Total Secrecy From Anti-Eavesdropping Channel Estimation

Shuo Wu, *Member, IEEE*, and Yingbo Hua, *Fellow, IEEE*

***Abstract*—Anti-eavesdropping channel estimation (ANECE) is useful for a network of cooperative full-duplex radio devices/users. Using ANECE, a secret key can be generated by each pair of users, and additional secret information can be transmitted between a pair of users. This paper analyzes the capacity of the secret key based on ANECE, and compares it with the conventional method for channel training. The paper also analyzes the secrecy capacity of information transmission using an one-way scheme, and compares it with a two-way scheme. It is shown that the total amount of secrecy generated from ANECE can be substantially larger than that based on the conventional method especially when an eavesdropper may have an unlimited number of antennas. The paper also formulates a total secure degree of freedom (TSDoF) of the ANECE based scheme, and compares it with a prior scheme of secret information transmission from a multi-antenna node to another against a multi-antenna eavesdropper where channel state information is unknown everywhere initially. The comparison shows that there is a substantial gain of TSDoF by exploiting full-duplex radios and reciprocal channels via ANECE. Most of the key insights are highlighted in twelve proven properties.**

***Index Terms*—Wireless networks, physical layer security, anti-eavesdropping, secret key generation, secret information transmission, total secure degree of freedom.**

I. Introduction

ENHANCEMENT of wireless network security is important for future applications in Internet-of-Things (IoT) and various battlefield networks. Wireless physical layer security, as the first line of defense against eavesdropping, aims to keep the information transmitted in open air between legitimate users safe from eavesdropper (Eve) even if the users do not have a pre-existing secret key for digital encryption.

There are two groups of methods for wireless physical layer security [1]. One is commonly called secret key generation, and the other can be referred to as secret information transmission. A method for secret key generation is a protocol for a pair of legitimate users to generate a common secret key from their observed signals that are correlated with each other, and this key can be applied later for digital encryption. A method for secret information transmission is a physical layer scheme that allows a user to directly transmit a secret to another without the necessity of a secret key already shared between them.

The secret key generated between two users can keep the information later transmitted between them secret from Eve who may have any number of antennas and/or any level of channel gain. This is because the secrecy of a transmission encrypted by a secret key is no less than the entropy of the key regardless of the number of antennas on Eve, e.g., see [2].

On the other hand, the secrecy of the traditional methods for secret information transmission diminishes to zero if the number of antennas on eavesdropper is sufficiently large [3]. A fundamental reason for this pessimistic phenomenon is because Eve is not prevented from knowing her receive channel state information (CSI) with respect to the transmitter (Alice). It is also known that if two legitimate users know their CSI but not Eve's CSI while Eve knows her CSI as well as the users' CSI, the secure degree of freedom (SDoF) of the system is zero if the number of antennas on Eve is larger than or equal to the smaller of the numbers of antennas on the users [4].

To prevent Eve from knowing her receive CSI with respect to Alice, some of the early ideas are to avoid any transmission of pilots from Alice, e.g., see [5] and [6]. But this strategy does not work well for wireless communications at high carrier frequencies (such as in gigahertz) for which pilots are essential for the intended receiver (Bob) to be able to perform carrier synchronization and subsequently detect phase shift keying (PSK) and/or quadrature amplitude modulated (QAM) symbols.[1]

More recently, a new strategy called anti-eavesdropping channel estimation (ANECE) was proposed in [7], which allows two or more cooperative full-duplex radio devices/users to estimate consistently their own receive CSI with respect to each other but at the same time prevents Eve from having a consistent estimate of her receive CSI. Full-duplex radio is an emerging technology, which allows a radio to transmit and receive at the same time on the same carrier. In this paper, we assume that the residual self-interference of full-duplex radio is relatively small (subject to a range of communication) and can be lumped into the channel noise term.

Some potential benefits from ANECE were studied in [7] and [8]. In [7], the capacity of Eve with any number of antennas to receive information from Alice, subject to a limited window per channel coherence period, was considered. That

Manuscript received May 1, 2021; revised October 13, 2021 and December 12, 2021; accepted January 17, 2022. Date of publication January 25, 2022; date of current version March 7, 2022. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Youngchul Sung. This work was supported in part by the Army Research Office under Grant W911NF-17-1-0581, and in part by the Department of Defense under Grant W911NF-20-2-0267. *(Corresponding author: Yingbo Hua.)*

The authors are with the Department of Electrical and Computer Engineering, University of California, Riverside, CA 92521 USA (e-mail: swu046@ucr.edu; yhua@ee.ucr.edu).

Digital Object Identifier 10.1109/TSP.2022.3145676

[1] If frequency shift keying (FSK) is used by Alice, then all receivers including Eve within range could rely on phase locked loop to detect all information symbols without pilot. So, FSK is not suitable in the context of this paper.

capacity was shown to be zero, approximately zero or at most bounded as the transmit power from Alice increases, under an assumption. That assumption is that Eve's receive channel matrix with respect to Alice is only known to Eve as a matrix of independent and identically distributed (i.i.d.) random elements with their absolute means much smaller than their deviations. That assumption does not take into account the fact that ANECE does not hide all information of Eve's receive CSI from Eve. In [8], the SDoF of information transmission between two multi-antenna full-duplex users applying a two-user ANECE, against Eve with unlimited number of antennas, was studied. However, the study shown in [8] does not address a network of more than two users.

In this paper, we present further contributions on the secrecy achievable from ANECE. We will first examine the capacity C_{key} of the secret key that can be generated by each pair of users in a multi-user network where ANECE is conducted. We will also examine the capacity C_{trans} of secret information transmission between a pair of users using ANECE-assisted channel estimates. Assuming that the channel matrices between users are independent of Eve's receive channel matrices, the total secrecy for a pair of users achievable from ANECE is the sum of the two capacities. To measure the total secrecy from ANECE, we will introduce a total secure degree of freedom (TSDoF) based on both C_{key} and C_{trans}.

To analyze C_{key}, we will consider two variations of ANECE for a network of more than two users: one is "pair-wise ANECE" and the other is "all-user ANECE". We will also consider C_{key} based on the conventional channel training. A major finding is that the SDoF of C_{key} with respect to the channel training energy is the same for all these training methods. The analysis of C_{key} builds on the prior work in [9] where optimal ANECE pilots were derived.

While C_{key} is invariant to, or unconditional on, the number N_E of antennas on Eve, C_{trans} is generally affected by N_E. Indeed, one might find such prior works as [3] and [4] suggesting that C_{trans} reduces to zero as N_E increases. However, we will show that, using ANECE (for a network of full-duplex radios with reciprocal channels), C_{trans} can be a positive value unconditional on N_E.

Most of the key results in this paper are highlighted in Properties 1-12, which should be easy to locate (and hopefully also easy to appreciate) by readers. For example, for a reciprocal channel between two full-duplex users each with N antennas in the presence of Eve with N_E antennas where CSI is unknown everywhere initially, the TSDoF of this system via ANECE (in bits per channel use per doubling of power) is shown to be $d_{new} = \frac{N^2+\eta}{N+K_2}$ with $\eta = (2K_2N - N_E(K_2 - \min(K_2, N)))^+$ with $K_2 \geq 0$ and $N + K_2$ being the number of channel uses per coherence period. This follows from (130). Furthermore, Property 12 shows that the TSDoF of ANECE is in general significantly larger than that shown in [10]. The latter assumes a conventional MIMO channel between two multi-antenna users where CSI is also assumed to be unknown everywhere initially. Property 12 quantifies a TSDoF advantage of utilizing full-duplex radios and reciprocal channels via ANECE.

The rest of the paper is organized as follows. In section II, we review the principle of ANECE, and also provide some key facts of ANECE for users as well as Eve. In section III, we analyze and compare the secret key capacities between each pair of users following a pair-wise ANECE, an all-user ANECE and the conventional method for channel estimation. In section IV, we analyze and compare the secrecy capacities of one-way and two-way information transmission between a pair of users following all-user ANECE. To make the analyses tractable and insightful, we will assume a symmetric network of users where each user has the same number of antennas and all elements of their channel matrices are i.i.d.. Furthermore, we focus on results under a large energy for channel training and a large power for secret information transmission. A Monte Carlo simulation is shown in section V. In section VI, we compare our results with [10] and [11]. The final conclusion is given in section VII.

Notations: The bold lower and upper cases represent vectors and matrices respectively. The $M \times N$ complex matrix space, the $N \times N$ identity matrix and the $N \times N$ zero matrix are denoted by $\mathbb{C}^{M\times N}$, $\mathbf{I}_N$ and $\mathbf{0}_N$ respectively. The transpose, conjugate, and conjugated transpose are ^{T}, *, and ^{H} respectively. The trace, determinant, Kronecker product, expectation, expectation over x only, diagonal matrix, base-2 logarithm and natural logarithm are Tr, $|\cdot|$, $\otimes$, $\mathcal{E}$, $\mathcal{E}_x$, $diag\{\cdot\}$, $\log_2$ and $\ln$, respectively. The functions of mutual information, conditional mutual information, differential entropy and conditional differential entropies are denoted by such forms as $I(\cdot;\cdot)$, $I(\cdot;\cdot|\cdot)$, $h(\cdot)$ and $h(\cdot|\cdot)$ respectively. Finally, $x^+ = \max(x, 0)$ and $\lfloor x \rfloor$ denotes the largest integer no larger than x.

II. Preliminaries

A. System Model

Assumption A: All channel gains in the system stay constant within each coherence period. All users are full-duplex with N antennas each. (Any residual self-interference is lumped into the additive noise.) The channels between users are reciprocal. All channel gains between users change as i.i.d. circular complex Gaussian random variables of zero mean and unit variance $\mathcal{CN}(0, 1)$ from one coherent period to another. All channel gains from users to Eve are independent from each other and from those between users. The channel gains from user j to Eve change as i.i.d. circular complex Gaussian random variables of zero mean and variance $\sigma^2_{E,j}$, i.e., $\mathcal{CN}(0, \sigma^2_{E,j})$, from one coherent period to another. All channel noise elements are i.i.d. $\mathcal{CN}(0, 1)$ from one sampling instant to another.

Each coherent period is utilized in two different phases. In phase 1, ANECE is conducted among $M \geq 2$ full-duplex users. In phase 2, there is a transmission of secret information between users. Illustrated in Fig. 1 is a scenario where user 1 broadcasts secret information to all other users in phase 2. This can be generalized into $M + 1$ phases, i.e., in each of the M phases after phase 1, one of the users broadcasts to all other users. The secrecy capacity of the broadcast transmission after phase 1 is analyzed in section IV-A. The secrecy capacity of another transmission scheme after phase 1 is analyzed in section IV-B.

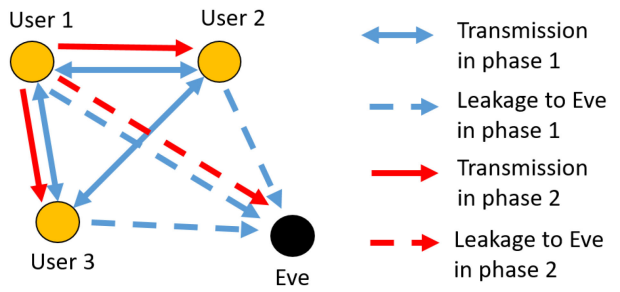


Fig. 1. Illustration of an ANECE based scheme where two or more users jointly perform ANECE in phase 1 and user 1 broadcasts secret information to other users in phase 2 against Eve with any number of antennas.

More specifically, we consider a wireless network of M full-duplex multi-antenna nodes/users where each user has N antennas. Over K_1 time slots in phase 1, all users transmit their pilots concurrently (with a synchronization precision at the symbol level rather than at the carrier level). User j transmits $\mathbf{p}_j(k) \in \mathbb{C}^{N\times 1}$ for $k = 1, \ldots, K_1$ and $j = 1, \ldots, M$. Let $\mathbf{P}_j = [\mathbf{p}_j(1), \ldots, \mathbf{p}_j(K_1)]$. Then the signal matrix received by user i in phase 1 can be written as

$$\begin{aligned}\mathbf{Y}_i &= \sum_{j\neq i}^{M} \mathbf{H}_{i,j}\mathbf{P}_j + \mathbf{N}_i \\ &= \mathbf{H}_{(i)}\mathbf{P}_{(i)} + \mathbf{N}_i \end{aligned} \tag{1}$$

where $\sum_{j\neq i}^{M}$ denotes the sum over all $j \in \{1, \ldots, M\}$ but $j \neq i$, $\mathbf{H}_{i,j}$ is the channel matrix from user j to user i, $\mathbf{H}_{(i)}$ is the horizontal stack of $\mathbf{H}_{i,j}$ for all $j \neq i$, $\mathbf{P}_{(i)}$ is the vertical stack of $\mathbf{P}_j$ for all $j \neq i$, and $\mathbf{N}_i$ is a noise matrix (including residual self-interference). For reciprocal channels, $\mathbf{H}_{i,j} = \mathbf{H}_{j,i}^T$ for all $j \neq i$.

