can verify that

$$F_{T_b}(t)=1-\int_{\frac{c_2c_3+c_4t}{c_1c_3-t^2}}^\infty dx\int_{\frac{c_4+tx}{c_3}}^{\frac{c_1x-c_2}{t}}f_{X_A}(x)f_{X_B}(y)dy \tag{30}$$

where the upper and lower limits of the inner integral over y are equal when x in the outer integral equals its lower limit. Then (28) follows from $f_{T_b}(t)=\frac{\partial F_{T_b}(t)}{\partial t}$ (where one of the three terms in the derivative is zero). It is obvious from (25) that T_b can be negative. But as P_J increases, c_2 and c_4 become negligible, so does the probability of $T_b<0$, and hence so does $f_{T_b}(t)$ for $t<0$. ■

For $M>1$, it seems hard to reduce (28) even if possible. But for $M=1$, it is known that $f_{X_A}(x)=f_{X_B}(x)=e^{-x}u(x)$, and one can verify after some tedious manipulations that (28) reduces to

$$f_{T_b}(t) = \left(\frac{c_3}{(c_3+t)^2} + \frac{c_2c_3 + c_4t}{(c_3+t)(c_1c_3 - t^2)} + \frac{c_1c_4 + c_2t}{(c_1+t)(c_1c_3-t^2)} + \frac{c_1}{(c_1+t)^2} \right) \times \exp\left\{ -\frac{c_1c_4 + c_2c_3 + (c_2+c_4)t}{c_1c_3 - t^2} \right\}. \tag{31}$$

For a large P_J, the self-interferences at both Alice and Bob dominate the background noise, i.e., $\rho|g_B|^2P_J \gg 1$ and $\rho|g_A|^2P_J \gg 1$, and hence t^* converges to its lower bound $\rho\frac{|g_A||g_B|}{|h|^2}$ which is proportional to ρ and invariant to P_J.

It is useful to know the mean of T_b under large P_J. Assume a large P_J for which c_2 and c_4 go to zero. In this case, (28) becomes

$$\begin{aligned} f_{T_b}(t) &\approx \int_0^\infty f_{X_A}(x)\left[\frac{x}{c_3} f_{X_B}\left(\frac{tx}{c_3}\right) + \frac{c_1x}{t^2} f_{X_B}\left(\frac{c_1x}{t}\right)\right]dx \\ &= \frac{\Gamma(2M)}{\Gamma^2(M)}\left[\frac{1}{c_3}\frac{(\frac{t}{c_3})^{M-1}}{(1+\frac{t}{c_3})^{2M}} + \frac{1}{t}\frac{(\frac{c_1}{t})^M}{(1+\frac{c_1}{t})^{2M}}\right] \end{aligned} \tag{32}$$

where $c_1 = \frac{a\rho|g_B|^2}{b|h|^2}$, $c_3 = \frac{b\rho|g_A|^2}{a|h|^2}$ and $0 < t < t^* = \sqrt{c_1c_3} = \frac{\rho|g_A||g_B|}{|h|^2}$. Furthermore, one can show that the expectation of T_b is

$$\begin{aligned} \mathcal{E}\{T_b\} &= \frac{\Gamma(2M)}{\Gamma^2(M)}\left[\int_0^{t^*}\frac{(\frac{t}{c_3})^M}{(1+\frac{t}{c_3})^{2M}}dt + \int_0^{t^*}\frac{(\frac{c_1}{t})^M}{(1+\frac{c_1}{t})^{2M}}dt\right] \\ &= \frac{\Gamma(2M)}{\Gamma^2(M)}\left[c_3\int_0^{\sqrt{\frac{c_1}{c_3}}}\frac{x^M}{(1+x)^{2M}}dx + c_1\int_0^{\sqrt{\frac{c_3}{c_1}}}\frac{x^M}{(1+x)^{2M}}dx\right]. \end{aligned} \tag{33}$$

where the change of variable $\frac{c_1}{t} = \frac{1}{x}$ has been applied to obtain the second term for the second equation. Note that $\sqrt{\frac{c_1}{c_3}} = \frac{a|g_B|}{b|g_A|}$. Using the change of variable $x = \frac{y}{1-y}$ in the previous integral, one has

$$\begin{aligned} \mathcal{E}\{T_b\} &= \frac{\Gamma(2M)}{\Gamma^2(M)}\left[c_3\int_0^{r_A} y^M(1-y)^{M-2}dy + c_1\int_0^{r_B} y^M(1-y)^{M-2}dy\right] \\ &= \frac{\Gamma(2M)}{\Gamma^2(M)}[c_3B(r_A;M+1,M-1) + c_1B(r_B;M+1,M-1)] \end{aligned} \tag{34}$$

where $r_A = \frac{\sqrt{\frac{c_1}{c_3}}}{1+\sqrt{\frac{c_1}{c_3}}} = \frac{\frac{a|g_B|}{b|g_A|}}{1+\frac{a|g_B|}{b|g_A|}}$, $r_B = 1 - r_A$, and $B(x;a,b) = \int_0^x t^{a-1}(1-t)^{b-1}dt$ which is known as incomplete beta function. Let $I(r;p,q) = \frac{B(r;p,q)}{B(p,q)}$ which is known as regularized incomplete beta function and upper bounded by one. Then, (34) becomes

$$\mathcal{E}\{T_b\} = \frac{M}{M-1}\frac{\rho}{|h|^2}\left[\frac{b}{a}|g_A|^2 I(r_A;M+1,M-1) + \frac{a}{b}|g_B|^2 I(r_B;M+1,M-1)\right]. \tag{35}$$

It is known from the discussion of the equation (3.15) in [27] that $\lim_{M\to\infty} I(r;M+1,M-1) = 0$ for a fixed $r < \frac{1}{2}$ and $\lim_{M\to\infty} I(r;M+1,M-1) = 1$ for a fixed $r > \frac{1}{2}$. One can then conclude the following:

Property 8:

$$\lim_{M\to\infty, P_J\to\infty} \mathcal{E}\{T_b\} = \frac{\rho|g_A||g_B|}{|h|^2}\min\left\{\frac{a'}{b'}, \frac{b'}{a'}\right\} \tag{36}$$

where $a' = \frac{a}{|g_A|}$ and $b' = \frac{b}{|g_B|}$. ■

Lemma 1: The PDF of Φ (the squared-cosine-magnitude of the angle between two complex Gaussian vectors) as defined in (24) is given by

$$f_\Phi(x) = \frac{\Gamma(M)}{\Gamma(M-1)}(1-x)^{M-2} = (M-1)(1-x)^{M-2} \tag{37}$$

where $0 \le x \le 1$ and $M \geq 2$. Note that the mean of Φ is $\frac{1}{M}$. ■

Proof: See Appendix A. ■

A generalization of Lemma 1 is:

Lemma 2: Let $\Phi = \frac{|\mathbf{a}^H\mathbf{b}|^2}{\|\mathbf{a}\|^2\|\mathbf{b}\|^2}$ where $\mathbf{a}$ is $\mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$ and $\mathbf{b}$ has any distribution subject to $\mathbf{b} \neq \mathbf{0}$ with probability one. Then, Φ has the beta distribution $\mathtt{Beta}(1, M-1)$ as shown in (37).

Proof: An outline of the proof is as follows. Conditional on $\mathbf{b}$, one can write $\Phi = \frac{X}{X+Y}$ where X and Y are independent, $X = |\mathbf{a}^H\frac{\mathbf{b}}{\|\mathbf{b}\|}|^2$ and $Y = \sum_{i=2}^M |\mathbf{a}^H\mathbf{e}_i|^2$ where $\mathbf{e}_i$ for $i = 2, \ldots, M$ are orthonormal and span the orthogonal complement of $\mathbf{b}$. Furthermore, for any $\mathbf{b} \neq \mathbf{0}$, X has the Chi-squared distribution with degree two, and Y has the Chi-squared distribution with degree $2(M-1)$. It follows that Φ has the beta distribution $\mathtt{Beta}(1, M-1)$. ■

With the knowledge of the PDFs of T_b and Φ, one can evaluate the probability p_0 of zero secrecy by

$$\begin{aligned} p_0 &= \mathcal{P}\{\Phi \le T_b\} \\ &= \int_0^{t^*} f_{T_b}(t)\int_0^{\min\{t,1\}} f_\Phi(\phi)d\phi dt \\ &= \int_0^1 f_\Phi(\phi)\int_\phi^{t^*} f_{T_b}(t)dtd\phi. \end{aligned} \tag{38}$$

This expression can be used to compute p_0 under any ρ, P_J, $|h|^2$, $|g_A|^2$ and $|g_B|^2$. A special case of the above is as follows:

Property 9: If P_J is large and ρ is small so that $t^* \ll 1$, and $M \geq 2$, then

$$\begin{aligned} p_0 &\approx \int_0^{t^*} f_{T_b}(t)(M-1)tdt = (M-1)\mathcal{E}\{T_b\} \\ &< (M-1)t^* = (M-1)\rho\frac{|g_A||g_B|}{|h|^2}, \end{aligned} \tag{39}$$

and for a large M,

$$p_0 \approx M\rho \frac{|g_A||g_B|}{|h|^2} \min\left\{\frac{a'}{b'}, \frac{b'}{a'}\right\}. \tag{40}$$

■

This property says that with a large P_J, one can always keep p_0 small by keeping $(M-1)\rho$ small. It is also important to note that the upper bound on p_0 shown in (39) is invariant to any CSI of Eves. The result of (40) implies that for large M, large P_J and small ρ, the most harmful Eve is at a location where $a' = b'$, i.e., $\frac{a}{|g_A|} = \frac{b}{|g_B|}$. When $|g_A| = |g_B|$, such a location is at the half-way between Alice and Bob. (Although p_0 here is constant for all Eves on the y-axis, the secrecy capacity S increases as Eve moves away along the y-axis from the origin.)

For the special case of $M = 1$, it is known that $\Phi = 1$, i.e., $f_\Phi(x) = \delta(x-1)$.

Property 10: For $M = 1$,

$$p_0 = \mathcal{P}\{1 \le T_b\} = \int_1^{t^*} f_{T_b}(t)dt \tag{41}$$

where $f_{T_b}(t)$ is given by (31). Here, p_0 is zero iff $t^* \le 1$. ■

Note that for $M = 1$, one can keep $p_0 = 0$ exactly by keeping $t^* \le 1$. But for $M > 1$, it is unavoidable to have $p_0 > 0$ although one can make p_0 small by having a small $\rho(M-1)$.

The above properties have shown something rather encouraging in terms of the secrecy capacity of two single-antenna users against multi-antenna Eves. In other words, provided that P_J is large enough and ρ is small enough, one can keep p_0 small. This motivates us to look into $\mathcal{S}$ based on the OMF at Eve, as shown next.

C. Properties of $\mathcal{S}$ Based on OMF

Now consider $\mathcal{S}$ defined in (14)–(16) with $SNR_{A,E}$ and $SNR_{B,E}$ replaced by $SNR_{A,E,o}$ and $SNR_{B,E,o}$ respectively. One can verify that $\mathcal{S} = 0$ iff $\Phi \le T_o$ where

$$T_o = \min\{Z_{A,B}, Z_{B,A}\} \tag{42}$$

where $Z_{A,B} = (\frac{c_0}{X_B} + \frac{1}{a})(a - \frac{c_B}{X_A})$, $Z_{B,A} = (\frac{c_0}{X_A} + \frac{1}{b})(b - \frac{c_A}{X_B})$, $c_0 = \frac{1}{abP_J}$, $c_B = \frac{|h|^2}{1+\rho|g_B|^2P_J}$ and $c_A = \frac{|h|^2}{1+\rho|g_A|^2P_J}$. Obviously, there is a nonzero probability for $T_o < 0$.