It is clear that as long as $\mathbf{P}_{(i)}$ has a full row rank, user i is able to obtain a consistent estimate of $\mathbf{H}_{i,j}$ for all $j \neq i$. Namely, the estimate of $\mathbf{H}_{i,j}$ for any $j \neq i$ by user i converges to $\mathbf{H}_{i,j}$ as the noise variance relative to the transmit power goes to zero.

Now consider an eavesdropper (Eve) with N_E antennas. If there are multiple eavesdroppers colluding with each other at the physical layer, we can treat all of these eavesdroppers as one Eve with N_E being the sum of the numbers of their antennas. If multiple eavesdroppers are present in isolation from each other, we can consider Eve as any one of them without loss of generality. The signal matrix received by Eve in phase 1 is

$$\begin{aligned}\mathbf{Y}_E &= \sum_{i=1}^{M} \mathbf{H}_{E,i}\mathbf{P}_i + \mathbf{N}_E \\ &= \bar{\mathbf{H}}_E\bar{\mathbf{P}} + \mathbf{N}_E \end{aligned} \tag{2}$$

where $\mathbf{H}_{E,i} \in \mathbb{C}^{N_E\times N}$ is the channel matrix from user i to Eve, $\bar{\mathbf{H}}_E$ is the horizontal stack of $\mathbf{H}_{E,i}$ for all i, $\bar{\mathbf{P}}$ is the vertical stack of $\mathbf{P}_i$ for all i, and $\mathbf{N}_E$ is a noise matrix. To prevent Eve from obtaining a consistent estimate of $\bar{\mathbf{H}}_E$, we need $\bar{\mathbf{P}}$ to have a reduced row rank $r \leq N_T - 1$ with $N_T = MN$.

Any set of pilots satisfying $rank(\mathbf{P}_{(i)}) = N_T - N$ for all i and $rank(\bar{\mathbf{P}}) = r \leq N_T - 1$ is called a set of ANECE pilots. Optimal designs of the ANECE pilots are discussed in [9].

Since the network under consideration in this paper is symmetric, an optimal pilot matrix $\bar{\mathbf{P}}$ is known [9] to be

$$\bar{\mathbf{P}} = \bar{\mathbf{F}}^*\bar{\mathbf{V}}^* \tag{3}$$

where $\bar{\mathbf{V}}$ is a $(M-1)N \times K_1$ orthonormal matrix satisfying $\bar{\mathbf{V}}\bar{\mathbf{V}}^H = \mathbf{I}_r$ with $r = (M-1)N$, and

$$\bar{\mathbf{F}} = \sqrt{\frac{K_1P_1}{N^2(M-1)}}\bar{\mathbf{Q}}_m \tag{4}$$

where P_1 is the transmit power from each user, m can be any integer in $[0, M-1]$, and $\bar{\mathbf{Q}}_m$ is the $MN \times (M-1)N$ matrix obtained by removing a set of N equally spaced columns from the $MN \times MN$ discrete Fourier transform (DFT) matrix $\mathbf{Q}_{DFT}$. Specifically, if we let the $(l+1, k+1)$th element of $\mathbf{Q}_{DFT}$ be w_{MN}^{lk} with $w_{MN} = e^{-j2\pi\frac{1}{MN}}$, $0 \leq l \leq MN-1$ and $0 \leq k \leq MN-1$, and stack horizontally N equally spaced columns of $\mathbf{Q}_{DFT}$ as follows:

$$\mathbf{Q}_m = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ w_{MN}^{m} & w_{MN}^{m+M} & \cdots & w_{MN}^{m+(N-1)M} \\ \vdots & & & \vdots \\ w_{MN}^{m(MN-1)} & w_{MN}^{(m+M)(MN-1)} & \cdots & w_{MN}^{(m+(N-1)M)(MN-1)} \end{bmatrix}. \tag{5}$$

then $\bar{\mathbf{Q}}_m$ results from removing all the columns in $\mathbf{Q}_m$ (for any fixed integer $m \in [0, M-1]$) from $\mathbf{Q}_{DFT}$. We will choose $m = 0$ unless mentioned otherwise.

B. MMSE Channel Estimation in Phase 1

In phase 1, user i obtains the $N \times K_1$ signal matrix $\mathbf{Y}_i$, and Eve obtains the $N_E \times K_1$ signal matrix $\mathbf{Y}_E$. The minimum mean squared error (MMSE) channel estimations by users and Eve will be needed. It is important to understand the key properties of these estimates, which are highlighted in Properties 1, 2 and 3 below in this section.

According to Assumption A, the elements in $\mathbf{H}_{i,j}$ for all i and $j \neq i$ are i.i.d. $\mathcal{CN}(0,1)$, and all elements in $\mathbf{N}_i$ for all i are also i.i.d. $\mathcal{CN}(0,1)$. And $\mathbf{H}_{E,j}$ for all j are independent from each other and from $\mathbf{H}_{i,m}$ for all i and $m \neq i$, and the elements in $\mathbf{H}_{E,j}$ are i.i.d. $\mathcal{CN}(0, \sigma_{E,j}^2)$. And the elements in $\mathbf{N}_E$ are i.i.d. $\mathcal{CN}(0,1)$.

1) Channel Estimation by Users: Applying $vec(\mathbf{XYZ}) = (\mathbf{Z}^T \otimes \mathbf{X})vec(\mathbf{Y})$, we can write $\mathbf{y}_i = vec(\mathbf{Y}_i)$ as

$$\begin{aligned}\mathbf{y}_i &= \sum_{j\neq i}^{M} (\mathbf{P}_j^T \otimes \mathbf{I}_N)\mathbf{h}_{i,j} + \mathbf{n}_i \\ &= \mathbf{Q}_{(i)}^H\mathbf{h}_{(i)} + \mathbf{n}_i \end{aligned}$$

where $\mathbf{h}_{i,j} = vec(\mathbf{H}_{i,j})$, and $\mathbf{n}_i = vec(\mathbf{N}_i)$, $\mathbf{h}_{(i)} = vec(\mathbf{H}_{(i)})$, and $\mathbf{Q}_{(i)} = \mathbf{P}_{(i)}^* \otimes \mathbf{I}_N$.

We will also use the selection matrices $\mathbf{S}_j$ and $\mathbf{S}_{(j)}$ which are defined to be such that $\mathbf{S}_j\bar{\mathbf{P}} = \mathbf{P}_j$ and $\mathbf{S}_{(j)}\bar{\mathbf{P}} = \mathbf{P}_{(j)}$. We denote $\mathbf{K}_{\mathbf{x},\mathbf{y}} = \mathcal{E}\{\mathbf{x}\mathbf{y}^H\}$ as the correlation matrix between two

random vectors $\mathbf{x}$ and $\mathbf{y}$. Also let $\mathbf{K}_{\mathbf{x}} = \mathbf{K}_{\mathbf{x},\mathbf{x}}$. Depending on the context, $\mathbf{K}_{\mathbf{x}}$ may be also used as the covariance matrix of $\mathbf{x}$.

The MMSE estimate of $\mathbf{h}_{(i)}$ by user i is

$$\begin{aligned}\hat{\mathbf{h}}_{(i)} &= \mathbf{K}_{\mathbf{h}_{(i)},\mathbf{y}_i}\mathbf{K}_{\mathbf{y}_i}^{-1}\mathbf{y}_i \\ &= \mathbf{Q}_{(i)}(\mathbf{Q}_{(i)}^H\mathbf{Q}_{(i)} + \mathbf{I}_{NK_1})^{-1}\mathbf{y}_i. \end{aligned} \tag{6}$$

and equivalently the MMSE estimate of $\mathbf{h}_{i,j}$ by user i has the following form

$$\begin{aligned}\hat{\mathbf{h}}_{i,j} &= \mathbf{K}_{\mathbf{h}_{i,j},\mathbf{y}_i}\mathbf{K}_{\mathbf{y}_i}^{-1}\mathbf{y}_i \\ &= (\mathbf{P}_j^* \otimes \mathbf{I}_N)(\mathbf{Q}_{(i)}^H\mathbf{Q}_{(i)} + \mathbf{I}_{NK_1})^{-1}\mathbf{y}_i. \end{aligned} \tag{7}$$

Let $\Delta\mathbf{h}_{(i)} = \mathbf{h}_{(i)} - \hat{\mathbf{h}}_{(i)}$. Then the correlation matrix of $\Delta\mathbf{h}_{(i)}$ is

$$\begin{aligned}\mathbf{K}_{\Delta\mathbf{h}_{(i)}} &= \mathcal{E}\{\Delta\mathbf{h}_{(i)}\Delta\mathbf{h}_{(i)}^H\} = \mathbf{K}_{\mathbf{h}_{(i)}} - \mathbf{K}_{\mathbf{h}_{(i)},\mathbf{y}_i}\mathbf{K}_{\mathbf{y}_i}^{-1}\mathbf{K}_{\mathbf{h}_{(i)},\mathbf{y}_i}^H \\ &= \mathbf{I}_{N(N_T-N)} - \mathbf{Q}_{(i)}(\mathbf{Q}_{(i)}^H\mathbf{Q}_{(i)} + \mathbf{I}_{NK_1})^{-1}\mathbf{Q}_{(i)}^H \\ &= \left(\mathbf{Q}_{(i)}\mathbf{Q}_{(i)}^H + \mathbf{I}_{N(N_T-N)}\right)^{-1}. \end{aligned} \tag{8}$$

Using the optimal pilot matrix in (3), one can verify (e.g., see equation (79) in [9]) that

$$\begin{aligned}\mathbf{K}_{\Delta\mathbf{h}_{(i)}} &= \left(\mathbf{I}_{MN-N} + \bar{\mathbf{S}}_{(i)}\bar{\mathbf{F}}\bar{\mathbf{F}}^H\bar{\mathbf{S}}_{(i)}^T\right)^{-1} \otimes \mathbf{I}_N \\ &= \frac{\mathbf{I}_{M-1} + \frac{N\beta}{1+N\beta}\mathbf{I}_{M,i}\mathbf{q}_m\mathbf{q}_m^H\mathbf{I}_{M,i}^T}{1+MN\beta} \otimes \mathbf{I}_{N^2}, \end{aligned} \tag{9}$$

where $\beta = \frac{K_1P_1}{N^2(M-1)}$, $\mathbf{I}_{M,i}$ is $\mathbf{I}_M$ without the ith row, and $\mathbf{q}_m = [1, w_M^m, \ldots, w_M^{(M-1)m}]^T$ with $w_M = e^{-j2\pi\frac{1}{M}}$.

The correlation matrix of $\Delta\mathbf{h}_{i,j} = \mathbf{h}_{i,j} - \hat{\mathbf{h}}_{i,j}$ is an $N^2 \times N^2$ diagonal block of $\mathbf{K}_{\Delta\mathbf{h}_{(i)}}$, which is

$$\begin{aligned}\mathbf{K}_{\Delta\mathbf{h}_{i,j}} &= \mathcal{E}\left\{\Delta\mathbf{h}_{i,j}\Delta\mathbf{h}_{i,j}^H\right\} \\ &= \alpha\mathbf{I}_{N^2}, \end{aligned} \tag{10}$$

where α is a diagonal element of $\frac{\mathbf{I}_{M-1}+\frac{N\beta}{1+N\beta}\mathbf{I}_{M,i}\mathbf{q}_m\mathbf{q}_m^H\mathbf{I}_{M,i}^T}{1+MN\beta}$, i.e.,

$$\alpha = \frac{1+2N\beta}{1+N(M+1)\beta+MN^2\beta^2}. \tag{11}$$

Property 1: As $E_1 = K_1P_1 \to \infty$, the correlation matrix $\mathbf{K}_{\Delta\mathbf{h}_{i,j}}$ of the channel estimation errors at each user converges to the zero matrix.

Proof: As shown already, $\mathbf{K}_{\Delta\mathbf{h}_{i,j}} = \alpha\mathbf{I}_{N^2}$. And it follows from (11) and $\beta = \frac{K_1P_1}{N^2(M-1)}$ that $\lim_{E_1\to\infty}\alpha = 0$.