Property 11: Let $F_{T_o}(t)$ be the CDF of T_o, and $f_{T_o}(t)$ be the PDF of T_o. Then,

$$\mathcal{P}\{\mathcal{S} = 0\} = \int_0^1 f_\Phi(\phi)d\phi \int_\phi^\infty f_{T_o}(t)dt$$
$$= 1 - \int_0^1 f_\Phi(\phi)F_{T_o}(\phi)d\phi \tag{43}$$

where $f_\Phi(x)$ is given by (37), and $F_{T_o}(t)$ for $t < 1$ is given by (49). Furthermore, $\mathcal{P}\{T_o > 1\} > 0$ iff $t^* > 1$ where t^* is defined in (27). ■

Proof: In the following, both the proof and a discussion are provided. The CDF of T_o is

$$F_{T_o}(t) = \mathcal{P}\{T_o \le t\} = 1 - \mathcal{P}\{T_o > t\} \tag{44}$$

where $T_o > t$ is equivalent to

$$\begin{cases} ac_0X_A - \frac{c_B}{a}X_B + (1-t)X_AX_B > c_0c_B \\ bc_0X_B - \frac{c_A}{b}X_A + (1-t)X_AX_B > c_0c_A \end{cases} \tag{45}$$

Assume $t < 1$. Then, the first inequality in (45) implies

$$X_B < \frac{ac_0X_A - c_0c_B}{\frac{c_B}{a} - (1-t)X_A} \tag{46}$$

if $0 \le X_A < \frac{c_B}{a(1-t)}$, or

$$X_B > \frac{ac_0X_A - c_0c_B}{\frac{c_B}{a} - (1-t)X_A} \tag{47}$$

if $X_A > \frac{c_B}{a(1-t)}$. The second inequality in (45) implies that

$$X_B > \frac{c_0c_A + \frac{c_A}{b}X_A}{bc_0 + (1-t)X_A} \tag{48}$$

for all $X_A \ge 0$. The above conditions are illustrated in Fig. 2. Let (x^*, y^*) be two positive numbers such that (45) hold with inequalities replaced by equalities and $(X_A, X_B) = (x^*, y^*)$. Namely, x^* is the positive solution to $k_2x^2 + k_1x + k_0 = 0$ where $k_2 = -\frac{c_A}{b}(1-t) - ac_0(1-t)$, $k_1 = -abc_0^2 + \frac{c_Ac_B}{ab} + c_0c_B(1-t) - c_0c_A(1-t)$ and $k_0 = bc_0^2c_B + \frac{c_0c_Ac_B}{a}$.

As illustrated in Fig. 2, one has $\frac{c_B}{a} < x^* < \frac{c_B}{a(1-t)}$. Therefore, for $t < 1$,

$$F_{T_o}(t) = 1 - \int_{x^*}^{\frac{c_B}{a(1-t)}} dx \int_{\frac{c_0c_A + \frac{c_A}{b}x}{bc_0+(1-t)x}}^{\frac{ac_0x - c_0c_B}{\frac{c_B}{a}-(1-t)x}} f_{X_A}(x)f_{X_B}(y)dy$$
$$- \int_{\frac{c_B}{a(1-t)}}^{\infty} dx \int_{\frac{c_0c_A + \frac{c_A}{b}x}{bc_0+(1-t)x}}^{\infty} f_{X_A}(x)f_{X_B}(y)dy. \tag{49}$$

Note that x^* is also function of t. The PDF $f_{T_o}(t)$ of T_o can be found from $\frac{\partial}{\partial t}F_{T_o}(t)$. But the result is too tedious to be insightful and hence omitted. Given $F_{T_o}(t)$, one can compute $\mathcal{P}\{\mathcal{S} = 0\}$ according to (43).

Now consider $t = 1$. In this case, the two nonlinear terms in (45) vanish, which results in two linear inequalities. Furthermore, one can verify that both conditions in (45) can be satisfied at the same time with a nonzero probability iff $t^* > 1$ (equivalently $a^2b^2c_0^2 - c_Ac_B > 0$). If $t^* > 1$, this nonzero probability is

$$\mathcal{P}\{T_o > 1\}$$
$$= \int_{\frac{bc_0c_B(abc_0+c_A)}{a^2b^2c_0^2 - c_Ac_B}}^{\infty} dx \int_{\frac{c_A(bc_0+x)}{b^2c_0}}^{\frac{ac_0(ax-c_B)}{c_B}} f_{X_A}(x)f_{X_B}(y)dy. \tag{50}$$

When $t > 1$, the first condition in (45) implies (46) for all $X_A \ge 0$, and the second condition in (45) implies (48) if $0 \le X_A < \frac{bc_0}{t-1}$, or the following:

$$X_B < \frac{c_0c_A + \frac{c_A}{b}X_A}{bc_0 + (1-t)X_A} \tag{51}$$

if $X_A > \frac{bc_0}{t-1}$. Illustrated in Fig. 3 are the two conditions in (45) with $t > 1$ and $t^* \le 1$, for which $\mathcal{P}\{T_o > t\} = 0$. ■

For the special case of $M = 1$, it is known that the properties of $\mathcal{S}$ must be independent of the choice of T_b or T_o since the BMF and the OMF are now equivalent. For $M = 1$, it follows that $\Phi = 1$ and hence $p_0 = \mathcal{P}\{1 \le T_b\} = \mathcal{P}\{1 \le T_o\}$, which holds despite the fact that T_b and T_o are still different random variables. Indeed, for $M = 1$, it follows that

$$p_0 = \int_1^{t^*} f_{T_b}(t)dt = \mathcal{P}\{T_o > 1\} = \begin{cases} > 0, & t^* > 1 \\ = 0, & t^* \le 1 \end{cases}. \tag{52}$$

For any integer $M \geq 1$, T_b is no larger than t^*, but T_o is not bounded by t^*. One can make t^* small by choosing large P_J and small ρ, which in turn makes T_b small statistically. But for T_o, one can show:

Property 12: With probability one,

$$\lim_{P_J \to \infty} T_o = 1 + \mathcal{O}\left(\frac{1}{P_J}\right) \tag{53}$$

and

$$\lim_{M \to \infty} T_o = 1 + \mathcal{O}\left(\frac{1}{M}\right). \tag{54}$$

■

Proof: Here, (53) follows from (42) and the fact that c_0, c_A and c_B all decrease towards zero in the order of $\mathcal{O}(\frac{1}{P_J})$ as P_J becomes larger. This holds regardless of any fixed value of ρ. (54) also follows from (42) where both X_A and X_B converges to their means equal to M statistically as M increases. ■

Since the mean of Φ is $\frac{1}{M}$, the probability of zero secrecy $\mathcal{P}\{\Phi \leq T_o\}$ based on the OMF at a multi-antenna Eve is generally high for any P_J. One can verify that if $P_J = 0$, $\mathcal{P}\{\Phi \leq T_o\} = (1 - F_{X_A}(\frac{|h|^2}{a}))(1 - F_{X_B}(\frac{|h|^2}{b}))$.

Note that for the OMF, Eve needs to know the complete CSI: $\mathbf{h}_A$, $\mathbf{h}_B$, a, b and P_J. One can further investigate the sensitivity of the secrecy capacity to the errors in the knowledge of the CSI. It is obvious from $\mathbf{R}_{A,E}^{-1}$ and $\mathbf{R}_{B,E}^{-1}$ (e.g., see (7)) that the above results are not sensitive to the errors in the knowledge of a, b and P_J provided that P_J is large or M is large. But the knowledge of $\mathbf{h}_A$ and $\mathbf{h}_B$ is essential for eavesdropping. This motivates the next section on ANECE.

III. ANECE

The insights shown in the previous section motivate the need of ANECE, for which a novel method is presented next.

A. Two-User Case

Let both Alice and Bob transmit their channel-estimation pilots simultaneously. (Such pilots cannot be secrets in general especially if they are standardized and hence accessible by third parties.) And they receive the following signals, respectively,

$$y_A(k) = h\sqrt{P_T} p_B(k) + \sqrt{\rho P_T} g_A w_A(k) + n_A(k) \tag{55}$$

$$y_B(k) = h\sqrt{P_T} p_A(k) + \sqrt{\rho P_T} g_B w_B(k) + n_B(k) \tag{56}$$

where $k = 1, \ldots, K$ is the time index, $\sqrt{P_T} p_A(k)$ is the training pilots sent from Alice, $\sqrt{P_T} p_B(k)$ is that from Bob, $\sqrt{\rho P_T} g_A w_A(k)$ is the RSI noise at Alice, $\sqrt{\rho P_T} g_B w_B(k)$ is that at Bob, and $n_A(k)$ and $n_B(k)$ are the background noises.

As discussed before, it is reasonable to assume that $w_A(k)$, $w_B(k)$, $n_A(k)$ and $n_B(k)$ are mutually independent, white, and have zero means and unit variances. Note that no assumption on the shapes of $p_A(k)$ and $p_B(k)$ has been made yet.

Alice is interested in the two unknowns: h and $|g_A|$, both of which are generally affected by the surrounding multipath. Define the vectors $\mathbf{y}_A$ and $\mathbf{p}_B$ stacked from $y_A(k)$ and $p_B(k)$, respectively, for all $k = 1, \ldots, K$. Assuming Gaussian distributions of all noises, the joint maximum likelihood estimation (MLE) estimates of h and $|g_A|$ given $\mathbf{y}_A$ and $\mathbf{p}_B$ at Alice can be shown (as in the Appendix) to be $\hat{h} = \frac{\mathbf{p}_B^H \mathbf{y}_A}{\sqrt{P_T}\|\mathbf{p}_B\|^2}$ and $|\hat{g}_A|^2 = \frac{1}{\rho P_T}(\frac{\|\mathbf{y}_A - \hat{h}\sqrt{P_T}\mathbf{p}_B\|^2}{K} - 1)$. Also, $\hat{h}$ is unbiased and has the variance:

$$var\{\hat{h}\} = \frac{\rho P_T |g_A|^2 + 1}{P_T \|\mathbf{p}_B\|^2}. \tag{57}$$

The MLE of h and $|g_A|$ are consistent as long as $\|\mathbf{p}_B\|^2$ increases with K. The simplest choice of $p_B(k)$ is such that $|p_B(k)| = 1$ for all $k = 1, \ldots, K$, for which $\|\mathbf{p}_B\|^2 = K$. In practice, ρ and K should be such that $var\{\hat{h}\}$ is small enough. Obviously, for a given required $var\{\hat{h}\}$, the smaller is ρ, the smaller is the value of K that one can allow.

At the same time as Alice performs the above estimation, Bob can perform a similar consistent estimation of h and $|g_B|$ based on $y_B(k)$ and $p_A(k) = 1$ for all $k = 1, \ldots, K$.

Now let us consider the signal received by Eve, which is

$$\mathbf{y}_E(k) = \mathbf{h}_A\sqrt{aP_T} p_A(k) + \mathbf{h}_B\sqrt{bP_T} p_B(k) + \mathbf{n}_E(k) \tag{58}$$

where $k = 1, \ldots, K$. If one uses $p_A(k) = p_B(k)$ for all k, then $\mathbf{y}_E(k) = (\sqrt{a}\mathbf{h}_A + \sqrt{b}\mathbf{h}_B)\sqrt{P_T} p_A(k) + \mathbf{n}_E(k)$. In this case, the two vectors $\mathbf{h}_A$ and $\mathbf{h}_B$ are not identifiable from $\mathbf{y}_E(k)$ with $k = 1, \ldots, K$ unless there is some prior knowledge of the two vectors. Only if $a \gg b$ (Eve is very close to Alice in terms of channel gain), then can Eve have a reasonable estimate of $\mathbf{h}_A$ but not $\mathbf{h}_B$. And only if $b \gg a$ (Eve is very close to Bob), then can Eve have a reasonable estimate of $\mathbf{h}_B$ but not $\mathbf{h}_A$. Otherwise, Eve cannot estimate any of the two vectors.

Note that if the reciprocal property of the channel between Alice and Bob holds, once Alice and Bob have found h, they already have some level of shared secret information unknown to Eve. Due to noise, any estimate of h needs to be quantized at Alice and Bob (although separately) according to a pre-determined scheme in order for them to obtain the same quantized estimate of h with high probability. The amount of secret bits shared by Alice and Bob after the estimation of h is upper bounded by the mutual information between $y_A(k)$ and $y_B(k)$. Such an analysis is available in [39].

In mobile applications, channel estimation pilot is in general an integral part of a data packet. This is because CSI varies significantly from one packet to the next due to mobility. The change of CSI is also often caused by imperfections of circuits (including carrier frequency offset). An example of a simplified packet structure based on the above channel estimation method is illustrated in Fig. 4.

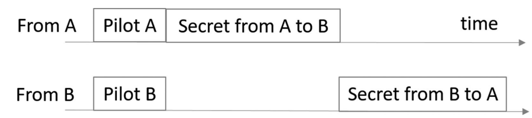

Fig. 4. Shown here is an example of the structure of two packets simultaneously sent by Alice and Bob. Each packet has a pilot and a payload. (The pilot and payload samples of each packet can be interleaved via a common permutation.) The two pilots (i.e., $p_A(k) = p_B(k)$) are used for channel estimation by both Alice and Bob at the same time, and the two payloads (i.e., $x_A(k)$ and $x_B(k)$) contain secret information. Eve also knows the pilots but is unable to estimate its CSI (e.g., see (58) with $p_A(k) = p_B(k)$). Obviously, when equipped with ideal or near-ideal full-duplex radios, the users would not need to schedule their payloads orthogonally in time, which hence further inhibits eavesdropping.

Let us revisit (2) with $P_J = 0$ but for Eves without the knowledge of their CSI with respect to Alice and Bob, i.e.,

$$\mathbf{y}_{A,E}(k) = \sqrt{aP_T}\mathbf{h}_A x_A(k) + \mathbf{n}_{A,E}(k) \tag{59}$$

where $\mathbf{h}_A$ is unknown and $k = 1, \ldots, K$. In order for Eve to obtain any information from Alice, Eve now has to do blind estimation of $x_A(k)$. However, with unknown $\mathbf{h}_A$, the sequence $\{x_A(k)\}$ is not identifiable from the sequence $\{\mathbf{y}_{A,E}(k)\}$ due to a complex scalar ambiguity [35]–[37]. For security against eavesdropping, the ambiguity of the blind problem should be maximally exploited to degrade the capacity of Eve. It is shown in Appendix B that the capacity of Eve without CSI to receive information from a transmitter can be degraded severely by some coding scheme used at the transmitter. In particular, if P_T is large, the secrecy capacity between users over a short time interval can generally be made relatively close to the channel capacity between users.

If Eve is very close to Alice such that $a \gg b$, Eve can estimate $\mathbf{h}_A$ but not $\mathbf{h}_B$. This Eve can potentially (if it also knows the code book used by Alice) receive the secret key sent from Alice in phase 1, but not the secret key sent from Bob in phase 2. A similar statement holds if Eve is very close to Bob. Hence, even in the extreme cases, no Eve can steal both secret keys.

If there is a sufficient radius around each of Alice and Bob within which there is no Eve (i.e., $\max\{\frac{a}{b}, \frac{b}{a}\}$ is not too large), then no Eve can have a good estimate of any of the two vectors $\mathbf{h}_A$ and $\mathbf{h}_B$. In this case, each of the two keys exchanged between Alice and Bob is secure even if all Eves try to collude. This is because physical layer collusion among Eves without knowing their CSI can be made infeasible as discussed in Appendix B.

Property 13: With the proposed two-user channel estimation method in an environment where the CSI of Eves is previously unknown to them, the secrecy capacity of the two users against all Eves over a short time interval can be made close to the channel capacity between the users. ∎

With a single antenna on each user, the short time interval corresponds to a single sample[2] (or two samples) if the payloads are scheduled orthogonally (or concurrently). To increase the amount of secret information, the users could increase their transmission power (or use a repetition code). Alternatively, both users could use smart antennas to change their CSI for each of several new cycles of channel estimation and secret key transmissions. Obviously, to precisely quantify the overall secrecy capacity, one also needs to take into account the overhead of channel estimation.

If Eve knows the large-scale fading of its CSI with respect to Alice and Bob (i.e., a and b) and also knows a reliable statistical model of $\mathbf{h}_A$ and $\mathbf{h}_B$, then Eve can try to estimate $\mathbf{h}_A$ and $\mathbf{h}_B$ based on $\mathbf{y}_E(k)$ received during training or based on an estimate of $\mathbf{h}_E \doteq \sqrt{a}\mathbf{h}_A + \sqrt{b}\mathbf{h}_B$. Such an example is shown in Appendix C. But in practice, the statistical model of a channel at an arbitrary location is very difficult to establish accurately.

With a partial knowledge of $\mathbf{h}_A$, for example, there is in theory a positive capacity $C_{A,E}$ from Alice to Eve. Assuming one sample per coherent period of $\mathbf{h}_A$, a lower bound and an upper bound of $C_{A,E}$ can be translated from equations (61) and (62) in [40] respectively. The lower bound diminishes to zero if the mean of $\mathbf{h}_A$ is much smaller than its variance, and the upper bound corresponds to the case where $\mathbf{h}_A$ is completely known to Eve.

On the other hand, if K in (59) is large (i.e., a large number of samples per coherent period of $\mathbf{h}_A$), then the averaged capacity from Alice to Eve without knowledge of $\mathbf{h}_A$ could in theory approach the capacity from Alice to Eve with knowledge of $\mathbf{h}_A$. This is because the amount of ambiguity associated with an unknown scalar becomes negligible when compared to the total amount of information carried by $x_A(k)$ for large K. One can verify this by reconsidering (88) in Appendix B. For example, if $M_A = M_E = 1$, and $\mathbf{x}(k)$ is i.i.d. and has constant modulus, then the equality in (88) holds (since $\mathbf{Hx}(k)$ is now Gaussian distributed), and the second term in (88) diminishes to zero as K increases, which leads to $C_{A,E} = \log_2(1 + \sigma_x^2\sigma_h^2)$. But it is important to note that this capacity is achievable only if there is a joint coding (between Alice and Eve) over many of such large independent blocks (each corresponding to an independent realization of $\mathbf{H}$). Blind detection is not a trivial problem as shown in [37] where a single block is considered, and no method can resolve the ambiguity inherent in the problem unless there is additional side information.

Once the secret keys have been shared between the users, they could switch back to conventional modes of communications and still maintain secrecy through network-layer cryptography. The frequency band used for the key exchange can also be different from that for normal communications, and hence the required full-duplex radio can be implemented in any frequency band.

In Appendix D, an extension of the above channel estimation method to the case of multi-input multiple-output (MIMO) user channels is presented.

B. Multi-User Case

This section considers a multi-user problem where multiple ($N \geq 3$) full-duplex single-antenna users want to share their secret keys with each other. If the number of users is large and multi-hop transmissions are necessary, the problem is complex and needs to be treated separately. Related works on this subject include [30]–[33]. But the focus here is on a single-hop multi-user wireless network, where every user is within a single transmission range of any other user.

Extending the channel estimation method presented in the previous section, there are now N users transmitting their packets simultaneously. The packet from the ith user consists of a pilot $p_i(k)$ and a payload $x_i(k)$ where $i = 1, \ldots, N$. At the same time that user i transmits $p_i(k)$ and $x_i(k)$, user i also receives

$$y_i(k) = \sum_{j \neq i} h_{j,i}\sqrt{P_T}p_j(k) + \sqrt{\rho P_T}g_i w_i(k) + n_i(k) \tag{60}$$

in the pilot region, and

$$y_i(l) = \sum_{j \neq i} h_{j,i}\sqrt{P_T}x_j(l) + \sqrt{\rho P_T}g_i w_i(l) + n_i(l) \tag{61}$$

in the payload region. Here, $h_{j,i}$ is the CSI (including both large-scale and small-scale fading) from user j to user i, and all other notations are obviously defined. Also at the same time, the signals received by any Eve with M antennas can be

[2] This does not prevent a user from using multiple symbols to transmit a key packet although only one symbol is totally protected.

expressed as

$$\mathbf{y}_E(k) = \sum_{j=1}^{N} \mathbf{h}_{j,E}\sqrt{P_T}p_j(k) + \mathbf{n}_E(k) \tag{62}$$

in the pilot region, and

$$\mathbf{y}_E(l) = \sum_{j=1}^{N} \mathbf{h}_{j,E}\sqrt{P_T}x_j(l) + \mathbf{n}_E(l) \tag{63}$$

in the payload region.

To simplify the notations, one can stack all $y_i(k)$, $k = 1, \ldots, K$, vertically into a column vector $\mathbf{y}_i$, and all $\mathbf{y}_E(k)$, $k = 1, \ldots, K$, horizontally into a matrix $\mathbf{Y}_E$, and rewrite (60) and (62) as

$$\mathbf{y}_i = \sqrt{P_T}\mathbf{P}_i\mathbf{h}_i + \sqrt{\rho P_T}g_i\mathbf{w}_i + \mathbf{n}_i \tag{64}$$

$$\mathbf{Y}_E = \sqrt{P_T}\mathbf{H}_E\mathbf{P}^T + \mathbf{N}_E \tag{65}$$

where $\mathbf{P} = [\mathbf{p}_1, \ldots, \mathbf{p}_N]$ is the $K \times N$ pilot matrix shared by all users, the ith column of which is the pilot sent from user i (i.e., $\mathbf{p}_i = [p_i(1), \ldots, p_i(K)]^T$), $\mathbf{P}_i$ of $K \times (N-1)$ is $\mathbf{P}$ with its ith column removed, $\mathbf{h}_i$ of $(N-1) \times 1$ is the CSI vector of user i with respect to all other users, and the jth column of the $M \times N$ matrix $\mathbf{H}_E$ is the CSI vector $\mathbf{h}_{j,E}$ of Eve with respect to user j. In order for all users to be able to identify their CSI locally, it is necessary and sufficient that each $\mathbf{P}_i$ has the full-column rank $N-1$. In order to make Eve unable to identify its CSI, it is necessary and sufficient that $\mathbf{P}$ (of N columns) has a rank no larger than $N-1$.