2) Channel Estimation by Eve: The signal vector $\mathbf{y}_E = vec(\mathbf{Y}_E)$ received by Eve in phase 1 is

$$\begin{aligned}\mathbf{y}_E &= \sum_{i=1}^{M}(\bar{\mathbf{P}}^T\mathbf{S}_i^T \otimes \mathbf{I}_{N_E})\mathbf{h}_{E,i} + \mathbf{n}_E \\ &= \left(\bar{\mathbf{P}}^T \otimes \mathbf{I}_{N_E}\right)\bar{\mathbf{h}}_E + \mathbf{n}_E \end{aligned} \tag{12}$$

with $\mathbf{h}_{E,i} = vec(\mathbf{H}_{E,i})$, $\bar{\mathbf{h}}_E = vec(\bar{\mathbf{H}}_E)$, and $\mathbf{n}_E = vec(\mathbf{N}_E)$.

The MMSE estimation of $\mathbf{h}_{E,i}$ by Eve is

$$\begin{aligned}\hat{\mathbf{h}}_{E,i} &= \mathbf{K}_{\mathbf{h}_{E,i},\mathbf{y}_E}\mathbf{K}_{\mathbf{y}_E}^{-1}\mathbf{y}_E \\ &= \sigma_{E,i}^2(\mathbf{S}_i\bar{\mathbf{P}}^* \otimes \mathbf{I}_{N_E})\left(\bar{\mathbf{P}}^T\mathbf{\Lambda}_E\bar{\mathbf{P}}^* \otimes \mathbf{I}_{N_E} + \mathbf{I}_{N_EK_1}\right)^{-1}\mathbf{y}_E, \end{aligned} \tag{13}$$

where $\mathbf{\Lambda}_E = diag\{\sigma_{E,1}^2\mathbf{I}_N, \ldots, \sigma_{E,M}^2\mathbf{I}_N\}$. The correlation matrix of $\hat{\mathbf{h}}_{E,i}$ is

$$\begin{aligned}\mathbf{K}_{\hat{\mathbf{h}}_{E,i}} &= \mathbf{K}_{\mathbf{h}_{E,i},\mathbf{y}_E}\mathbf{K}_{\mathbf{y}_E}^{-1}\mathbf{K}_{\mathbf{h}_{E,i},\mathbf{y}_E}^H \\ &= \sigma_{E,i}^4(\mathbf{S}_i\bar{\mathbf{P}}^* \otimes \mathbf{I}_{N_E})\left(\bar{\mathbf{P}}^T\mathbf{\Lambda}_E\bar{\mathbf{P}}^* \otimes \mathbf{I}_{N_E} + \mathbf{I}_{N_EK_1}\right)^{-1} \\ &\quad \cdot (\bar{\mathbf{P}}^T\mathbf{S}_i^T \otimes \mathbf{I}_{N_E}) \\ &= \sigma_{E,i}^4(\mathbf{S}_i\mathbf{\Lambda}_E^{-\frac{1}{2}}\mathbf{\Psi}\mathbf{\Lambda}_E^{-\frac{1}{2}}\mathbf{S}_i^T \otimes \mathbf{I}_{N_E}) \\ &= \sigma_{E,i}^2(\mathbf{S}_i\mathbf{\Psi}\mathbf{S}_i^T \otimes \mathbf{I}_{N_E}), \end{aligned}$$

with

$$\mathbf{\Psi} = \mathbf{\Lambda}_E^{\frac{1}{2}}\bar{\mathbf{P}}^*\left(\bar{\mathbf{P}}^T\mathbf{\Lambda}_E\bar{\mathbf{P}}^* + \mathbf{I}_{K_1}\right)^{-1}\bar{\mathbf{P}}^T\mathbf{\Lambda}_E^{\frac{1}{2}}. \tag{14}$$

Applying $(\mathbf{I}+\mathbf{AB})^{-1} = \mathbf{I} - \mathbf{A}(\mathbf{I}+\mathbf{BA})^{-1}\mathbf{B}$, we have

$$\mathbf{\Psi} = \mathbf{I}_{MN} - \left(\mathbf{I}_{MN} + \mathbf{\Lambda}_E^{\frac{1}{2}}\bar{\mathbf{P}}^*\bar{\mathbf{P}}^T\mathbf{\Lambda}_E^{\frac{1}{2}}\right)^{-1}. \tag{15}$$

We can further simplify the correlation matrix $\mathbf{K}_{\hat{\mathbf{h}}_{E,i}}$ with the optimal pilot matrix $\bar{\mathbf{P}}^* = \bar{\mathbf{F}}\bar{\mathbf{V}}$. First, we define $\mathbf{\Sigma}_E = diag\{\sigma_{E,1}, \ldots, \sigma_{E,M}\}$. Then, one can verify that

$$\bar{\mathbf{P}}^*\bar{\mathbf{P}}^T = \bar{\mathbf{F}}\bar{\mathbf{F}}^H = \beta\left(MN\mathbf{I}_M - N\mathbf{q}_m\mathbf{q}_m^H\right) \otimes \mathbf{I}_N \tag{16}$$

and hence

$$\begin{aligned}&\mathbf{S}_i\mathbf{\Psi}\mathbf{S}_i^T \\ &= \mathbf{I}_N - \left[\left(\mathbf{I}_M + \mathbf{\Sigma}_E\left(MN\beta\mathbf{I}_M - N\beta\mathbf{q}_m\mathbf{q}_m^H\right)\mathbf{\Sigma}_E\right)^{-1}\right]_{i,i}\mathbf{I}_N \\ &= \mathbf{I}_N - \left[\left(\mathbf{\Sigma}_{E'}^2 - N\beta\mathbf{\Sigma}_E\mathbf{q}_m\mathbf{q}_m^H\mathbf{\Sigma}_E\right)^{-1}\right]_{i,i}\mathbf{I}_N \end{aligned} \tag{17}$$

where $\mathbf{\Sigma}_{E'}^2 = \mathbf{I}_M + MN\beta\mathbf{\Sigma}_E^2$. Applying the Sherman-Morrison identity $(\mathbf{A}+\mathbf{bc}^T)^{-1} = \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{bc}^T\mathbf{A}^{-1}}{1+\mathbf{c}^T\mathbf{A}^{-1}\mathbf{b}}$, we have

$$\begin{aligned}\mathbf{S}_i\mathbf{\Psi}\mathbf{S}_i^T &= \left(1 - \left(1+MN\beta\sigma_{E,i}^2\right)^{-1}\right. \\ &\quad \left. - \frac{N\beta\left(1+MN\beta\sigma_{E,i}^2\right)^{-2}\sigma_{E,i}^2}{1-N\beta\sum_{m=1}^{M}\left[\left(1+MN\beta\sigma_{E,m}^2\right)^{-1}\sigma_{E,m}^2\right]}\right)\mathbf{I}_N \\ &= \frac{\hat{\alpha}_{E,i}}{\sigma_{E,i}^2}\mathbf{I}_N. \end{aligned} \tag{18}$$

where $\hat{\alpha}_{E,i}$ is defined obviously above. Then, from (14), we have

$$\mathbf{K}_{\hat{\mathbf{h}}_{E,i}} = \hat{\alpha}_{E,i}\mathbf{I}_{NN_E}. \tag{19}$$

Furthermore, the correlation matrix of $\Delta\mathbf{h}_{E,i} = \mathbf{h}_{E,i} - \hat{\mathbf{h}}_{E,i}$ is

$$\mathbf{K}_{\Delta\mathbf{h}_{E,i}} = \mathbf{K}_{\mathbf{h}_{E,i}} - \mathbf{K}_{\mathbf{h}_{E,i},\mathbf{y}_E}\mathbf{K}_{\mathbf{y}_E}^{-1}\mathbf{K}_{\mathbf{h}_{E,i},\mathbf{y}_E}^H$$

$$= \sigma_{E,i}^2 \mathbf{I}_{NN_E} - \mathbf{K}_{\hat{\mathbf{h}}_{E,i}}$$
$$= (\sigma_{E,i}^2 - \hat{\alpha}_{E,i}) \mathbf{I}_{NN_E}. \quad (20)$$

Property 2: If $\sigma_{E,i}^2 = \sigma_E^2$ (i.e., invariant to i), then as $E_1 = K_1 P_1 \to \infty$, the correlation matrix $\mathbf{K}_{\Delta \mathbf{h}_{E,i}}$ of the channel estimation errors at Eve converges to the constant nonzero matrix $\frac{\sigma_E^2}{M} \mathbf{I}_{NN_E}$.

Proof: As shown before, $\mathbf{K}_{\Delta \mathbf{h}_{E,i}} = (\sigma_{E,i}^2 - \hat{\alpha}_{E,i}) \mathbf{I}_{NN_E}$. Then, one can verify from the definition of $\hat{\alpha}_{E,i}$ in (18) that

$$\lim_{E_1 \to \infty} (\sigma_{E,i}^2 - \hat{\alpha}_{E,i}) = \frac{\sigma_E^2}{M}. \quad (21)$$

The cross-correlation matrix between $\Delta \mathbf{h}_{E,i}$ and $\Delta \mathbf{h}_{E,j}$ for $i \neq j$ is

$$\begin{aligned}
&\mathbf{K}_{\Delta \mathbf{h}_{E,i}, \Delta \mathbf{h}_{E,j}} \\
&= -\mathbf{K}_{\mathbf{h}_{E,i}, \mathbf{y}_E} \mathbf{K}_{\mathbf{y}_E}^{-1} \mathbf{K}_{\mathbf{h}_{E,j}, \mathbf{y}_E}^H \\
&= -\sigma_{E,i}^2 \sigma_{E,j}^2 \left(\mathbf{S}_i \bar{\mathbf{P}}^* \otimes \mathbf{I}_{N_E} \right) \left(\bar{\mathbf{P}}^T \mathbf{\Lambda}_E \bar{\mathbf{P}}^* \otimes \mathbf{I}_{N_E K_1} + \mathbf{I}_{N_E K_1} \right)^{-1} \\
&\quad \cdot \left(\bar{\mathbf{P}}^T \mathbf{S}_j^T \otimes \mathbf{I}_{N_E} \right) \\
&= -\sigma_{E,i}^2 \sigma_{E,j}^2 \left(\mathbf{S}_i \mathbf{\Lambda}_E^{-\frac{1}{2}} \mathbf{\Psi} \mathbf{\Lambda}_E^{-\frac{1}{2}} \mathbf{S}_j^T \otimes \mathbf{I}_{N_E} \right) \\
&= -\sigma_{E,i} \sigma_{E,j} \left(\mathbf{S}_i \mathbf{\Psi} \mathbf{S}_j^T \otimes \mathbf{I}_{N_E} \right). \quad (22)
\end{aligned}$$

Define

$$\alpha_{E,i,j} = \sigma_{E,i}^2 \sigma_{E,j}^2 \cdot \frac{N\beta \left(1 + MN\beta \sigma_{E,i}^2\right)^{-1} \left(1 + MN\beta \sigma_{E,j}^2\right)^{-1}}{1 - N\beta \sum_{m=1}^{M} \left[\left(1 + MN\beta \sigma_{E,m}^2\right)^{-1} \sigma_{E,m}^2 \right]}. \quad (23)$$

Using (15) and (16), we have

$$\begin{aligned}
&\mathbf{S}_i \mathbf{\Psi} \mathbf{S}_j^T \\
&= -\left[\left(\mathbf{I}_M + \mathbf{\Sigma}_E \left(MN\beta \mathbf{I}_M - N\beta \mathbf{q}_m \mathbf{q}_m^H \right) \mathbf{\Sigma}_E \right)^{-1} \right]_{i,j} \mathbf{I}_N \\
&= -\frac{\alpha_{E,i,j}}{\sigma_{E,i} \sigma_{E,j}} \mathbf{I}_N. \quad (24)
\end{aligned}$$

It then follows from (22) that $\mathbf{K}_{\Delta \mathbf{h}_{E,i}, \Delta \mathbf{h}_{E,j}} = \alpha_{E,i,j} \mathbf{I}_{NN_E}$.

Property 3: If $\sigma_{E,i}^2 = \sigma_E^2$ (i.e., invariant to i), then as $E_1 = K_1 P_1 \to \infty$, the cross-correlation matrix $\mathbf{K}_{\Delta \mathbf{h}_{E,i}, \Delta \mathbf{h}_{E,j}}$ for the channel estimation errors at Eve for $i \neq j$ also converges to the nonzero matrix $\frac{\sigma_E^2}{M} \mathbf{I}_{NN_E}$.