An ideal choice of $\mathbf{P}$ is as follows. Let $K = k_0(N-1)$ where k_0 is an integer and $\mathbf{P}$ is

$$\mathbf{P} = [\mathbf{Q}^T, \ldots, \mathbf{Q}^T]^T \tag{66}$$

where the (m,l)th element of the $(N-1) \times N$ matrix $\mathbf{Q}$ is

$$(\mathbf{Q})_{m,l} = e^{-j2\pi\frac{(m-1)(l-1)}{N}} \tag{67}$$

with $1 \le m \le N-1, 1 \le l \le N$, and $j = \sqrt{-1}$. (j is also used elsewhere as an integer which should be clear in the context.) It is easy to verify that the rank of $\mathbf{P}$ is $N-1$; every column of $\mathbf{P}$ has the same norm $\sqrt{K} = \sqrt{k_0(N-1)}$; and the normalized inner product between the ith and lth (distinct) columns of $\mathbf{P}$ is

$$\frac{\mathbf{p}_i^H\mathbf{p}_l}{\|\mathbf{p}_i\|\|\mathbf{p}_l\|} = \frac{-1}{N-1}e^{j2\pi\frac{(i-l)(N-1)}{N}}, \tag{68}$$

the magnitude of which is the same for all pairs of the columns of $\mathbf{P}$. In a sense, all columns of $\mathbf{P}$ are equally spaced from each other in a $(N-1)$-dimensional subspace. This is why the above $\mathbf{P}$ is considered to be ideal.

The MLE of $\mathbf{h}_i$ from $\mathbf{y}_i$ at user i is

$$\hat{\mathbf{h}}_i = \frac{1}{\sqrt{P_T}}\mathbf{P}_i^+\mathbf{y}_i \tag{69}$$

where $\mathbf{P}_i^+ = (\mathbf{P}_i^H\mathbf{P}_i)^{-1}\mathbf{P}_i^H$. Furthermore, one can verify that

$$\mathbf{P}_i^H\mathbf{P}_i = k_0(N\mathbf{I} - \mathbf{e}_i\mathbf{e}_i^H) \tag{70}$$

where $(\mathbf{e}_i)_l = e^{j2\pi\frac{m_{i,l}(N-1)}{N}}$ with $m_{i,l}$ being an integer dependent on $i = 1, \ldots, N$ and $l = 1, \ldots, N-1$. It follows that

$$(\mathbf{P}_i^H\mathbf{P}_i)^{-1} = \frac{1}{k_0N}(\mathbf{I} + \mathbf{e}_i\mathbf{e}_i^H) \tag{71}$$

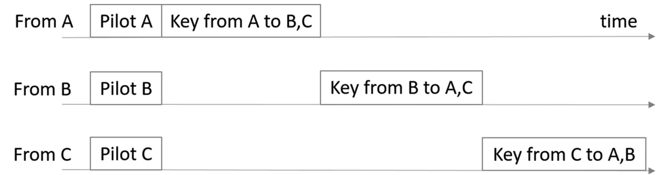

Fig. 5. An example of structure of packets from three users where the payload structure is such that each user can transmit a secret key to all other users in a given payload region.

and hence the computation for $\hat{\mathbf{h}}_i$ is very simple.

The estimate $\hat{\mathbf{h}}_i$ is unbiased and has the covariance matrix $\frac{\rho P_T|g_i|^2+1}{P_T}(\mathbf{P}_i^H\mathbf{P}_i)^{-1}$. The MLE of $|g_i|^2$ is

$$|\hat{g}_i|^2 = \frac{\frac{1}{K}\|\mathbf{y}_i - \sqrt{P_T}\mathbf{P}_i\hat{\mathbf{h}}_i\|^2 - 1}{\rho P_T}. \tag{72}$$

Both $\hat{\mathbf{h}}_i$ and $|\hat{g}_i|^2$ are consistent estimates with respect to k_0.

Because of $\mathbf{P}$ being of rank $N-1$, there is a null vector $\mathbf{h}_0$ such that $\mathbf{P}\mathbf{h}_0 = 0$. Then, $\mathbf{Y}_E$ is unchanged if $\mathbf{H}_E$ is replaced by $\mathbf{H}_E + \mathbf{c}_0\mathbf{h}_0^T$ where $\mathbf{c}_0$ is any $M \times 1$ complex vector. Hence, given $\mathbf{Y}_E$ at Eve, $\mathbf{H}_E$ is not identifiable. If the range of $\mathbf{H}_E^T$ does not include $\mathbf{h}_0$, then $\mathbf{H}_E$ would be identifiable. But for a random $\mathbf{H}_E$, that assumption holds with probability zero.

It should be noted that the above discussion assumes that Eve has no statistical knowledge of $\mathbf{H}_E$, which is often the case in practice especially if Eve does not know its large-scale fading with respect to the users. Otherwise, if an Eve knows the statistics of $\mathbf{H}_E$, the linear minimum mean squared error (LMMSE) estimation of it, for example, can be easily applied, and there is a nonzero capacity from any of the users to the Eve although this capacity can be degraded greatly by the uncertainty of $\mathbf{H}_E$.

If the payloads $x_i(l)$ from all users are not overlapping in time, one can write from (61) that

$$y_i(l_j) = h_{j,i}\sqrt{P_T}x_j(l_j) + n_i(l_j) \tag{73}$$

where $i \neq j$, l_j is the time index where $x_j(l_j) \neq 0$ (but $x_m(l_j) = 0$ with $m \neq j$). Also note that $w_i(l_j) = 0$ due to $x_i(l_j) = 0$. The secrecy capacity from user j to user i against any Eve who does not know $\mathbf{h}_{j,E}$ can be made relatively close to $\log_2(1 + P_T|h_{j,i}|^2)$ subject to a high power and a short period of time as in the two-users case. This choice of non-overlapping payloads is effective for each user to broadcast a secret key to all other users. An example of the packet structure is shown in Fig. 5.

With the multi-user channel estimation method shown above, each user obtains its CSI with respect to all other users. With the reciprocal property, the CSI obtained by each user is the CSI of both to and from all other users. Namely, $h_{j,i}$ for all $j \neq i$ obtained by user i are the same as $h_{i,j}$ for all $j \neq i$. If all users broadcast their estimated CSI using non-overlapping payloads, then every user will know the global CSI. This could allow any subset of users to transmit to another user in a multiple-access fashion in another set of transmissions. An illustration of this idea is shown in Fig. 6. For example, if users 1 and 2 simultaneously send x_1 of power P_1 and x_2 of power P_2 to user 3, then user 3 can receive the information from them at the rate given by $\log_2(1 + |h_{1,3}|^2P_1 + |h_{2,3}|^2P_2)$ [29]. Ideally, the users could also share each other's received signals secretively to

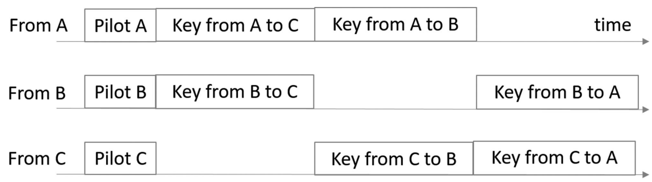

Fig. 6. An example of structure of packets from three users where the payloads are overlapped to allow multiple access transmissions.

form virtual MIMO among them. But this could require excessive overheads due to relatively long pilots required for ANECE. But if all channels stay static after the initial ANECE, and the users can receive each other's information without further use of pilots, then the above is a feasible notion. Naturally, this is under the assumption that Eve cannot perform its channel estimation without receiving a pilot-driven signal.

Shown in Appendix E is an extension of the above channel estimation idea to the MIMO case. Shown in Appendix F is a broadband OFDM-based MIMO channel estimation method against eavesdropping.

In practice, users cannot rely on full-duplex for all transmissions. For example, in order to initiate the proposed ANECE, users may need to communicate with each other to set up the time for channel estimation.[3] If users do not change to a new frequency or new antenna positions (including possibly antenna orientations and polarizations) before they start the ANECE, the packets they sent previously (which typically contain pilots) could be used by Eve for channel estimation. So, users must do something to alter the CSI of Eves before they transmit packets simultaneously for ANECE. If one also needs to prevent Eves from finding their large-scale fading with respect to the users, the users (such as drones) must also move to new locations before the ANECE is conducted at a pre-agreed time. Of course, such a requirement will have an impact on the scope of possible applications.

As shown earlier, no Eve is able to identify its CSI with respect to all users when the users apply the ANECE. But if an Eve is much closer to a user than to all other users, the Eve may be able to approximately identify its CSI with respect to this user and hence may succeed in eavesdropping this user. Furthermore, if an Eve is much closer to a subset of users than to all other users, then the Eve may succeed in eavesdropping this subset of users. In order to avoid this situation, the distribution of the users should be such that the ratio of the maximum and minimum distances (or max-min ratio) between the users is below some pre-specified value. For a given number of points in 2D space, the minimum achievable max-min ratio is studied in [34]. For example, for 30 points in 2D space, the currently known minimum max-min ratio is 5.18. This ratio is generally small enough to prevent any Eve from being much closer to more than one users than to all other users. For a given number of points, the minimum max-min ratio in 3D space is obviously even smaller than that in 2D space. With a fixed path loss exponent, the max-min ratio can be measured in terms of the large-scale fading gains between users. With a fixed distribution of users, a process of measuring the large-scale fading gains and computing the max-min ratio is desirable before the users are selected to participate in a session of the ANECE and the corresponding sharing of secret information.

Provided that the max-min ratio of the users is not too high, the only way that an Eve can succeed in eavesdropping is when the Eve is very close to one of the users. Specifically, around user i, there is a region $\mathcal{A}_i$ within which an Eve could succeed in identifying its CSI with respect to this user and hence in stealing the secret sent by this user. And $\mathcal{A}_i$, and its size $|\mathcal{A}_i|$, are governed by the users that are the closest to user i. The closer are the nearby users around user i, the smaller is $|\mathcal{A}_i|$.

If some Eves are directional, the success of the ANECE requires the directional Eves to be sufficiently far away from the users so that the users appear clustered together from the perspective of any of the directional Eves. If some users are directional, then the directional users should either all face toward or all face away from any given Eve. This appears difficult to realize in practice. This paper is mainly concerned with omnidirectional users and Eves. The following property is easy to prove.

Property 14: Assume a sufficient constraint on the max-min ratio of the user distribution and the use of the proposed multiuser ANECE method. Also assume that the CSI of every Eve changes (or is previously unknown to itself) before the pilots are sent by all N users, the users send secret keys to each other with orthogonal scheduling[4], and user i sends a secret key with the rate R_i (in bits/s/Hz) to all other users over a short time interval sufficient to suppress the capacity of Eves without CSI. Then:

1) If no Eve is close enough to any of the users to be able to identify its CSI, then the rate of secrecy shared by all users (not counting the overhead of pilots) against all Eves which may try to collude is $\frac{1}{N}\sum_{i=1}^{N} R_i$.
2) With respect to non-colluding Eves, some of which may be close enough to some of users, the rate of secrecy shared by all users is no smaller than $\frac{1}{N}(\sum_{i=1}^{N} R_i - \max_j R_j)$.
3) With respect to any group of colluding Eves, if not every user has an Eve close enough to identify its CSI, then the secret from at least one user is safe from eavesdropping and hence the rate of secrecy shared by all users is no smaller than $\frac{1}{N}\min_i R_i$.
4) If Eves are distributed randomly according to the Poisson distribution with λ being the averaged number of Eves in a unit area, and $\mathcal{A}_i$ is the only region where an Eve can steal the secret sent from user i, then the probability for the overall secrecy to be zero, subject to collusion among Eves, is the probability that there is an Eve in every $\mathcal{A}_i$. This probability is $\Pi_{i=1}^{N}(1-e^{-\lambda|\mathcal{A}_i|})$, which goes to zero as N increases.

■

IV. Conclusions

This paper has provided a novel perspective of the potential of full-duplex radio for securing wireless network against passive multi-antenna Eves at unknown locations. The paper shows that if a multi-antenna Eve (or a network of colluding Eves) is allowed to know its CSI with respect to the legitimate users, the secrecy capacity of the single-antenna users could degrade rapidly as the total number of antennas on Eve increases, regardless of the quality of full-duplex radio. This detrimental phenomenon of

[3] An alternative could be that users periodically perform the ANECE.

[4] In the case of concurrent scheduling for all users with ideal full-duplex radios, the factor $\frac{1}{N}$ shown next should be removed.

decreasing secrecy capacity against Eve with increasing number of antennas is similar to that of all conventional setups where Eve knows its CSI, including those of multi-antenna users as in [44]–[45]. Motivated by this finding, this paper also presents a novel method for ANECE. This method allows all users to estimate their own CSI, but at the same time denies any Eve from finding their CSI with respect to any of the users subject to a mild constraint. This paper also shows that without knowing its CSI, Eve with any number of antennas can be virtually disabled over a time window. In other words, by using ANECE, the secrecy capacity between users against Eve with any number of antennas can increase without bound over such time window as the transmission power increases.

To mimic the ANECE property of full-duplex radio by using half-duplex (HD) radio, it would require one or more collaborative HD nodes to perform jamming against Eves[5] during the transmission of pilot from a HD transmitter to a HD receiver. But the unknown locations of Eves and the constraints on jamming interference to the legitimate HD receivers would make the deployment of such collaborative HD nodes highly infeasible. Therefore, it seems reasonable to think that the ANECE method shown in this paper is a ground-breaking discovery useful to drive a full development of full-duplex radio for wireless network security.

Finally, it seems useful to note that explicit channel estimation is not always necessary for users to exchange secret information. Such an example in high SNR channel environment is available in [43]. Further development of this idea with exploitation of full-duplex radio is definitely a logical direction of research. One should also address such important questions as: whether and how can other approaches including [43] challenge the ANECE-based approach against Eve (or a network of colluding Eves) with virtually unlimited number of antennas, and how can the ANECE-based approach be further understood and improved?

APPENDIX

A. Proof of Lemma 1

It is known that $\Phi X_B = \frac{|\mathbf{h}_A^H \mathbf{h}_B|^2}{\|\mathbf{h}_A\|^2}$ which is the magnitude-squared of $\frac{\mathbf{h}_A^H \mathbf{h}_B}{\|\mathbf{h}_A\|} \doteq G_B$. The distribution of G_B is

$$f_{G_B}(g_B) = \int f_{G_B|\mathbf{h}_A}(g_B|\mathbf{h}_A) f_{\mathbf{h}_A}(\mathbf{h}_A) d\mathbf{h}_A. \quad (74)$$

However, given $\mathbf{h}_A$, G_B is complex Gaussian distributed with zero mean and unit variance. (A similar argument was made earlier.) In other words, the distribution of G_B conditional on $\mathbf{h}_A$ is invariant to $\mathbf{h}_A$. Therefore, it follows from (74) that $f_{G_B}(g_B) = f_{G_B|\mathbf{h}_A}(g_B|\mathbf{h}_A)$ and its distribution is $\mathcal{CN}(0,1)$. Hence, $2\Phi X_B$ is the standard Chi-squared of degree two, i.e.,

$$f_{2\Phi X_B}(z) = \frac{1}{2} e^{-\frac{z}{2}}. \quad (75)$$

Since Φ and X_B are independent, the PDF of their product can be shown to be $f_{2\Phi X_B}(z) = \int_0^1 f_\Phi(x) f_{2X_B}(\frac{z}{x}) \frac{1}{x} dx$. Therefore,

$$\frac{1}{2} e^{-\frac{z}{2}} = \int_0^1 f_\Phi(x) \frac{1}{2^M \Gamma(M)} \left(\frac{z}{x}\right)^{M-1} e^{-\frac{z}{2x}} \frac{1}{x} dx. \quad (76)$$

Let $\mathbf{v}_A$ and $\mathbf{v}_B$ be two $N \times 1$ real-valued Gaussian random vectors with the distribution $\mathcal{N}(\mathbf{0}, \mathbf{I})$, it is known from equation (22) in [26] that $Q_v \doteq \frac{\mathbf{v}_A^T \mathbf{v}^B}{\|\mathbf{v}_A\| \|\mathbf{v}_B\|}$ has the following distribution:

$$f_{Q_v}(x) = \frac{1}{\sqrt{\pi}} \frac{\Gamma(\frac{N}{2})}{\Gamma(\frac{N-1}{2})} (1-x^2)^{\frac{N-3}{2}} \quad (77)$$

where $|x| < 1$ (and $N \geq 2$), and hence $P_v \doteq Q_v^2$ has the following distribution

$$f_{P_v}(x) = f_{Q_v}(\sqrt{x}) \frac{1}{\sqrt{x}} = \frac{1}{\sqrt{\pi}} \frac{\Gamma(\frac{N}{2})}{\Gamma(\frac{N-1}{2})} (1-x)^{\frac{N-3}{2}} \frac{1}{\sqrt{x}} \quad (78)$$

where $0 < x < 1$. Note that the first equation in the above has applied the fact that $f_{Q_v}(x)$ is symmetric. Furthermore, it is known that $P_v \|\mathbf{v}_B\|^2 = \frac{|\mathbf{v}_A^T \mathbf{v}_B|^2}{\|\mathbf{v}_A\|^2}$ which is the standard Chi-squared of degree one, i.e.,

$$f_{P_v \|\mathbf{v}_B\|^2}(x) = \frac{1}{2^{\frac{1}{2}} \Gamma(\frac{1}{2})} x^{-\frac{1}{2}} e^{-\frac{x}{2}} \quad (79)$$

and $\|\mathbf{v}_B\|^2$ is the standard Chi-squared of degree N, i.e.,

$$f_{\|\mathbf{v}_B\|^2}(x) = \frac{1}{2^{\frac{N}{2}} \Gamma(\frac{N}{2})} x^{\frac{N}{2}-1} e^{-\frac{x}{2}}. \quad (80)$$

It follows from $f_{P_v \|\mathbf{v}_B\|^2}(z) = \int_0^1 f_{P_v}(x) f_{\|\mathbf{v}_B\|^2}\left(\frac{z}{x}\right) \frac{1}{x} dx$ that

$$\frac{1}{2^{\frac{1}{2}} \Gamma(\frac{1}{2})} z^{-\frac{1}{2}} e^{-\frac{z}{2}} = \int_0^1 \frac{1}{\sqrt{\pi}} \frac{\Gamma(\frac{N}{2})}{\Gamma(\frac{N-1}{2})} (1-x)^{\frac{N-3}{2}} \times \frac{1}{\sqrt{x}} \frac{1}{2^{\frac{N}{2}} \Gamma(\frac{N}{2})} \left(\frac{z}{x}\right)^{\frac{N}{2}-1} e^{-\frac{z}{2x}} \frac{1}{x} dx \quad (81)$$

which is equivalent to

$$\frac{1}{2} e^{-\frac{z}{2}} = \int_0^1 \frac{1}{2^{\frac{N+1}{2}}} \frac{1}{\Gamma(\frac{N-1}{2})} (1-x)^{\frac{N-3}{2}} \left(\frac{z}{x}\right)^{\frac{N-1}{2}} e^{-\frac{z}{2x}} \frac{1}{x} dx. \quad (82)$$

Comparing (76) against (82) with $N = 2M - 1$ yields (37).

B. Capacity of Multi-Antenna Eve Without CSI

As a generalization of (59), assume that an Eve of M_E antennas receives the following signals from Alice with M_A antennas:

$$\mathbf{y}(k) = \mathbf{H}\mathbf{x}(k) + \mathbf{n}(k) \quad (83)$$

where $k = 1, \ldots, K$, the k-independent $M_E \times M_A$ channel matrix $\mathbf{H}$ is unknown to Eve but modelled to be such that $\mathbf{h} = vec(\mathbf{H})$ (stacking columns of $\mathbf{H}$ vertically) is $\mathcal{CN}(\mathbf{m}, \mathbf{R}_h)$, the noise $\mathbf{n}(k)$ is $\mathcal{CN}(\mathbf{0}, \mathbf{I}_{M_E})$, and $\mathbf{x}(k)$ has zero mean and the correlation matrix $\mathcal{E}\{\mathbf{x}(l)\mathbf{x}(m)^H\} = \sigma_x^2 \delta_{l,m} \mathbf{I}_{M_A}$. By stacking the K equations in (83) horizontally, one has $\mathbf{Y} = \mathbf{HX} + \mathbf{N}$. Like $\mathbf{h}$, also define $\mathbf{y} = vec(\mathbf{Y})$, $\mathbf{x} = vec(\mathbf{X})$ and $\mathbf{n} = vec(\mathbf{N})$.

[5]This paper focuses on passive Eves. An active Eve could use jamming to prevent the legitimate users from finding their CSI. But a wrong estimate of CSI by a receiver can be easily detected if the packet has some embedded check-bits.

It follows that

$$\mathbf{y} = (\mathbf{X}^T \otimes \mathbf{I}_{M_E})\mathbf{h} + \mathbf{n} \quad (84)$$

$$= (\mathbf{I}_K \otimes \mathbf{H})\mathbf{x} + \mathbf{n}. \quad (85)$$

From [38], one knows that the capacity in bits/s/Hz from Alice to Eve is $C_{A,E} = \frac{1}{K} I(\mathbf{x};\mathbf{y}) = \frac{1}{K}(h(\mathbf{y}) - h(\mathbf{y}|\mathbf{x})) \leq \frac{1}{K}(\log_2 |\mathbf{R}_y| - \mathcal{E}\{\log_2 |\mathbf{R}_y(\mathbf{x})|\})$ where the inequality comes from the differential entropy of $\mathbf{y}$, i.e., $h(\mathbf{y}) \leq \log_2((2\pi e)^{M_E K}|\mathbf{R}_y|)$ as $\mathbf{HX}$ is likely non-Gaussian. Here, $\mathbf{R}_y$ is the covariance matrix of $\mathbf{y}$, i.e., $\mathbf{R}_y = \sigma_x^2(\mathbf{I}_K \otimes \mathcal{E}\{\mathbf{H}\mathbf{H}^H\}) + \mathbf{I}_{M_E K}$ which follows from (85), and $\mathbf{R}_y(\mathbf{x})$ is the covariance matrix of $\mathbf{y}$ conditional on $\mathbf{x}$, i.e., $\mathbf{R}_y(\mathbf{x}) = (\mathbf{X}^T \otimes \mathbf{I}_{M_E})\mathbf{R}_h(\mathbf{X}^* \otimes \mathbf{I}_{M_E}) + \mathbf{I}_{M_E K}$ which follows from (84). Let the mean of $\mathbf{H}$ be $\mathbf{M}$, and the ith $M_E \times M_E$ diagonal block of $\mathbf{R}_h$ be $\mathbf{R}_{h,i}$. Then, $\mathcal{E}\{\mathbf{H}\mathbf{H}^H\} = \mathbf{M}\mathbf{M}^H + \sum_{i=1}^{M_A} \mathbf{R}_{h,i}$. It follows that

$$|\mathbf{R}_y| = |\sigma_x^2\left(\mathbf{M}\mathbf{M}^H + \sum_{i=1}^{M_A}\mathbf{R}_{h,i}\right) + \mathbf{I}_{M_E}|^K \quad (86)$$

$$|\mathbf{R}_y(\mathbf{x})| = |\mathbf{R}_h^{1/2}(\mathbf{X}^*\mathbf{X}^T \otimes \mathbf{I}_{M_E})\mathbf{R}_h^{1/2} + \mathbf{I}_{M_E M_A}| \quad (87)$$

where for the second equation, the fact $|\mathbf{A}\mathbf{A}^T + \mathbf{I}| = |\mathbf{A}^T\mathbf{A} + \mathbf{I}|$ has been applied.