Proof: This property follows from $\mathbf{K}_{\Delta \mathbf{h}_{E,i}, \Delta \mathbf{h}_{E,j}} = \alpha_{E,i,j} \mathbf{I}_{NN_E}$ and (23), i.e.,

$$\lim_{E_1 \to \infty} \alpha_{E,i,j} = \frac{\sigma_E^2}{M}. \quad (25)$$

We see from Properties 1, 2 and 3 that while the channel estimation errors at all users go to zero as the training energy E_1 per user increases, the variances and some cross-correlations of channel estimation errors at Eve converge to a non-zero constant as E_1 increases. We also see that as the number of users participating in the M-user ANECE increases, the non-zero constant $\frac{\sigma_E^2}{M}$ decreases. But we will show that this does not weaken some important advantages of all-user ANECE over pair-wise ANECE.

III. Capacity of Secret Key From Estimated Channel Matrices in Phase 1

In this section, we will study the capacity of the secret key that can be generated between each pair of users after ANECE has been conducted in phase 1. In the next section, we will study the secrecy capacity of information transmission between a given pair of users in phase 2.

The channel estimates by each pair of users following ANECE can be used to generate a secret key via secret key generation protocol [12]–[16]. In this section, we analyze and compare the capacities of these secret keys in three different cases.

As stated in Assumption A, Eve's receive channel matrix is independent of the users' channel matrices and all channel matrices stay constant within each coherence period but change independently from one coherence period to another, the capacity of secret key generated by users i and j based on their respective observations $\mathbf{Y}_i$ and $\mathbf{Y}_j$ in bits per channel coherence period is known [13, Th. 4.1] to be

$$C_{key}(i,j) \doteq I(\mathbf{Y}_i; \mathbf{Y}_j) = I(\mathbf{y}_i; \mathbf{y}_j) \quad (26)$$

which is the mutual information between $\mathbf{y}_i$ and $\mathbf{y}_j$.

From [9, Lemma 1], we know that if $\mathbf{S}_j \bar{\mathbf{P}}^*$ and $\mathbf{S}_i \bar{\mathbf{P}}^*$ both have full row ranks, then $I(\mathbf{y}_i; \mathbf{y}_j) = I(\hat{\mathbf{h}}_{i,j}; \hat{\mathbf{h}}_{j,i})$ where $i \neq j$, $\hat{\mathbf{h}}_{i,j}$ is the MMSE estimate of the channel vector $\mathbf{h}_{i,j} \doteq vec(\mathbf{H}_{i,j})$ by user i, and $\hat{\mathbf{h}}_{j,i}$ is the MMSE estimate of $\mathbf{h}_{j,i} \doteq vec(\mathbf{H}_{j,i}) = vec(\mathbf{H}_{i,j}^T)$ by user j. Furthermore, it is known (see (41) in [9]) that subject to $\bar{\mathbf{P}}^* = \bar{\mathbf{F}} \bar{\mathbf{V}}$ with any $\bar{\mathbf{F}}$ but $\bar{\mathbf{V}} \bar{\mathbf{V}}^H = \mathbf{I}$,

$$I(\mathbf{y}_i; \mathbf{y}_j) = -\log_2 |\mathbf{I}_{N^2} - \mathbf{\Gamma}_{i,j} \mathbf{\Gamma}_{T,j,i}|, \quad (27)$$

where

$$\begin{aligned}
\mathbf{\Gamma}_{i,j} &= (\mathbf{S}_j \bar{\mathbf{F}} \bar{\mathbf{F}}^H \mathbf{S}_j^T \otimes \mathbf{I}_N) - \Big[(\mathbf{S}_j \bar{\mathbf{F}} \bar{\mathbf{F}}^H \bar{\mathbf{S}}_{(i)}^T) \\
&\cdot (\mathbf{I}_N + \bar{\mathbf{S}}_{(i)} \bar{\mathbf{F}} \bar{\mathbf{F}}^H \bar{\mathbf{S}}_{(i)}^T)^{-1} (\bar{\mathbf{S}}_{(i)} \bar{\mathbf{F}} \bar{\mathbf{F}}^H \mathbf{S}_j^T) \Big] \otimes \mathbf{I}_N, \quad (28) \\
\mathbf{\Gamma}_{T,j,i} &= (\mathbf{I}_N \otimes \mathbf{S}_i \bar{\mathbf{F}} \bar{\mathbf{F}}^H \mathbf{S}_i^T) - \mathbf{I}_N \otimes \Big[(\mathbf{S}_i \bar{\mathbf{F}} \bar{\mathbf{F}}^H \mathbf{S}_{(j)}^T) \\
&\cdot (\mathbf{I}_N + \mathbf{S}_{(j)} \bar{\mathbf{F}} \bar{\mathbf{F}}^H \mathbf{S}_{(j)}^T)^{-1} (\mathbf{S}_{(j)} \bar{\mathbf{F}} \bar{\mathbf{F}}^H \mathbf{S}_i^T) \Big]. \quad (29)
\end{aligned}$$

Note that $\bar{\mathbf{F}}$ in (4) is optimal for the case where ANECE is applied simultaneously to $M \geq 2$ users. We will refer to this case of ANECE as "all-user ANECE". We can also apply ANECE to each pair of users sequentially for all pairs in phase 1, which will be referred to as "pair-wise ANECE". For pair-wise ANECE, the optimal $\bar{\mathbf{F}}$ will be shown next.

A. Using Pair-Wise ANECE

For a fair comparison between all-user ANECE and pair-wise ANECE for $M > 2$, we set the total energy consumed by each user for channel estimation (in phase 1) to be E_1. It follows that $E_1 = P_1 K_1$ with P_1 being the power consumed by each user for all-user ANECE.

For pair-wise ANECE, there are $\binom{M}{2} = \frac{1}{2}M(M-1)$ distinct pairs sharing K_1 time slots in phase 1 orthogonally, and each user needs to be active sequentially for $M-1$ distinct pairs. Let P_1' be the power consumed by each user in an active pair, and $K_1' = \frac{2K_1}{M(M-1)}$ be the number of slots used by each active pair. Then we need to set $P_1'K_1'(M-1) = P_1K_1 = E_1$ or equivalently

$$P_1' = \frac{1}{2}MP_1. \tag{30}$$

Note that $\sqrt{\frac{K_1'P_1'}{N^2}} = \sqrt{\frac{K_1P_1}{N^2(M-1)}}$ which is the same as the scalar in (4). Therefore, the optimal $\bar{\mathbf{F}}$ for pair-wise ANECE is a special case of (4) and equals to

$$\bar{\mathbf{F}}' = \sqrt{\frac{K_1P_1}{N^2(M-1)}}\bar{\mathbf{Q}}_l' \tag{31}$$

with

$$\bar{\mathbf{Q}}_l' = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ w_{2N}^{l} & w_{2N}^{l+2} & \cdots & w_{2N}^{l+2(N-1)} \\ \vdots & & & \vdots \\ w_{2N}^{l(2N-1)} & w_{2N}^{(l+2)(2N-1)} & \cdots & w_{2N}^{(l+2(N-1))(2N-1)} \end{bmatrix} \tag{32}$$

where l is either 0 or 1. Note that l here relates to m (in the previous $\bar{\mathbf{Q}}_m$ with $M=2$) as $l = (m+1)_{modulo-2}$.

Given the symmetry of the network, we have $I(\mathbf{y}_i;\mathbf{y}_j) = I(\mathbf{y}_1;\mathbf{y}_2)$ for all $i \neq j$. Therefore, it follows from (27) that the capacity of the secret key generated from the estimated channel matrices by each pair of users following pair-wise ANECE is

$$C_{key}^{(pair)} \doteq -\log_2 |\mathbf{I}_{N^2} - \mathbf{\Gamma}_{1,2}\mathbf{\Gamma}_{T,2,1}|, \tag{33}$$

with

$$\mathbf{\Gamma}_{1,2} = (\mathbf{S}_2\bar{\mathbf{F}}'\bar{\mathbf{F}}'^H\mathbf{S}_2^T \otimes \mathbf{I}_N) - \left[(\mathbf{S}_2\bar{\mathbf{F}}'\bar{\mathbf{F}}'^H\mathbf{S}_2^T) \cdot (\mathbf{I}_N + \mathbf{S}_2\bar{\mathbf{F}}'\bar{\mathbf{F}}'^H\mathbf{S}_2^T)^{-1}(\mathbf{S}_2\bar{\mathbf{F}}'\bar{\mathbf{F}}'^H\mathbf{S}_2^T)\right] \otimes \mathbf{I}_N, \tag{34}$$

$$\mathbf{\Gamma}_{T,2,1} = (\mathbf{I}_N \otimes \mathbf{S}_1\bar{\mathbf{F}}'\bar{\mathbf{F}}'^H\mathbf{S}_1^T) - \mathbf{I}_N \otimes \left[(\mathbf{S}_1\bar{\mathbf{F}}'\bar{\mathbf{F}}'^H\mathbf{S}_1^T) \cdot (\mathbf{I}_N + \mathbf{S}_1\bar{\mathbf{F}}'\bar{\mathbf{F}}'^H\mathbf{S}_1^T)^{-1}(\mathbf{S}_1\bar{\mathbf{F}}'\bar{\mathbf{F}}'^H\mathbf{S}_1^T)\right] \tag{35}$$

where $\mathbf{S}_1 = [\mathbf{I}_N, \mathbf{0}_N]$ and $\mathbf{S}_2 = [\mathbf{0}_N, \mathbf{I}_N]$.

Furthermore, one can verify that

$$\mathbf{S}_1\bar{\mathbf{F}}'(\bar{\mathbf{F}}')^H\mathbf{S}_1^T = k\mathbf{I}_N, \tag{36}$$

$$\mathbf{S}_2\bar{\mathbf{F}}'(\bar{\mathbf{F}}')^H\mathbf{S}_2^T = k\mathbf{I}_N, \tag{37}$$

with

$$k = \frac{E_1}{N(M-1)}. \tag{38}$$

And hence $\mathbf{\Gamma}_{1,2} = \frac{k}{1+k}\mathbf{I}_{N^2}$ and $\mathbf{\Gamma}_{T,2,1} = \frac{k}{1+k}\mathbf{I}_{N^2}$. Therefore, the following property is ready to be verified.

Property 4: The capacity (in bits per channel coherence period) of the secret key generated by each pair of users following pair-wise ANECE is

$$C_{key}^{(pair)} = N^2\log_2\frac{(1+k)^2}{2k+1}. \tag{39}$$

with k defined in (38). For large M or small E_1,

$$C_{key}^{(pair)} \approx (\log_2 e)\frac{E_1^2}{(M-1)^2}. \tag{40}$$

For large E_1,

$$C_{key}^{(pair)} \approx N^2\log_2\frac{E_1}{2(M-1)N}. \tag{41}$$

Furthermore,

$$d_{key}^{(pair)} \doteq \lim_{E_1\to\infty}\frac{C_{key}^{(pair)}}{\log_2 E_1} = N^2. \tag{42}$$

Proof: (39) follows from the previous discussion, i.e., the simplification of (33). (40) follows from the second order approximation of (39) for small $\frac{E_1}{M-1}$. The others are also easy to verify.

Note that $d_{key}^{(pair)}$ can be called the degree of freedom in $C_{key}^{(pair)}$ with respect to $\log_2 E_1$. In other words, $d_{key}^{(pair)}$ is the increased number of bits of secret key for every doubling of E_1 when E_1 is large.

B. Using All-User ANECE

We now consider the capacity of secret key generated by each pair of users following all-user ANECE applied to the symmetric network, which is given by $C_{key}^{(all)} \doteq I(\mathbf{y}_j;\mathbf{y}_i) = I(\mathbf{y}_1;\mathbf{y}_2)$ in (27) along with (28), (29) and (4).

It is known (see the discussion below equation (100) in [9] where $N\alpha_d$ is equivalent to k below) that

$$(\mathbf{I}_{N^2} - \mathbf{\Gamma}_{i,j}\mathbf{\Gamma}_{T,j,i}) = \left(1 - \frac{\left(Mk - \frac{k}{1+k}\right)^2}{(1+Mk)^2}\right)\mathbf{I}_{N^2}. \tag{43}$$

Using this result in (27), the following property can be shown.

Property 5: The capacity (in bits per channel coherence period) of secret key generated by each pair of users following all-user ANECE is

$$C_{key}^{(all)} = -\log_2|\mathbf{I}_{N^2} - \mathbf{\Gamma}_{i,j}\mathbf{\Gamma}_{T,j,i}| = N^2\log_2\left(\frac{k_1}{k_2}\right), \tag{44}$$

with

$$k_1 = (1+Mk)^2, \tag{45}$$

$$k_2 = 1 + 2Mk + \frac{2Mk^2}{1+k} - \frac{k^2}{(1+k)^2}, \tag{46}$$

and $k = \frac{E_1}{N(M-1)}$. Furthermore,

$$\lim_{M\to\infty} C_{key}^{(all)} = N^2\log_2\frac{\left(1+\frac{E_1}{N}\right)^2}{1+2\frac{E_1}{N}}. \tag{47}$$

For small E_1,

$$C_{key}^{(all)} \approx (\log_2 e)E_1^2. \tag{48}$$

For large E_1,

$$C_{key}^{(all)} \approx N^2\log_2\left(\frac{ME_1}{4(M-1)N}\right). \tag{49}$$

Finally,

$$d_{key}^{(all)} \doteq \lim_{E_1\to\infty} \frac{C_{key}^{(all)}}{\log_2 E_1} = N^2. \tag{50}$$

Proof: (44) follows from the previous discussion. (47) follows by using $k \to 0$ and $Mk \to \frac{E_1}{N}$ as $M \to \infty$. In fact, the first-order approximation of $C_{key}^{(all)}$ in terms of $1/M$ as $M \to \infty$ can be shown to be

$$C_{key}^{(all)} \approx N^2 \log_2 \left[a \left(1 - \frac{b}{M} \right) \right] \tag{51}$$

with $a = \frac{(1+c)^2}{1+2c}$, $b = \frac{2c(1+c^2)}{(1+c)(1+2c)}$ and $c = E_1/N$. (48) follows from the second order approximation for small E_1. (49) follows by using $k_1 \to (Mk)^2$ and $k_2 \to 4Mk$ as $E_1 \to \infty$. The rest is obvious.

We see from Properties 4 and 5 that the degree of freedom of secret key generated by each pair of users following either pair-wise ANECE or all-user ANECE is the same N^2. Furthermore, we have the following.

Property 6: For $M > 2$, the gap between $C_{key}^{(all)}$ and $C_{key}^{(pair)}$ is

$$\Delta C_{key}^{(all)} \doteq C_{key}^{(all)} - C_{key}^{(pair)} = N^2 \log_2 \left(1 + \frac{1}{\nu - 1} \right) > 0 \tag{52}$$

with

$$\nu = \frac{(1+Mk)^2}{(M-2)Mk^2} > 1 \tag{53}$$

and $k = \frac{E_1}{N(M-1)}$. As M or E_1 increases, ν decreases and hence $\Delta C_{key}^{(all)}$ increases. Furthermore,

$$\lim_{M\to\infty} \nu = \left(\frac{N}{E_1} + 1 \right)^2 > 1, \tag{54}$$

$$\lim_{E_1\to\infty} \nu = \frac{M}{M-2} > 1. \tag{55}$$

Proof: (52) follows from Propositions 4 and 5 with straightforward but slightly tedious steps. We can also write

$$\nu = \frac{(M-1)^2}{(M-2)M} \frac{\left(1 + \frac{M}{M-1}\frac{E_1}{N}\right)^2}{\left(\frac{E_1}{N}\right)^2} \tag{56}$$

where we see that both $\frac{M}{M-1}$ and $\frac{(M-1)^2}{(M-2)M}$ are decreasing functions of M. So, ν decreases as M increases. We also see from (56) that ν decreases as $\frac{E_1}{N}$ increases. The limit of ν as M or E_1 increases is also easy to verify. The inequalities in (52) and (53) follow from the decreasing nature of ν and the inequalities in (54) and (55).

The above property suggests that using all-user ANECE is advantageous over using pair-wise ANECE in terms of secret key capacity (i.e., $C_{key}^{(all)}$ versus $C_{key}^{(pair)}$) subject to the same total transmit energy E_1 per user for each coherence period.

C. Using Conventional Method

Now, we consider the conventional method for channel training. In this case, each user sequentially broadcasts a pilot matrix $\mathbf{P}$ with orthonormal rows and the total energy E_1, i.e., $\mathbf{P}\mathbf{P}^H = \frac{E_1}{N}\mathbf{I}_N$. The total number K_1 of time slots needed for all users must now satisfy $K_1 \geq MN$.

Although the conventional method does not allow multiple users transmit their pilots concurrently, we still can use the same formula (27) to compute the capacity of secret key generated by each pair of users. Specifically, to obtain $I(\mathbf{y}_1;\mathbf{y}_2)$ based on the conventional channel estimation, we can choose $\bar{\mathbf{P}} = \bar{\mathbf{F}} = \sqrt{\frac{E_1}{N}}[\mathbf{I}_N, \mathbf{I}_N]^T, \mathbf{S}_1 = [\mathbf{I}_N, \mathbf{0}_N], \mathbf{S}_2 = [\mathbf{0}_N, \mathbf{I}_N], \bar{\mathbf{S}}_{(1)} = \mathbf{S}_2$ and $\bar{\mathbf{S}}_{(2)} = \mathbf{S}_1$ for $I(\mathbf{y}_1;\mathbf{y}_2)$ in (27). Therefore, the capacity of secret key generated by each pair of users following the conventional channel training is

$$C_{key}^{(conv)} \doteq -\log_2 |\mathbf{I}_{N^2} - \mathbf{\Gamma}_{1,2}\mathbf{\Gamma}_{T,2,1}|, \tag{57}$$

$$\begin{aligned}\mathbf{\Gamma}_{1,2} &= \frac{E_1}{N}\mathbf{I}_{N^2} \\ &\quad - \left(\frac{E_1}{N}\mathbf{I}_N \left(\left(1 + \frac{E_1}{N} \right) \mathbf{I}_N \right)^{-1} \frac{E_1}{N}\mathbf{I}_N \right) \otimes \mathbf{I}_N \\ &= \frac{E_1}{N+E_1}\mathbf{I}_{N^2}, \end{aligned} \tag{58}$$

$$\begin{aligned}\mathbf{\Gamma}_{T,2,1} &= \frac{E_1}{N}\mathbf{I}_{N^2} \\ &\quad - \mathbf{I}_N \otimes \left(\frac{E_1}{N}\mathbf{I}_N \left(\left(1 + \frac{E_1}{N} \right) \mathbf{I}_N \right)^{-1} \frac{E_1}{N}\mathbf{I}_N \right) \\ &= \frac{E_1}{N+E_1}\mathbf{I}_{N^2}. \end{aligned} \tag{59}$$

Therefore, the following property is ready to be verified.

Property 7: The capacity of secret key generated by each pair of users following the conventional channel training is

$$\begin{aligned} C_{key}^{(conv)} &= -\log_2 \left| \left(1 - \frac{E_1^2}{(N+E_1)^2} \right) \mathbf{I}_{N^2} \right| \\ &= N^2 \log_2 \frac{(1+\frac{E_1}{N})^2}{(1+2\frac{E_1}{N})}. \end{aligned} \tag{60}$$

For small E_1,

$$C_{key}^{(conv)} \approx (\log_2 e) E_1^2. \tag{61}$$

For large E_1,

$$C_{key}^{(conv)} \approx N^2 \log_2 \frac{E_1}{2N}. \tag{62}$$

Finally,

$$d_{key}^{(conv)} \doteq \lim_{E_1\to\infty} \frac{C_{key}^{(conv)}}{\log_2 E_1} = N^2. \tag{63}$$

Proof: (60) follows from the previous discussion. (61) follows from the second-order approximation for small E_1. The rest is obvious.

We see that $d_{key}^{(conv)} = d_{key}^{(all)} = d_{key}^{(pair)} = N^2$. Furthermore, the following is easy to verify.

Property 8: Let $\Delta C_{key}^{(conv)} \doteq C_{key}^{(conv)} - C_{key}^{(all)}$. Then,

$$\lim_{M\to\infty} \Delta C_{key}^{(conv)} = 0, \tag{64}$$

$$\lim_{E_1\to\infty} \Delta C_{key}^{(conv)} = N^2 \log_2 \left(2\frac{M-1}{M}\right). \tag{65}$$

And for small E_1,

$$\Delta C_{key}^{(conv)} \le \mathcal{O}(E_1^3). \tag{66}$$

Proof: (64) follows from (47) and (60). More generally, it follows from (51) and (60) that for large M,

$$\Delta C_{key}^{(all)} = N^2(\log_2 e)\frac{b}{M} \tag{67}$$

with b defined below (51). (65) follows from (49) and (62). (66) follows from (48) and (61).

The conventional method does not have the rank constraint on the overall pilot matrix $\bar{\mathbf{P}}$. It is expected that $\Delta C_{key}^{(conv)} > 0$ for $M > 2$. But it is a surprising result that as M increases subject to any fixed E_1, the gap $\Delta C_{key}^{(conv)}$ approaches zero. This property could be useful (for example) for a swarm of drones which apply all-user ANECE for channel estimation.

We also see here that with large E_1, the gap $\Delta C_{key}^{(conv)}$ for any $M > 2$ users is no larger than N^2. For the case of single-antenna users, the gap $\Delta C_{key}^{(conv)}$ at high power is no larger than 1.

It is important to note here that by the conventional method for channel training, Eve is not made blind to her receive channel matrix and hence any subsequent transmission of information between users always has a zero secrecy when the number of antennas on Eve is sufficiently large [3].

Before we discuss the unconditional secrecy of ANECE assisted subsequent transmissions between users, shown below is another useful property.

Property 9: The minimum number $K_{1,min}$ of time slots (or sampling instants) required for channel training for all users in phase 1 is as follows. For pair-wise ANECE,

$$K_{1,min}^{(pair)} = \frac{1}{2}M(M-1)N. \tag{68}$$

For all-user ANECE,

$$K_{1,min}^{(all)} = (M-1)N. \tag{69}$$

For the conventional method,

$$K_{1,min}^{(conv)} = MN. \tag{70}$$

Proof: For pair-wise ANECE, each pair needs at least N time slots, and there are $\frac{1}{2}M(M-1)$ pairs. For all-user ANECE, the pilot matrices (each of $(M-1)N$ or more columns) transmitted by all users concurrently must occupy $(M-1)N$ or more time slots. For the conventional method, each transmit user requires at least N time slots.

It is interesting to see that $K_{1,min}^{(all)} < K_{1,min}^{(conv)} < K_{1,min}^{(pair)}$ for $M > 2$, i.e., using all-user ANECE causes the least amount of delay for consistent channel estimation at users.

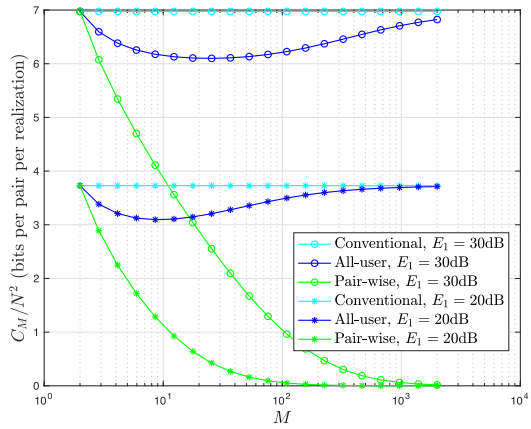


Fig. 2. $\frac{C_{key}}{N^2}$ versus M.

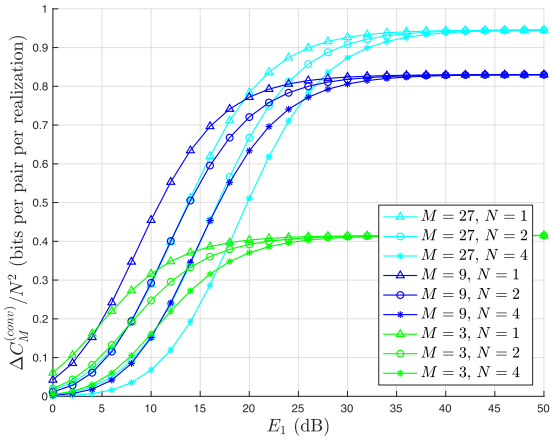


Fig. 3. $\frac{\Delta C_{key}^{(conv)}}{N^2}$ versus E_1.

D. Numerical Illustration

Fig. 2 shows $\frac{C_{key}}{N^2}$ versus M where $\sigma^2 = 1$, $N = 4$ and $E_1 = 20, 30$ dB. We see that $C_{key}^{(conv)} > C_{key}^{(all)} > C_{key}^{(pair)}$ as expected from the previous discussions. It is interesting to observe that $C_{key}^{(all)}$ decreases initially and then increases later as M increases. Consistent with (64), $C_{key}^{(all)}$ approaches $C_{key}^{(conv)}$ as M becomes large.