For unknown $\mathbf{H}$, one can assume that $\mathbf{M}\mathbf{M}^H \ll \sum_{i=1}^{M_A}\mathbf{R}_{h,i}$. For convenience, let us also assume that $\mathbf{R}_h = \sigma_h^2\mathbf{I}_{M_E M_A}$. It then follows that

$$C_{A,E} \leq C_{A,E,up} \doteq \frac{1}{K}(\log_2|\mathbf{R}_y| - \mathcal{E}\{\log_2|\mathbf{R}_y(\mathbf{x})|\}) \quad (88)$$

where $|\mathbf{R}_y| = (M_A\sigma_x^2\sigma_h^2 + 1)^{M_E K}$ and $|\mathbf{R}_y(\mathbf{x})| = |\sigma_h^2\mathbf{X}^*\mathbf{X}^T + \mathbf{I}_{M_A}|^{M_E}$.

One can now verify the following:

1) If $K = M_A = 1$ and the symbol from Alice has a constant modulus, then $C_{A,E,up} = 0$.
2) If $K = M_A \gg 1$, then $\mathbf{X}^*\mathbf{X}^T \approx K\sigma_x^2\mathbf{I}_{M_A}$ and hence $C_{A,E,up} \approx 0$.
3) For any fixed $K = M_A$ but a large $\sigma_h^2\sigma_x^2$, $C_{A,E,up}$ becomes a constant independent of σ_h^2 and σ_x^2 while $C_{A,B}$ is always independent of σ_h^2 and scales with σ_x^2 as $\mathcal{O}(\log_2 \sigma_x^2)$, and hence one can achieve $C_{A,B} \gg C_{A,E,up}$.

The above suggests that for Eve with unknown CSI, its capacity to receive information over a short period of time (i.e., $K = M_A$) can be degraded severely (if not completely) regardless of the number M_E of antennas on Eve. Also note that the above analysis applies to the case where multiple or many (colluding) Eves combine their received signals to form a large antenna array. Such collusion would result in the same signal model as in (83) except that M_E is increased.

For $K > M_A$, there are still coding schemes that prevent Eve without CSI from obtaining any information. For example, if a constant modulus repetition code with any $K > M_A = 1$ is used, $C_{A,E}$ can be shown to be zero.

On the other hand, if $\mathbf{x}(k)$ for $k = 1, \ldots, K$ are independent and K is large, then the amount of information (or ambiguity) carried by $\mathbf{H}$ becomes less significant compared to that carried by $\mathbf{x}(k)$ for $k = 1, \ldots, K$. To suppress the capacity of Eve, K should be small in general. One may choose $K = M_A$ but some large σ_x^2 to achieve a sufficient amount of secrecy.

C. Estimation of $\mathbf{h}_A$ and $\mathbf{h}_B$ From Estimate of $\mathbf{h}_E$

Let a and b be known to Eve, and $\hat{\mathbf{h}}_E$ be the estimate of $\mathbf{h}_E \doteq \sqrt{a}\mathbf{h}_A + \sqrt{b}\mathbf{h}_B$ obtained by Eve from $\{\mathbf{y}_E(k), k = 1, \ldots, K\}$ in (58). Assume $\hat{\mathbf{h}}_E = \sqrt{a}\mathbf{h}_A + \sqrt{b}\mathbf{h}_B + \mathbf{e}$ where $\mathbf{h}_A$ and $\mathbf{h}_B$ are independent of each other, and each has the distribution $\mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$. Also assume $\mathbf{e}$ is $\mathcal{CN}(\mathbf{0}, \sigma_e^2\mathbf{I}_M)$. Then, one can verify that the minimum-mean-squared-error (MMSE) estimate of $\mathbf{h}_A$ given $\hat{\mathbf{h}}_E$ is $\hat{\mathbf{h}}_A = \frac{\sqrt{a}}{a+b+\sigma_e^2}\hat{\mathbf{h}}_E$, and its MSE matrix is $\mathcal{E}\{(\hat{\mathbf{h}}_A - \mathbf{h}_A)(\hat{\mathbf{h}}_A - \mathbf{h}_A)^H\} = \frac{b+\sigma_e^2}{a+b+\sigma_e^2}\mathbf{I}_M$. Similar expressions hold for the MMSE estimation of $\mathbf{h}_B$ given $\hat{\mathbf{h}}_E$. Note that only when a and b are known to Eve, is the above estimation applicable by Eve.

D. ANECE for Two-User MIMO Channel

Let Alice, Bob and Eve all have multiple antennas. The channel matrix from Alice to Bob is $\mathbf{H}_{A,B}$ (including both large-scale and small-scale fading), and those between other nodes are defined similarly. Let Alice and Bob transmit the pilot vectors $\mathbf{p}_A(k)$ and $\mathbf{p}_B(k)$ respectively and simultaneously. Then, Alice, Bob and Eve receive the following signals:

$$\mathbf{y}_A(k) = \sqrt{P_T}\mathbf{H}_{B,A}\mathbf{p}_B(k) + \sqrt{\rho P_T}\mathbf{G}_A\mathbf{w}_A(k) + \mathbf{n}_A(k) \quad (89)$$

$$\mathbf{y}_B(k) = \sqrt{P_T}\mathbf{H}_{A,B}\mathbf{p}_A(k) + \sqrt{\rho P_T}\mathbf{G}_B\mathbf{w}_B(k) + \mathbf{n}_B(k) \quad (90)$$

$$\mathbf{y}_E(k) = \sqrt{P_T}\mathbf{H}_{A,E}\mathbf{p}_A(k) + \sqrt{P_T}\mathbf{H}_{B,E}\mathbf{p}_B(k) + \mathbf{n}_E(k) \quad (91)$$

where $\sqrt{\rho P_T}\mathbf{G}_A\mathbf{w}_A(k)$ is the residual self-interference noise at Alice after the cancellation of the self-interference caused by $\mathbf{p}_A(k)$, $\mathbf{n}_A$ is the background noise, and other notations are similarly defined. Also assume that $\mathbf{w}_A$ is white and independent of $\mathbf{p}_A(k)$, and $\mathbf{w}_B$ is white and independent of $\mathbf{p}_B(k)$. Due to mutual couplings between antennas, $\mathbf{G}_A$ and $\mathbf{G}_B$ are not diagonal in general.

Let us first consider $\mathbf{y}_E(k)$ with $k = 1, \ldots, K$, which can be rewritten into:

$$\mathbf{Y}_E = \sqrt{P_T}\mathbf{H}_E\mathbf{P}_E + \mathbf{N}_E \quad (92)$$

where the kth column of $\mathbf{Y}_E$ is $\mathbf{y}_E(k)$, $\mathbf{H}_E = [\mathbf{H}_{A,E}, \mathbf{H}_{B,E}]$, and the kth column of $\mathbf{P}_E$ is $\left[\mathbf{p}_A^T(k), \mathbf{p}_B^T(k)\right]^T$. The necessary and sufficient condition for Eve to identify $\mathbf{H}_E$ from $\mathbf{Y}_E$ and $\mathbf{P}_E$ (uniquely in the absence of noise) is that $\mathbf{P}_E$ has a full row rank. If one chooses $\mathbf{p}_A(k) = \mathbf{p}_B(k)$, then the rank of $\mathbf{P}_E$ is no larger than half of the number of rows in $\mathbf{P}_E$, and hence $\mathbf{H}_E$ is not identifiable by Eve. Therefore, Alice and Bob can easily follow some simple protocol to make sure that the CSI at Eve is not identifiable by Eve. If Alice has M_A antennas and Bob has $M_B \leq M_A$ antennas, one can choose $\mathbf{p}_A(k)$ and $\mathbf{p}_B(k)$ such that the first M_A rows of $\mathbf{P}_E$ are independent, the last M_B rows of $\mathbf{P}_E$ are also independent, but the span of the latter belongs to that of the former.

The signals received by Alice can be rewritten as

$$\mathbf{Y}_A = \sqrt{P_T}\mathbf{H}_{B,A}\mathbf{P}_B + \mathbf{V}_A \tag{93}$$

where the kth column of $\mathbf{Y}_A$ is $\mathbf{y}_A(k)$, the kth column of $\mathbf{P}_B$ is $\mathbf{p}_B(k)$, and $\mathbf{V}_A = \sqrt{\rho P_T}\mathbf{G}_A\mathbf{W}_A + \mathbf{N}_A$ contains all the noises. Obviously, provided that rows of $\mathbf{P}_B$ are independent, $\mathbf{H}_{B,A}$ is identifiable from $\mathbf{Y}_A$ and $\mathbf{P}_B$. Similar statements can be made for estimation of $\mathbf{H}_{A,B}$ at Bob. If the reciprocal property holds, one can write $\mathbf{H}_{B,A} = \mathbf{H}_{A,B}^T$ which is however not required in this paper. (The reciprocal property is known to hold for the electro-magnetics in air. Although it may not hold for a radio transceiver, the channel gains on a transceiver can be pre-calibrated and compensated.)

Now consider the channel estimation at Alice, where both $\mathbf{H}_{B,A}$ and $\mathbf{G}_A$ are unknowns, and $\mathbf{w}_A(k)$ and $\mathbf{n}_A(k)$ are independent $\mathcal{CN}(\mathbf{0}, \mathbf{I})$. Obviously, given $\mathbf{y}_A(k)$ and $\mathbf{p}_B(k)$ for all k, $\mathbf{G}_A$ is ambiguous up to a right unitary matrix. One can further simplify $\mathbf{y}_A(k)$ as follows:

$$\mathbf{y}_A(k) = \sqrt{P_T}\mathbf{H}_{B,A}\mathbf{p}_B(k) + \mathbf{v}_A(k) \tag{94}$$

where $\mathbf{v}_A(k)$ is $\mathcal{CN}(\mathbf{0}, \mathbf{R}_A)$ with $\mathbf{R}_A = \rho P_T \mathbf{G}_A\mathbf{G}_A^H + \mathbf{I}$ being unknown covariance matrix. For any given $\mathbf{R}_A$, one can write

$$\begin{aligned}\mathbf{R}_A^{-\frac{1}{2}}\mathbf{y}_A(k) &= \sqrt{P_T}\mathbf{R}_A^{-\frac{1}{2}}\mathbf{H}_{B,A}\mathbf{p}_B(k) + \mathbf{R}_A^{-\frac{1}{2}}\mathbf{v}_A(k)\\ &= \sqrt{P_T}(\mathbf{p}_B^T(k) \otimes \mathbf{R}_A^{-\frac{1}{2}})\mathbf{h}_{B,A} + \tilde{\mathbf{v}}_A(k)\end{aligned} \tag{95}$$

where $\otimes$ is the Kronecker product, $\mathbf{h}_{B,A} = vec\{\mathbf{H}_{B,A}\}$ (stacking the columns of $\mathbf{H}_{B,A}$ vertically), and $\tilde{\mathbf{v}}_A(k)$ is $\mathcal{CN}(\mathbf{0}, \mathbf{I})$. And the MLE of $\mathbf{h}_{B,A}$ is

$$\begin{aligned}\hat{\mathbf{h}}_{B,A} &= \frac{1}{\sqrt{P_T}}\left(\sum_{k=1}^{K}(\mathbf{p}_B^*(k)\mathbf{p}_B^T(k)) \otimes \mathbf{R}_A^{-1}\right)^{-1}\\ &\quad \times \left(\sum_{k=1}^{K}\mathbf{p}_B^*(k) \otimes \mathbf{R}_A^{-1}\mathbf{y}_A(k)\right)\\ &= \frac{1}{\sqrt{P_T}}\sum_{l=1}^{K}\left(\mathbf{S}_B^{-1}\mathbf{p}_B^*(l) \otimes \mathbf{y}_A(l)\right)\end{aligned} \tag{96}$$

where $\mathbf{S}_B = \sum_{k=1}^{K}\mathbf{p}_B^*(k)\mathbf{p}_B^T(k) = (\mathbf{P}_B\mathbf{P}_B^H)^*$. It turns out that $\hat{\mathbf{h}}_{B,A}$ is invariant to $\mathbf{R}_A$. With any given $\mathbf{h}_{B,A} = vec\{\mathbf{H}_{B,A}\}$, the MLE of $\mathbf{R}_A$ is

$$\begin{aligned}\hat{\mathbf{R}}_A &= \frac{1}{K}\sum_{k=1}^{K}(\mathbf{y}_A(k) - \sqrt{P_T}\mathbf{H}_{B,A}\mathbf{p}_B(k))\\ &\quad \times (\mathbf{y}_A(k) - \sqrt{P_T}\mathbf{H}_{B,A}\mathbf{p}_B(k))^H.\end{aligned} \tag{97}$$