Fig. 3 shows $\frac{\Delta C_{key}^{(conv)}}{N^2}$ versus E_1 in dB (i.e., $10\log_{10} E_1$) where $\sigma^2 = 1$, $N = 1, 2, 4$ and $M = 3, 9, 27$. The values of $\frac{\Delta C_{key}^{(conv)}}{N^2}$ at large E_1 in this figure are consistent with (65).

IV. Secrecy Capacity of Information Transmission in Phase 2 After ANECE

In this section, we consider the secrecy capacity of information transmission between users in phase 2 following ANECE in phase 1. If a conventional method is used for channel training in phase 1, Eve is able to obtain a consistent estimate of her receive channel matrix with respect to any user who has sent a pilot. In this case, Eve with an unlimited number of antennas is able to detect all information transmitted by those users, and hence the unconditional secrecy of information transmission following a

conventional training method is zero. However, if ANECE is applied by users for channel training, Eve is unable to obtain a consistent estimate of her receive channel matrix with respect to any of these users, and hence there is a nonzero unconditional secrecy in information transmission following ANECE.

The results from this section will provide a quantitative measure of the secrecy of information transmission following ANECE. We will consider two types of secret information transmission following ANECE. The first is the broadcast transmission as shown in Fig. 1. We will also refer to this as one-way transmission from a user to another. In this case, the secrecy capacity between a transmit user and a receive user is the same as that between the transmit user and any of the other receive users. Notice our assumption of symmetric user network. Without loss of generality, we will just focus on one pair of users, which is done in section IV-A. However, if the M users sequentially broadcast their secret information in M phases after phase 1, the obtained results in section IV-A do not apply to each pair of users in each of the M phases. This is because the M channel matrices from the M users to Eve are no longer independent from each other given the knowledge that Eve has obtained in phase 1. In section IV-B, we will study a different transmission scheme after phase 1, called two-way scheme.

To obtain useful insights, we will focus on asymptotical results under a large energy E_1 in phase 1 and a large power P_2 in phase 2. The secrecy capacity C_{trans} of interest here is also measured in bits per channel coherence period. In addition to Assumption A, we will use:

Assumption B: All transmitted information symbols in phase 2 are i.i.d. $\mathcal{CN}(0, \frac{P_2}{N})$ with P_2 being the transmit power by the transmit user in phase 2. The power P_1 used by each user in phase 1 is so large that each user has practically obtained its exact CSI.

A. One-Way Transmission After All-User ANECE

We now consider the broadcast or one-way transmission of information from one user to other users in phase 2 as illustrated in Fig. 1. we will consider a particular pair of users, i.e., user i and user j each having N antennas, where user i transmits the information-carrying signal vectors $\mathbf{x}_i(k)$ for $k = 1, \ldots, K_2$ and user j accordingly receives

$$\mathbf{Y}'_j = \mathbf{H}_{j,i}\mathbf{X}_i + \mathbf{N}_j,' \tag{71}$$

where $\mathbf{Y}'_j = [\mathbf{y}'_i(1), \ldots, \mathbf{y}'_i(K_2)] \in \mathbb{C}^{N\times K_2}$, $\mathbf{X}_i = [\mathbf{x}_i(1), \ldots, \mathbf{x}_i(K_2)] \in \mathbb{C}^{N\times K_2}$, $\mathbf{H}_{j,i}$ is the same channel matrix between users i and j in phase 1, and $\mathbf{N}'_j$ is the channel noise matrix at user j in phase 2 but has the same statistics as in phase 1. According to Assumption B, the elements in $\mathbf{X}_i$ for all coherence periods are i.i.d. $\mathcal{CN}(0, \frac{P_2}{N})$ with P_2 being the power of $\mathbf{x}_i(k)$.

Corresponding to $\mathbf{X}_i$, Eve with N_E antennas receives

$$\mathbf{Y}'_E = \mathbf{H}_{E,i}\mathbf{X}_i + \mathbf{N}_E,' \tag{72}$$

where $\mathbf{H}_{E,i} \in \mathbb{C}^{N_E\times N}$ is the same channel matrix from user i to Eve in phase 1, and $\mathbf{N}'_E$ is the channel noise at Eve in phase 2 but has the same statistics as in phase 1.

The vector-form expressions of (71) and (72) are respectively

$$\mathbf{y}'_j = (\mathbf{I}_{K_2} \otimes \mathbf{H}_{j,i})\mathbf{x}_i + \mathbf{n}_j,' \tag{73}$$

$$\mathbf{y}'_E = (\mathbf{I}_{K_2} \otimes \mathbf{H}_{E,i})\mathbf{x}_i + \mathbf{n}'_E \tag{74}$$

where $\mathbf{y}'_j = vec(\mathbf{Y}'_j)$, $\mathbf{x}_i = vec(\mathbf{X}_i)$, etc.

According to Assumption B, E_1 is so large that the estimate errors of $\mathbf{H}_{j,i}$ by both users i and j are negligible compared to the effect of the channel noise (see Property 1). But the estimate of $\mathbf{H}_{E,i}$ by Eve equals to $\hat{\mathbf{H}}_{E,i} = \mathbf{H}_{E,i} - \Delta\mathbf{H}_{E,i}$ where all entries in $\Delta\mathbf{H}_{E,i}$ can be treated as uncorrelated with each other and having the variance $\frac{\sigma_E^2}{M}$ (see Property 2).

Hence, we have

$$\mathbf{y}'_j = (\mathbf{I}_{K_2} \otimes \mathbf{H}_{j,i})\mathbf{x}_i + \mathbf{n}_j,' \tag{75}$$

$$\mathbf{y}'_E = (\mathbf{I}_{K_2} \otimes \hat{\mathbf{H}}_{E,i})\mathbf{x}_i + (\mathbf{I}_{K_2} \otimes \Delta\mathbf{H}_{E,i})\mathbf{x}_i + \mathbf{n}'_E. \tag{76}$$

Then the capacity in bits per coherence period of the secret transmitted from user i to user j can be defined as

$$C_{trans}^{(one)} = \left[I(\mathbf{y}'_j; \mathbf{x}_i|\mathbf{h}_{j,i}) - I(\mathbf{y}'_E; \mathbf{x}_i|\hat{\mathbf{h}}_{E,i})\right]^+. \tag{77}$$

Next, we will analyze the two terms in $C_{trans}^{(one)}$ in order to find a lower bound of $C_{trans}^{(one)}$.

1) Analysis of $I(\mathbf{y}'_j; \mathbf{x}_i|\mathbf{h}_{j,i})$: Since $\mathbf{x}_i$ consists of NK_2 i.i.d. $\mathcal{CN}(0, \frac{P_2}{K_2})$ entries, we first write

$$I(\mathbf{y}'_j; \mathbf{x}_i|\mathbf{h}_{j,i}) = h(\mathbf{x}_i) - h(\mathbf{x}_i|\mathbf{y}_j,' \mathbf{h}_{j,i}) \tag{78}$$

where $h(\mathbf{x}_i) = \log_2[(\pi e)^{NK_2}|\frac{P_2}{N}\mathbf{I}_{NK_2}|]$. For the second term in (78), we can use the fact $h(\mathbf{x}|\mathbf{y}) \le \log[(\pi e)^n|\mathbf{K}_{\mathbf{x}|\mathbf{y}}|]$ with $\mathbf{x}$ being a random vector in $\mathbb{C}^{n\times 1}$ and $\mathbf{K}_{\mathbf{x}|\mathbf{y}} = \mathbf{K}_{\mathbf{x}} - \mathbf{K}_{\mathbf{x},\mathbf{y}}\mathbf{K}_{\mathbf{y}}^{-1}\mathbf{K}_{\mathbf{x},\mathbf{y}}$ [17]. Therefore,

$$h(\mathbf{x}_i|\mathbf{y}_j,' \mathbf{h}_{j,i}) \le \mathcal{E}_h\left\{\log|\mathbf{K}_{\mathbf{x}_i|\mathbf{y}_j,'\mathbf{h}_{j,i}}|\right\} + NK_2\log(\pi e) \tag{79}$$

where the expectation $\mathcal{E}_h$ is over the distribution of $\mathbf{h}_{j,i}$. With given $\mathbf{h}_{j,i}$, it follows from (75) that

$$\mathbf{K}_{\mathbf{x}_i|\mathbf{y}_j,'\mathbf{h}_{j,i}} = \frac{P_2}{N}\mathbf{I}_{NK_2} - \frac{P_2^2}{N^2}(\mathbf{I}_{K_2} \otimes \mathbf{H}_{j,i}^H) \cdot \left(\frac{P_2}{N}(\mathbf{I}_{K_2} \otimes \mathbf{H}_{j,i}\mathbf{H}_{j,i}^H) + \mathbf{I}_{NK_2}\right)^{-1}(\mathbf{I}_{K_2} \otimes \mathbf{H}_{j,i}). \tag{80}$$

Applying $|\mathbf{I} + \mathbf{AB}| = |\mathbf{I} + \mathbf{BA}|$ and hence $|a\mathbf{I} - a^2(\mathbf{I} \otimes \mathbf{A}^H)(a(\mathbf{I} \otimes \mathbf{AA}^H) + \mathbf{I})^{-1}(\mathbf{I} \otimes \mathbf{A})| = |a\mathbf{I} - a^2(a(\mathbf{I} \otimes \mathbf{AA}^H) + \mathbf{I})^{-1}(\mathbf{I} \otimes \mathbf{AA}^H)| = |(a(\mathbf{I} \otimes \mathbf{AA}^H) + \mathbf{I})^{-1}| \cdot |a\mathbf{I}|$, one can verify that

$$\log|\mathbf{K}_{\mathbf{x}_i|\mathbf{y}_j,'\mathbf{h}_{j,i}}| = NK_2\log_2\frac{P_2}{N} - K_2\log_2\left|\mathbf{I}_N + \frac{P_2}{N}\mathbf{H}_{j,i}\mathbf{H}_{j,i}^H\right|. \tag{81}$$

Combining (78), (79) and (81), we have

$$I(\mathbf{y}'_j; \mathbf{x}_i|\mathbf{h}_{j,i}) \ge K_2\mathcal{E}_h\left\{\log_2\left|\mathbf{I}_N + \frac{P_2}{N}\mathbf{H}_{j,i}\mathbf{H}_{j,i}^H\right|\right\}. \tag{82}$$

According to [18], if $\mathbf{A} \in \mathbb{C}^{m\times n}$ consists of i.i.d. $\mathcal{CN}(0,1)$ entries and c is a constant, then $\mathcal{E}\{\log_2|\mathbf{I}_m + \frac{P}{c}\mathbf{A}\mathbf{A}^H|\} \to \min\{m,n\}\log_2 P + o(\log_2 P)$ as $P\to\infty$.