Therefore, the (exact) joint MLE of $\mathbf{H}_{B,A}$ and $\mathbf{R}_A$ can be computed in two steps. In step 1, compute the MLE of $\mathbf{h}_{B,A} = vec\{\mathbf{H}_{B,A}\}$ by (96). In step 2, compute the MLE of $\mathbf{R}_A$ by (97) with $\mathbf{H}_{B,A}$ replaced by its MLE. One can verify that $\mathcal{E}\{\hat{\mathbf{h}}_{B,A}\} = \mathbf{h}_{B,A}$ and

$$Cov\{\hat{\mathbf{h}}_{B,A}\} = \frac{1}{P_T}(\mathbf{P}_B\mathbf{P}_B^H)^{*-1} \otimes \mathbf{R}_A. \tag{98}$$

The optimal choice of the pilot matrix $\mathbf{P}_B$ subject to unknown $\mathbf{R}_A$ is such that $\mathbf{P}_B\mathbf{P}_B^H$ is proportional to the identity matrix, e.g., $\mathbf{P}_B\mathbf{P}_B^H = K\mathbf{I}$.

The channel estimation algorithm carried out by Bob is symmetrical to that by Alice. The algorithm shown in Section A is a special case of the above.

E. ANECE for Multi-User MIMO Channels

Now consider $N \geq 3$ users where user i has n_i antennas. For channel estimation, all users send their pilots simultaneously. Let user j send $\sqrt{P_T}\mathbf{p}_j(k)$ of $n_j \times 1$ with $k = 1, \dots, K$ and $j = 1, \dots, N$. Then, user i receives

$$\mathbf{y}_i(k) = \sqrt{P_T}\sum_{j \neq i}\mathbf{H}_{j,i}\mathbf{p}_j(k) + \sqrt{\rho P_T}\mathbf{G}_i\mathbf{w}_i(k) + \mathbf{n}_i(k) \tag{99}$$

where the notations are defined in a similar way as in the previous section. The signal received by any given Eve is

$$\mathbf{y}_E(k) = \sqrt{P_T}\sum_{j=1}^{N}\mathbf{H}_{j,E}\mathbf{p}_j(k) + \mathbf{n}_E(k). \tag{100}$$

The matrix forms of the above two equations are

$$\mathbf{Y}_i = \sqrt{P_T}\mathbf{H}_i\mathbf{P}^{(i)} + \mathbf{V}_i \tag{101}$$

$$\mathbf{Y}_E = \sqrt{P_T}\mathbf{H}_E\mathbf{P}^T + \mathbf{N}_E \tag{102}$$

where the kth column of $\mathbf{Y}_i$ is $\mathbf{y}_i(k)$, the kth column of $\mathbf{Y}_E$ is $\mathbf{y}_E(k)$, $\mathbf{H}_i$ is stacked horizontally from all $\mathbf{H}_{j,i}$ with all $j \neq i$, $\mathbf{H}_E$ is stacked horizontally from all $\mathbf{H}_{j,E}$ with all $j = 1, \dots, N$, $\mathbf{P}^T$ is of $N_T \times K$ with $N_T = \sum_{i=1}^{N} n_i$, $\mathbf{P}^T = [\mathbf{P}_1, \dots, \mathbf{P}_N]^T$, $\mathbf{P}_i^T = [\mathbf{p}_i(1), \dots, \mathbf{p}_i(K)]$, and $\mathbf{P}^{(i)}$ is $\mathbf{P}^T$ with its ith block $\mathbf{P}_i^T$ removed. Also, $\mathbf{V}_i = \sqrt{\rho P_T}\mathbf{G}_i\mathbf{W}_i + \mathbf{N}_i$ and $\mathbf{N}_E$ are the noises.

The structure of (101) is identical to (93). Hence, the MLE method shown there can be directly used here. Now, it is only necessary to focus on the conditions required for $\mathbf{P}^{(i)}$ and $\mathbf{P}$. Clearly, one needs $\mathbf{P}$ of $K \times N_T$ to be of rank less than N_T so that $\mathbf{H}_E$ of $n_E \times N_T$ is not identifiable from $\mathbf{Y}_E$. On the other hand, for any $i = 1, \dots, N$, one needs $\mathbf{P}^{(i)}$ of $(N_T - n_i) \times K$ to be of the full-column rank $N_T - n_i$ so that $\mathbf{H}_i$ of $n_i \times (N_T - n_i)$ can be identified from $\mathbf{Y}_i$. Also, the smaller is the condition number (i.e., the ratio of the largest singular value over the smallest singular value) of $\mathbf{P}^{(i)}$, the more robust is the MLE of $\mathbf{H}_i$ against noise. This is because the covariance matrix of the MLE of $vec\{\mathbf{H}_i\}$ is proportional to the conjugate of $(\mathbf{P}^{(i)}\mathbf{P}^{(i)H})^{-1}$. See (98).

To meet the above constraints, it is proposed to construct $\mathbf{P}$ as follows. Let $1 \leq \tilde{n} \leq \min_i n_i$ and $K = k_0(N_T - \tilde{n})$ with k_0 being an integer. Then, choose $\mathbf{P}^T = [\mathbf{Q}^T, \dots, \mathbf{Q}^T]$ with k_0 identical blocks. Choose the (m, l)th element of the $(N_T - \tilde{n}) \times N_T$ matrix $\mathbf{Q}$ to be $e^{-j2\pi\frac{(m-1)(l-1)}{N_T}}$ with $1 \leq m \leq (N_T - \tilde{n})$ and $1 \leq l \leq N_T$. This $\mathbf{P}$ has the deficient rank $N_T - \tilde{n}$, which can be verified by using a property of Vandermonde matrix. ($\mathbf{Q}$ is a submatrix of a Vandermonde matrix.) Furthermore, any of its sub-matrices, $\mathbf{P}^{(i)}$ with $i = 1, \dots, N$, has the full-column rank $N_T - n_i$. One can also verify that the normalized inner product between the mth and lth (distinct) columns of $\mathbf{P}$ is $\frac{-1}{N_T-\tilde{n}}\sum_{k=1}^{\tilde{n}} e^{j2\pi\frac{(m-l)(N_T-k)}{N_T}}$.

If $\tilde{n}$ is replaced by one, then every pair of columns of $\mathbf{P}$ has the smallest magnitude $\frac{1}{N_T-1}$ of the normalized inner product, $\mathbf{P}$ has the deficient rank N_T-1, and $\mathbf{P}^{(i)}$ has the full rank N_T-n_i and also, as shown next, the best possible condition (i.e., smallest condition number). To prove that $\mathbf{P}^{(i)}$ has the smallest condition number when $\tilde{n}=1$, let us consider the $(N_T-n_i)\times(N_T-n_i)$ matrix $\mathbf{P}^{(i)}\mathbf{P}^{(i)H}$ which can be written as

$$\mathbf{P}^{(i)}\mathbf{P}^{(i)H} = k_0 N_T \mathbf{I} - k_0 \sum_{k=1}^{\tilde{n}} \mathbf{e}_i^{(k)}\mathbf{e}_i^{(k)H} \tag{103}$$

where $(\mathbf{e}_i^{(k)})_l = e^{-j2\pi\frac{m_{i,l}(N_T-k)}{N_T}}$ and $m_{i,l}$ is an integer dependent on $i=1,\ldots,N$ and $l=1,\ldots,N_T-n_i$. It is obvious that the largest and smallest eigenvalues of $\mathbf{P}^{(i)}\mathbf{P}^{(i)H}$ are $\lambda_{max}(\mathbf{P}^{(i)}\mathbf{P}^{(i)H}) = k_0N_T$ and $\lambda_{min}(\mathbf{P}^{(i)}\mathbf{P}^{(i)H}) = k_0N_T - k_0\lambda_{max}(\sum_{k=1}^{\tilde{n}}\mathbf{e}_i^{(k)}\mathbf{e}_i^{(k)H}) \le k_0N_T - k_0\lambda_{max}(\mathbf{e}_i^{(1)}\mathbf{e}_i^{(1)H}) = k_0N_T - k_0(N_T-n_i) = k_0n_i$. Then, $\frac{\lambda_{max}(\mathbf{P}^{(i)}\mathbf{P}^{(i)H})}{\lambda_{min}(\mathbf{P}^{(i)}\mathbf{P}^{(i)H})} \ge \frac{N_T}{n_i}$ where the equality holds when $\tilde{n}=1$.

If $\min_i n_i > 1$, then there are trade-offs as $\tilde{n}$ varies from 1 to $\min_i n_i$. As $\tilde{n}$ increases, the condition number of $\mathbf{P}^{(i)}$ increases (bad for users) but there are more degrees of freedoms in $\mathbf{H}_E$ given $\mathbf{Y}_E$ and $\mathbf{P}$ (bad for Eves).

F. ANECE for Multi-User Broadband MIMO Channels

If the channels between users are convolutive, then one can adopt the OFDM approach as follows. Each packet sent from a user has two regions: the pilot region for channel estimation and the payload region containing secret information. The pilot region is divided into K epoches, and the payload region is also divided into multiple epoches. Each epoch of the pilot region has N_s time slots, and each epoch of the payload region has N_c+N_s time slots, where $N_c \gg N_s$, N_c is the number of subcarriers, and N_sT_s is the maximum possible delay spread of the channel responses with $\frac{1}{T_s}$ being the channel bandwidth utilized. All time slots of the kth epoch in the pilot region of user i contain zeros except that the first slot is assigned with $\mathbf{p}_i(k)$. The (discrete-time) channel response from user j to user i is $\tilde{\mathbf{H}}_{j,i}(l)$ with $l=0,1,\ldots,N_s-1$. The N_c-point DFT of this response is $\mathbf{H}_{j,i}(c)$ where $c=0,1,\ldots,N_c-1$. One can verify that in the kth epoch of the pilot region, user i receives

$$\tilde{\mathbf{y}}_i(k,l) = \sqrt{P_T}\sum_{j\neq i}\tilde{\mathbf{H}}_{j,i}(l)\mathbf{p}_j(k) + \sqrt{\rho P_T}\tilde{\mathbf{G}}_{j,i}(l)\mathbf{w}_j(k) + \tilde{\mathbf{n}}_i(k,l) \tag{104}$$

with $l=0,1,\ldots,N_s-1$ and $k=1,\ldots,K$. Taking the N_c-point DFT of $\tilde{\mathbf{y}}_i(k,l)$ with respect to l yields

$$\mathbf{y}_i(k,c) = \sqrt{P_T}\sum_{j\neq i}\mathbf{H}_{j,i}(c)\mathbf{p}_j(k) + \sqrt{\rho P_T}\mathbf{G}_{j,i}(c)\mathbf{w}_j(k) + \mathbf{n}_i(k,c) \tag{105}$$

where $c=0,1,\ldots,N_c-1$. Similarly, in the frequency domain, Eve receives

$$\mathbf{y}_E(k,c) = \sqrt{P_T}\sum_{j=1}^{N}\mathbf{H}_{j,E}(c)\mathbf{p}_j(k) + \mathbf{n}_E(k,c). \tag{106}$$

For each c, the channel estimation problem is the same as that of (99) and (100).