Therefore, it follows from (82) that as $P_2 \to \infty$,

$$I(\mathbf{y}'_j;\mathbf{x}_i|\mathbf{h}_{j,i}) \geq K_2 N \log_2 P_2 + o(\log_2 P_2). \quad (83)$$

2) Analysis of $I(\mathbf{y}'_E;\mathbf{x}_i|\hat{\mathbf{h}}_{E,i})$: We can write

$$I(\mathbf{y}'_E;\mathbf{x}_i|\hat{\mathbf{h}}_{E,i}) = h(\mathbf{y}'_E|\hat{\mathbf{h}}_{E,i}) - h(\mathbf{y}'_E|\mathbf{x}_i,\hat{\mathbf{h}}_{E,i}) \quad (84)$$

where

$$h(\mathbf{y}'_E|\hat{\mathbf{h}}_{E,i}) \leq \mathcal{E}_{\hat{h}}\{\log[(\pi e)^{N_E K_2}|\mathbf{K}_{\mathbf{y}'_E|\hat{\mathbf{h}}_{E,i}}|]\}, \quad (85)$$

and the expectation $\mathcal{E}_{\hat{h}}$ is over the distribution of $\hat{\mathbf{h}}_{E,i}$. The inequality is because of the non-Gaussian nature of $\mathbf{y}'_E$ (see the second term in (76)) given $\hat{\mathbf{h}}_{E,i}$. It follows from (76) that the covariance matrix $\mathbf{K}_{\mathbf{y}'_E|\hat{\mathbf{h}}_{E,i}}$ of $\mathbf{y}'_E$ given $\hat{\mathbf{h}}_{E,i}$ is

$$\begin{aligned}\mathbf{K}_{\mathbf{y}'_E|\hat{\mathbf{h}}_{E,i}} &= \frac{P_2}{N}(\mathbf{I}_{K_2}\otimes\hat{\mathbf{H}}_{E,i}\hat{\mathbf{H}}^H_{E,i}) \\ &\quad + \mathbf{M}_{E,i} + \mathbf{I}_{N_E K_2} \end{aligned} \quad (86)$$

with

$$\begin{aligned}\mathbf{M}_{E,i} &= \mathcal{E}_{x,\Delta h}\{(\mathbf{I}_{K_2}\otimes\Delta\mathbf{H}_{E,i})\mathbf{x}_i\mathbf{x}_i^H(\mathbf{I}_{K_2}\otimes\Delta\mathbf{H}_{E,i})^H\} \\ &= \mathcal{E}_{\Delta h}\{(\mathbf{I}_{K_2}\otimes\Delta\mathbf{H}_{E,i})\frac{P_2}{N}\mathbf{I}_{NK_2}(\mathbf{I}_{K_2}\otimes\Delta\mathbf{H}_{E,i})^H\} \\ &= \frac{P_2}{N}(\mathbf{I}_{K_2}\otimes\mathcal{E}_{\Delta h}\{\Delta\mathbf{H}_{E,i}\Delta\mathbf{H}^H_{E,i}\}) \end{aligned} \quad (87)$$

and $\mathcal{E}_{x,\Delta h}$ is the expectation over $\mathbf{x}_i$ and $\Delta\mathbf{H}_{E,i}$. Using Property 2 for large E_1, one can verify

$$\mathbf{M}_{E,i} = \frac{P_2\sigma_E^2}{M}\mathbf{I}_{N_E K_2}. \quad (88)$$

Since $\mathbf{y}'_E$ is Gaussian when conditioned on $\mathbf{x}_i$ and $\hat{\mathbf{h}}_{E,i}$, the second term in (84) is

$$h(\mathbf{y}'_E|\mathbf{x}_i,\hat{\mathbf{h}}_{E,i}) = \mathcal{E}_x\{\log[(\pi e)^{N_E K_2}|\mathbf{K}_{\mathbf{y}'_E|\mathbf{x}_i,\hat{\mathbf{h}}_{E,i}}|]\}. \quad (89)$$

where $\mathcal{E}_x$ is the expectation over $\mathbf{x}_i$, and $\mathbf{K}_{\mathbf{y}'_E|\mathbf{x}_i,\hat{\mathbf{h}}_{E,i}}$ is the covariance matrix of $\mathbf{y}'_E$ given $\mathbf{x}_i$ and $\hat{\mathbf{h}}_{E,i}$.

We can rewrite $\mathbf{y}'_E$ in (76) as

$$\mathbf{y}'_E = (\mathbf{X}_i^T\otimes\mathbf{I}_{N_E})\hat{\mathbf{h}}_{E,i} + (\mathbf{X}_i^T\otimes\mathbf{I}_{N_E})\Delta\mathbf{h}_{E,i} + \mathbf{n}'_E. \quad (90)$$

Therefore,

$$\begin{aligned}\mathbf{K}_{\mathbf{y}'_E|\mathbf{x}_i,\hat{\mathbf{h}}_{E,i}} &= (\mathbf{X}_i^T\otimes\mathbf{I}_{N_E})\mathbf{K}_{\Delta\mathbf{h}_{E,i}}(\mathbf{X}_i^*\otimes\mathbf{I}_{N_E}) + \mathbf{I}_{N_E K_2} \\ &= \frac{\sigma_E^2}{M}\mathbf{X}_i^T\mathbf{X}_i^*\otimes\mathbf{I}_{N_E} + \mathbf{I}_{N_E K_2} \end{aligned} \quad (91)$$

where we have applied Property 2 for large E_1.

Define that $\bar{\mathbf{X}}_i = \sqrt{\frac{N}{P_2}}\mathbf{X}_i$ if $K_2 \geq N$, or $\bar{\mathbf{X}}_i = \sqrt{\frac{N}{P_2}}\mathbf{X}_i^H$ if $K_2 < N$. Also define $n' = \min\{N,K_2\}$. So, all elements in $\bar{\mathbf{X}}_i$ are i.i.d. $\mathcal{CN}(0,1)$, and $\bar{\mathbf{X}}_i^T\bar{\mathbf{X}}_i^*$ has the full rank n'.

Then,

$$\mathcal{E}_x\{\log[|\mathbf{K}_{\mathbf{y}'_E|\mathbf{x}_i,\hat{\mathbf{h}}_{E,i}}|]\} = N_E\mathcal{E}_x\left\{\log_2\left|\frac{P_2\sigma_E^2}{MN}\bar{\mathbf{X}}_i^T\bar{\mathbf{X}}_i^* + \mathbf{I}_{n'}\right|\right\}. \quad (92)$$

Applying the matrix Minkowskis inequality $|\mathbf{A}+\mathbf{B}|^{\frac{1}{n}} \geq |\mathbf{A}|^{\frac{1}{n}} + |\mathbf{B}|^{\frac{1}{n}}$ for any positive definite $\mathbf{A}$ and $\mathbf{B} \in \mathbb{C}^{n\times n}$ [19], we have

$$\begin{aligned}&\mathcal{E}_x\{\log[|\mathbf{K}_{\mathbf{y}'_E|\mathbf{x}_i,\hat{\mathbf{h}}_{E,i}}|]\} \\ &\geq N_E n'\mathcal{E}_x\left\{\log_2\left(1+\left|\frac{P_2\sigma_E^2}{MN}\bar{\mathbf{X}}_i^T\bar{\mathbf{X}}_i^*\right|^{\frac{1}{n'}}\right)\right\} \\ &= N_E n'\mathcal{E}_x\left\{\log_2\left(1+\frac{P_2\sigma_E^2}{MN}\exp\left(\frac{1}{n'}\ln\left|\bar{\mathbf{X}}_i^T\bar{\mathbf{X}}_i^*\right|\right)\right)\right\}. \end{aligned} \quad (93)$$

Note that $\log(1+a\exp(bx))$ for real a and b is a convex function of x. Then $\mathcal{E}\{\log(1+a\exp(bx))\} \geq \log(1+a\exp(b\mathcal{E}\{x\}))$ due to the Jensen's inequality [20]. Then

$$\mathcal{E}\{\log[|\mathbf{K}_{\mathbf{y}_E|\mathbf{x}_i,\hat{\mathbf{h}}_{E,i}}|]\} \geq N_E n'\log_2\left(1+\frac{P_2\sigma_E^2}{MN}\theta\right), \quad (94)$$

where $\theta = \exp(\frac{1}{n'}\mathcal{E}_x\ln|\bar{\mathbf{X}}_i^T\bar{\mathbf{X}}_i^*|)$ which is invariant to P_2. Furthermore, based on [21, Th. 1], we can write $\theta = \exp(\frac{1}{n'}\sum_{k=1}^{n'}\sum_{n=1}^{\max\{N,K_2\}-k}\frac{1}{n} - \gamma)$ with γ being the Euler's constant.

Combining the above results (i.e., (85), (86), (88), (89) and (94)) into (84), we have

$$\begin{aligned}&I(\mathbf{y}'_E;\mathbf{x}_i|\hat{\mathbf{h}}_{E,i}) \\ &\leq \mathcal{E}_{\hat{h}}\{\log[|\mathbf{K}_{\mathbf{y}'_E|\hat{\mathbf{h}}_{E,i}}|]\} - \mathcal{E}_x\{\log[|\mathbf{K}_{\mathbf{y}'_E|\mathbf{x}_i,\hat{\mathbf{h}}_{E,i}}|]\} \\ &= K_2\mathcal{E}_{\hat{h}}\left\{\log_2\left|\mathbf{I}_{N_E}+\frac{P_2}{N\left(1+\frac{P_2\sigma_E^2}{M}\right)}\hat{\mathbf{H}}_{E,i}\hat{\mathbf{H}}^H_{E,i}\right|\right\} \\ &\quad + N_E K_2\log_2\left(1+\frac{P_2\sigma_E^2}{M}\right) - N_E n'\log_2\left(1+\frac{P_2\sigma_E^2}{MN}\theta\right). \end{aligned} \quad (95)$$

As $P_2\to\infty$, the first term in (95) approaches a constant, but the last two terms in (95) approaches either a constant if $n' \doteq \min(N,K_2) = K_2$ or $N_E(K_2-N)\log_2 P_2$ plus a constant if $n' \doteq \min(N,K_2) = N$.

Therefore, we conclude that as $P_2\to\infty$,

$$I(\mathbf{y}'_E;\mathbf{x}_i|\hat{\mathbf{h}}_{E,i}) \leq N_E(K_2-n')\log_2 P_2 + o(\log_2 P_2) \quad (96)$$

Property 10: For one-way information transmission from one user to another following all-user ANECE, the secrecy capacity $C_{trans}^{(one)}$ in bits per channel coherence period has the following property. For $E_1 = P_1K_1\to\infty$ and $P_2 = \frac{E_2}{K_2}\to\infty$,

$$C_{trans}^{(one)} \geq \eta_1\log_2 P_2 + o(\log_2 P_2) \quad (97)$$

with

$$\eta_1 = [K_2N - N_E(K_2-n')]^+ \quad (98)$$

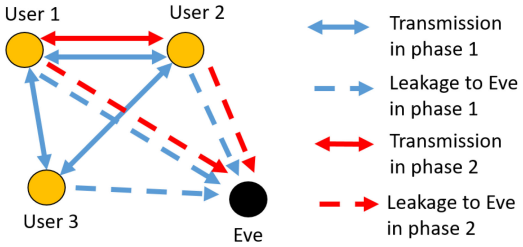


Fig. 4. Illustration of an ANECE based scheme where two or more users jointly perform ANECE in phase 1 but only users 1 and 2 exchange secret information with each other in phase 2 against Eve with any number of antennas.

and $n' = \min(N, K_2)$. For $K_2 \leq N$, $\eta_1 = K_2 N$ regardless of N_E. For $K_2 > N, \eta_1 = [K_2 N - N_E(K_2 - N)]^+$ which is positive if and only if $N_E < \frac{K_2 N}{K_2 - N}$.

Proof: Using (83) and (96) in (77) yields (97). The rest is also easy to verify.

We see that to ensure a positive unconditional secrecy for large E_1 and P_2 (unconditional on N_E), we need to choose $K_2 \leq N$. This property is unique from those shown in prior works such as [4] and [10].

We also see that for $K_2 \geq N$ and large E_1 and P_2, $\frac{C_{trans}^{(one)}}{K_2 \log_2 P_2} \gtrsim (\frac{N N_E}{K_2} + N - N_E)^+$. This means that to ensure a positive non-diminishing degree of freedom in the secrecy rate $\frac{C_{trans}^{(one)}}{K_2}$ in bits/s/Hz for large E_1, P_2 and K_2, we need to choose $N > N_E$ under which $\lim_{K_2 \to \infty} \frac{C_{trans}^{(one)}}{K_2 \log_2 P_2} \gtrsim N - N_E > 0$.

Finally, we note that Property 10 is equivalent to (60) in [8] although the latter was derived for a single pair of nodes in both phases 1 and 2. In other words, the all-user ANECE for any $M \geq 2$ users does not change the result of $C_{trans}^{(one)}$ provided that E_1 and P_2 are sufficiently large.

B. Two-Way Transmission After All-User ANECE

In this section, we consider a two-way transmission between one pair of full-duplex users following all-user ANECE. This is illustrated in Fig. 4. Specifically, in phase 2, users i and j transmit two independent signals $\mathbf{x}_i(k)$ and $\mathbf{x}_j(k)$ respectively to each other for $k = 1, \ldots, K_2$. The distributions of the elements in the signals are the same as in the one-way case, i.e., i.i.d. $\mathcal{CN}(0, \frac{P_2}{N})$. Ignoring the residual self-interference, the signals received by users i and j have the same form as in the case of one-way transmission, i.e., (75).

But the signal received by Eve now has contributions from both users i and j, i.e.,

$$\mathbf{Y}'_E = \mathbf{H}_{E,i}\mathbf{X}_i + \mathbf{H}_{E,j}\mathbf{X}_j + \mathbf{N}_E,' \tag{99}$$

or equivalently in vector form,

$$\mathbf{y}'_E = (\mathbf{I}_{K_2} \otimes \mathbf{H}_{E,i})\mathbf{x}_i + (\mathbf{I}_{K_2} \otimes \mathbf{H}_{E,j})\mathbf{x}_j + \mathbf{n}'_E \tag{100}$$

which differs from (76).