For each epoch of the payload region, each user encodes its information by following the conventional OFDM fashion, but the information from different users should be non-overlapping in time (or otherwise multiple-access coding is required). The conventional OFDM-based channel equalization applies here straightforwardly.

References

[1] T. Riihonen, S. Werner, and R. Wichman, "Mitigation of loopback self-interference in full-duplex MIMO relays," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5983–5993, Dec. 2011.

[2] Y. Hua, P. Liang, Y. Ma, A. Cirik, and Q. Gao, "A method for broadband full-duplex MIMO radio," *IEEE Signal Process. Lett.*, vol. 19, no. 12, pp. 793–796, Dec. 2012.

[3] Y.-S. Choi and H. Shirani-Mehr, "Simultaneous transmission and reception: Algorithm, design and system level performance," *IEEE Trans. Wireless Commun.*, vol. 12, no. 12, pp. 5992–6014, Dec. 2013.

[4] Y. Hua, Y. Ma, A. Gholian, Y. Li, A. Ciric, and P. Liang, "Radio self-interference cancellation by transmit beamforming, all-analog cancellation and blind digital tuning," *Signal Process.*, vol. 108, pp. 322–340, 2015.

[5] E. Ahmed, A. M. Eltawil, Z. Li, and B. A. Cetiner, "Full-duplex systems using multireconfigurable antennas," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 5971–5983, Nov. 2015.

[6] H. Krishnaswamy and G. Zussman, "1 chip 2x the bandwidth," *IEEE Spectrum*, vol. 53, no. 7, pp. 38–54, Jul. 2016.

[7] B. P. Day, A. R. Margetts, D. W. Bliss, and P. Schniter, "Full-duplex bidirectional MIMO: Achievable rates under limited dynamic range," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3702–3713, Jul. 2012.

[8] D. Nguyen, L.-N. Tran, P. Pirinen, and M. Latva-aho, "Precoding for full duplex multiuser MIMO systems: Spectral and energy efficiency maximization," *IEEE Trans. Signal Process.*, vol. 61, no. 16, pp. 4038–4050, Aug. 2013.

[9] A. Cirik, Y. Rong, and Y. Hua, "Achievable rates and QoS considerations of full-duplex MIMO radios for fast fading channels with imperfect channel estimation," *IEEE Trans. Signal Process.*, vol. 62, no. 15, pp. 3874–3886, Aug. 2014.

[10] A. Cirik, R. Wang, Y. Rong, and Y. Hua, "MSE based transceiver designs for full-duplex MIMO cognitive radios," *IEEE Trans. Commun.*, vol. 63, no. 6, pp. 2056–2070, Jun. 2015.

[11] L. Chen, W. Meng, and Y. Hua, "Optimal power allocation for a full-duplex multicarrier decode-forward relay system with or without direct link," *Signal Process.*, vol. 137, pp. 177–191, 2017.

[12] 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. 2013.

[13] G. Chen, Y. Gong, P. Xiao, and J. A. Chambers, "Physical layer network security in the full-duplex relay system," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 3, pp. 574–583, Mar. 2015.

[14] G. Chen, Y. Gong, P. Xiao, and J. A. Chambers, "Dual antenna selection in secure cognitive radio networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 7993–8002, Oct. 2016.

[15] 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. 2017.

[16] T.-X. Zheng, H.-M. Wang, Q. Yang, and M. H. Lee, "Safeguarding decentralized wireless networks using full-duplex jamming receivers," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 278–292, Jan. 2017.

[17] M. R. Abedi, N. Mokari, H. Saeedi, and H. Yanikomeroglu, "Robust resource allocation to enhance physical layer security in systems with full-duplex receivers: Active adversary," *IEEE Trans. Wireless Commun.*, vol. 16, no. 2, pp. 885–899, Feb. 2017.

[18] X. Tang, P. Ren, Y. Wang, and Z. Han, “Combating full-duplex active eavesdropper: A hierarchical game perspective,” *IEEE Trans. Commun.*, vol. 65, no. 3, pp. 1379–1395, Mar. 2017.

[19] B. Akgun, O. O. Koyluoglu, and M. Krunz, “Exploiting full-duplex receivers for achieving secret communications in multiuser MISO networks,” *IEEE Trans. Commun.*, vol. 65, no. 2, pp. 956–968, Feb. 2017.

[20] G. Chen, J. P. Coon, and M. D. Renzo, “Secrecy outage analysis for downlink transmissions in the presence of randomly located eavesdroppers,” 2017. [Online]. Available: https://arxiv.org/pdf/1701.00982.pdf

[21] T.-X. Zheng, H.-M. Wang, J. Yuan, Z. Han, and M. Lee, “Physical layer security in wireless ad hoc networks under a hybrid full-/half-duplex receiver deployment strategy,” *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 3827–3829, Jun. 2017.

[22] M. Abedi, N. Mokari, and H. Saeedi, “How to manage resources to provide physical layer security: Active versus passive adversary,” *Phys. Commun.*, vol. 27, pp. 143–149, Apr. 2018.

[23] Y. Sun, D. Ng, J. Zhu, and R. Schober, “Robust and secure resource allocation for full-duplex MISO multicarrier NOMA systems,” *IEEE Trans. Commun*, doi: 10.1109/TCOMM.2018.2830325.

[24] N. Ferguson, B. Schneier, and T. Kohno, *Cryptography Engineering*. Hoboken, NJ, USA: Wiley, 2010.

[25] Y. Hua, Q. Zhu, and R. Sohrabi, “Fundamental properties of full-duplex radio for secure wireless communications,” 2017. [Online]. Available: http://arxiv.org/abs/1711.10001

[26] T. T. Cai and T. Jiang, “Phase transition in limiting distributions of coherence of high-dimensional random matrices,” *J. Multivariate Anal.*, vol. 107, pp. 24–39, 2012.

[27] N. M. Temme, “The uniform asymptotic expansion of a class of integrals related to cumulative distribution functions,” *SIAM J. Math. Anal.*, vol. 13, no. 2, pp. 239–253, Mar. 1982.

[28] 2012. [Online]. Available: https://math.stackexchange.com/questions/200140/how-to-prove-asymptotic-limit-of-an-incomplete-gamma-function

[29] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[30] Y. Yu, Y. Huang, B. Zhao, and Y. Hua, “Further development of synchronous array method for ad hoc wireless networks,” *EURASIP J. Adv. Signal Process., Special Issue on Cross-Layer Design for the Physical, MAC, and Link Layer in Wireless Systems*, vol. 2009, Sep. 2008, Art. no. 873202.

[31] B. Zhao and Y. Hua, “A distributed medium access control scheme for a large network of wireless routers,” *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1614–1622, May 2008.

[32] X. Tang and Y. Hua, “Capacity of ultra-wideband power-constrained ad hoc networks,” *IEEE Trans. Inf. Theory*, vol. 54, no. 2, pp. 916–920, Feb. 2008.

[33] K. Hong, Y. Hua, and A. Swami, “Distributed and cooperative link scheduling for large-scale multi-hop wireless networks,” *EURASIP J. Wireless Commun. Netw.*, vol. 2007, Art. no. 034716.

[34] C. Audet, X. Fournier, P. Hansen, and F. Messine, “A note on diameters of point sets,” *Optimization Lett.*, vol. 4, no. 4, pp. 585–595, Nov. 2010.

[35] Y. Hua, “Fast maximum likelihood for blind identification of multiple FIR channels,” *IEEE Trans. Signal Process.*, vol. 44, no. 3, pp. 661–672, Mar. 1996.

[36] Y. Hua, and M. Wax, “Strict identifiability of multiple FIR channels driven by an unknown arbitrary sequence,” *IEEE Trans. Signal Process.*, vol. 44, no. 3, pp. 756–758, Mar. 1996.

[37] D. J. Ryan, I. V. L. Clarkson, and I. B. Collings, “Blind detection of PAM and QAM in fading channels,” *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 1197–1206, Mar. 2006.

[38] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY, USA: Wiley, 1991.

[39] R. Wilson, D. Tse, and R. A. Scholtz, “Channel identification: Secret sharing using reciprocity in ultrawildeband channels,” *IEEE Trans. Inf. Forensics Security*, vol. 2, no. 3, Sep. 2007.

[40] M. Medard, “The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel,” *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 933–946, May 2000.

[41] T.-H. Chang, W.-C. Chiang, Y.-W. Hong, and C.-Y. Chi, “Training sequence design for discriminatory channel estimation in wireless MIMO systems,” *IEEE Trans. Signal Process.*, vol. 58, no. 12, pp. 6223–6237, Dec. 2010.

[42] H.-M. Wang, T. Zheng, and X.-G. Xia, “Secure MISO wiretap channels with multiantenna passive eavesdropper: Artificial noise vs. artificial fast fading,” *IEEE Trans. Wireless Commun.*, vol. 14, no. 1, pp. 94–106, Jan. 2015.

[43] D. E. Simmons, N. Bhargav, J. P. Coon, and S. L. Cotton, “Physical layer security over OFDM-based links: Conjugate-and-return,” in *Proc. IEEE 81st Veh. Technol. Conf.*, Glasgow, U.K., doi: 10.1109/VTC-Spring.2015.7146015.

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

[45] 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.

Yingbo Hua (S’86–M’88–SM’92–F’02) received the B.S. degree in February 1982 from Southeast University, Nanjing, China, and the Ph.D. degree in 1988 from Syracuse University, Syracuse, NY, USA. He was on the faculty with the University of Melbourne, Australia, during 1990–2000. Following a sabbatical leave with Hong Kong University of Science and Technology in 1999–2000, and a consulting experience with Microsoft Research, WA, in summer 2000, he joined the University of California, Riverside, in 2001, where he was promoted to a Senior Full Professor in 2009. He has authored or coauthored more than 330 articles in the fields of signal processing, wireless communications and sensor networks, including such topics as high-resolution methods, sensor array processing, blind source separation, blind system identification, reduced rank estimation, principal component analysis, subspace tracking, MIMO relay beamforming, MIMO channel estimation, multi-hop networks, resource allocation, full-duplex radio, and wireless network security. Since 1994, he has served in various capacities on Editorial Boards such as the IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE SIGNAL PROCESSING LETTERS, *EURASIP Signal Processing*, IEEE SIGNAL PROCESSING MAGAZINE, IEEE WIRELESS COMMUNICATIONS LETTERS, and IEEE TRANSACTIONS ON SIGNAL AND INFORMATION PROCESSING OVER NETWORKS. He also served as a Guest Editor for a number of journals including IEEE JOURNAL OF SELECTED AREAS IN COMMUNICATIONS. He was a General Chair of IEEE ChinaSIP2015 and the Lead Chair of IEEE GlobalSIP2018 Symposium on Signal Processing for Wireless Network Security. He is a former member of several IEEE SPS Technical Committees. He is a Fellow of AAAS.