As in the one-way case, we assume a large E_1 so that both users i and j effectively know $\mathbf{H}_{i,j}$ and the channel estimation errors by Eve follows Properties 2 and 3.

Using $\mathbf{H}_{E,i} = \hat{\mathbf{H}}_{E,i} + \Delta\mathbf{H}_{E,i}$, (100) becomes

$$\begin{aligned}\mathbf{y}'_E = &(\mathbf{I}_{K_2} \otimes \hat{\mathbf{H}}_{E,i})\mathbf{x}_i + (\mathbf{I}_{K_2} \otimes \Delta\mathbf{H}_{E,i})\mathbf{x}_i \\ &+ (\mathbf{I}_{K_2} \otimes \hat{\mathbf{H}}_{E,j})\mathbf{x}_j + (\mathbf{I}_{K_2} \otimes \Delta\mathbf{H}_{E,j})\mathbf{x}_j + \mathbf{n}'_E.\end{aligned} \tag{101}$$

The secrecy capacity in bits per coherence period of the two-way transmission between users i and j can be defined as

$$\begin{aligned}C_{trans}^{(two)} = &\left[I\left(\mathbf{y}'_j; \mathbf{x}_i | \mathbf{h}_{j,i}\right) + I\left(\mathbf{y}'_i; \mathbf{x}_j | \mathbf{h}_{i,j}\right)\right. \\ &\left. - I\left(\mathbf{y}'_E; \mathbf{x}_i, \mathbf{x}_j | \hat{\mathbf{h}}_{E,i}, \hat{\mathbf{h}}_{E,j}\right)\right]^+.\end{aligned} \tag{102}$$

The analysis of the first two terms in (102) follow the same steps as for the one-way case. Namely, each of the two terms is governed by (83). We will need to focus on the third term $I(\mathbf{y}'_E; \mathbf{x}_i, \mathbf{x}_j | \hat{\mathbf{h}}_{E,i}, \hat{\mathbf{h}}_{E,j})$ in (102).

1) Analysis of $I(\mathbf{y}'_E; \mathbf{x}_i, \mathbf{x}_j | \hat{\mathbf{h}}_{E,i}, \hat{\mathbf{h}}_{E,j})$: We first write

$$\begin{aligned}I(\mathbf{y}'_E; \mathbf{x}_i, \mathbf{x}_j | \hat{\mathbf{h}}_{E,i}, \hat{\mathbf{h}}_{E,j}) = &\, h(\mathbf{y}'_E | \hat{\mathbf{h}}_{E,i}, \hat{\mathbf{h}}_{E,j}) \\ &- h(\mathbf{y}'_E | \mathbf{x}_i, \mathbf{x}_j, \hat{\mathbf{h}}_{E,i}, \hat{\mathbf{h}}_{E,j}).\end{aligned} \tag{103}$$

where the first term is

$$h\left(\mathbf{y}'_E | \hat{\mathbf{h}}_{E,i}, \hat{\mathbf{h}}_{E,j}\right) \leq \mathcal{E}\left\{\log_2\left[(\pi e)^{N_E K_2}\left|\mathbf{K}_{\mathbf{y}_E | \hat{\mathbf{h}}_{E,i}, \hat{\mathbf{h}}_{E,j}}\right|\right]\right\}, \tag{104}$$

with

$$\begin{aligned}\mathbf{K}_{\mathbf{y}'_E | \hat{\mathbf{h}}_{E,i}, \hat{\mathbf{h}}_{E,j}} = &\frac{P_2}{N}\left(\mathbf{I}_{K_2} \otimes \hat{\mathbf{H}}_{E,i}\hat{\mathbf{H}}_{E,i}^H\right) \\ &+ \frac{P_2}{N}\left(\mathbf{I}_{K_2} \otimes \hat{\mathbf{H}}_{E,j}\hat{\mathbf{H}}_{E,j}^H\right) \\ &+ \mathbf{M}_{E,i} + \mathbf{M}_{E,j} + \mathbf{I}_{N_E K_2}.\end{aligned} \tag{105}$$

Note that the nonzero cross-correlation between $\Delta\mathbf{H}_{E,i}$ and $\Delta\mathbf{H}_{E,j}$ as shown in Property 3 does not matter in the above due to independence between $\mathbf{x}_i$ and $\mathbf{x}_j$. Using (88) in (105), we have

$$\begin{aligned}\mathbf{K}_{\mathbf{y}'_E | \hat{\mathbf{h}}_{E,i}, \hat{\mathbf{h}}_{E,j}} = \mathbf{I}_{K_2} \otimes &\left(\frac{P_2}{N}\hat{\mathbf{H}}_{E,i}\hat{\mathbf{H}}_{E,i}^H + \frac{P_2}{N}\hat{\mathbf{H}}_{E,j}\hat{\mathbf{H}}_{E,j}^H\right. \\ &\left. + \left(\frac{2P_2\sigma_E^2}{M} + 1\right)\mathbf{I}_{N_E}\right).\end{aligned} \tag{106}$$

Then, for $P_2 \to \infty$,

$$\begin{aligned}&\mathcal{E}\{\log|\mathbf{K}_{\mathbf{y}'_E | \hat{\mathbf{h}}_{E,i}, \hat{\mathbf{h}}_{E,j}}|\} \\ &= K_2 \mathcal{E}\left\{\log_2\left|\frac{\frac{P_2}{N}\hat{\mathbf{H}}_{E,i}\hat{\mathbf{H}}_{E,i}^H + \frac{P_2}{N}\hat{\mathbf{H}}_{E,j}\hat{\mathbf{H}}_{E,j}^H}{\frac{2P_2\sigma_E^2}{M} + 1} + \mathbf{I}_{N_E}\right|\right\} \\ &\quad + N_E K_2 \log_2\left(\frac{2P_2\sigma_E^2}{M} + 1\right) \\ &= N_E K_2 \log_2 P_2 + o(\log_2 P_2).\end{aligned} \tag{107}$$

The second term in (103) is

$$h\left(\mathbf{y}'_E | \mathbf{x}_i, \mathbf{x}_j, \hat{\mathbf{h}}_{E,i}, \hat{\mathbf{h}}_{E,j}\right)$$

$$= \mathcal{E}\left\{\log_2\left[(\pi e)^{N_E K_2}\left|\mathbf{K}_{\mathbf{y}'_E|\mathbf{x}_i,\mathbf{x}_j,\hat{\mathbf{h}}_{E,i},\hat{\mathbf{h}}_{E,j}}\right|\right]\right\}. \quad (108)$$

where $\mathbf{K}_{\mathbf{y}'_E|\mathbf{x}_i,\mathbf{x}_j,\hat{\mathbf{h}}_{E,i},\hat{\mathbf{h}}_{E,j}}$ is the covariance matrix of $\mathbf{y}'_E$ given $\mathbf{x}_i$, $\mathbf{x}_j$, $\hat{\mathbf{h}}_{E,i}$ and $\hat{\mathbf{h}}_{E,j}$. We now rewrite $\mathbf{y}'_E$ in (100) as

$$\begin{aligned}\mathbf{y}'_E &= \left(\mathbf{X}_i^T \otimes \mathbf{I}_{N_E}\right)\hat{\mathbf{h}}_{E,i} + \left(\mathbf{X}_i^T \otimes \mathbf{I}_{N_E}\right)\Delta\mathbf{h}_{E,i} \\ &\quad + \left(\mathbf{X}_j^T \otimes \mathbf{I}_{N_E}\right)\hat{\mathbf{h}}_{E,j} + \left(\mathbf{X}_j^T \otimes \mathbf{I}_{N_E}\right)\Delta\mathbf{h}_{E,j} + \mathbf{n}_E.\end{aligned} \quad (109)$$

where the first and third terms do not affect $\mathbf{K}_{\mathbf{y}_E|\mathbf{x}_i,\mathbf{x}_j,\hat{\mathbf{h}}_{E,i},\hat{\mathbf{h}}_{E,j}}$, i.e.,

$$\begin{aligned}&\mathbf{K}_{\mathbf{y}'_E|\mathbf{x}_i,\mathbf{x}_j,\hat{\mathbf{h}}_{E,i},\hat{\mathbf{h}}_{E,j}} \\ &= \left(\mathbf{X}_i^T \otimes \mathbf{I}_{N_E}\right)\mathbf{K}_{\Delta\mathbf{h}_{E,i}}\left(\mathbf{X}_i^* \otimes \mathbf{I}_{N_E}\right) \\ &\quad + \left(\mathbf{X}_j^T \otimes \mathbf{I}_{N_E}\right)\mathbf{K}_{\Delta\mathbf{h}_{E,j}}\left(\mathbf{X}_j^* \otimes \mathbf{I}_{N_E}\right) \\ &\quad + \left(\mathbf{X}_i^T \otimes \mathbf{I}_{N_E}\right)\mathbf{K}_{\Delta\mathbf{h}_{E,i},\Delta\mathbf{h}_{E,j}}\left(\mathbf{X}_j^* \otimes \mathbf{I}_{N_E}\right) \\ &\quad + \left(\mathbf{X}_j^T \otimes \mathbf{I}_{N_E}\right)\mathbf{K}_{\Delta\mathbf{h}_{E,i},\Delta\mathbf{h}_{E,j}}\left(\mathbf{X}_i^* \otimes \mathbf{I}_{N_E}\right) + \mathbf{I}_{N_E K_2} \\ &= \left(\frac{\sigma_E^2}{M}\mathbf{X}_i^T\mathbf{X}_i^* + \frac{\sigma_E^2}{M}\mathbf{X}_j^T\mathbf{X}_j^* + \frac{\sigma_E^2}{M}\mathbf{X}_i^T\mathbf{X}_j^*\right. \\ &\quad \left. + \frac{\sigma_E^2}{M}\mathbf{X}_j^T\mathbf{X}_i^* + \mathbf{I}_{K_2}\right) \otimes \mathbf{I}_{N_E}.\end{aligned} \quad (110)$$

Here the nonzero cross-correlation between $\Delta\mathbf{h}_{E,i}$ and $\Delta\mathbf{h}_{E,j}$ in Property 3 has been applied. (An error about this occurred in Appendix A.3 in [8].)

It then follows that

$$\begin{aligned}&\mathcal{E}\left\{\log_2\left|\mathbf{K}_{\mathbf{y}'_E|\mathbf{x}_i,\mathbf{x}_j,\hat{\mathbf{h}}_{E,i},\hat{\mathbf{h}}_{E,j}}\right|\right\} \\ &= N_E\mathcal{E}\left\{\log_2\left|\frac{\sigma_E^2}{M}\left(\mathbf{X}_i^T + \mathbf{X}_j^T\right)\left(\mathbf{X}_i^* + \mathbf{X}_j^*\right) + \mathbf{I}_{K_2}\right|\right\}.\end{aligned} \quad (111)$$

Define $\mathbf{X}_{i,j} = \sqrt{\frac{N}{2P_2}}(\mathbf{X}_i^T + \mathbf{X}_j^T)$, whose entries are i.i.d. $\mathcal{CN}(0,1)$. Then,

$$\begin{aligned}&\mathcal{E}\left\{\log_2\left|\mathbf{K}_{\mathbf{y}'_E|\mathbf{x}_i,\mathbf{x}_j,\hat{\mathbf{h}}_{E,i},\hat{\mathbf{h}}_{E,j}}\right|\right\} \\ &= N_E\mathcal{E}\left\{\log_2\left|\mathbf{I}_{K_2} + \frac{2P_2\sigma_E^2}{MN}\mathbf{X}_{i,j}\mathbf{X}_{i,j}^H\right|\right\}.\end{aligned} \quad (112)$$

Following a similar analysis as shown before for $\mathcal{E}\{\log_2|\mathbf{K}_{\mathbf{y}'_E|\mathbf{x}_i,\hat{\mathbf{h}}_{E,i}}|\}$ in the one-way case, we have that as $P_2 \to \infty$,

$$\begin{aligned}&\mathcal{E}\left\{\log_2\left|\mathbf{K}_{\mathbf{y}_E|\mathbf{x}_i,\mathbf{x}_j,\hat{\mathbf{h}}_{E,i},\hat{\mathbf{h}}_{E,j}}\right|\right\} \\ &= N_E n' \log_2 P_2 + o(\log_2 P_2).\end{aligned} \quad (113)$$

with $n' = \min(N, K_2)$.

Applying the results (104), (107), (108) and (113) into (103), we have that as $P_2 \to \infty$,

$$\begin{aligned}&I\left(\mathbf{y}'_E; \mathbf{x}_i, \mathbf{x}_j | \hat{\mathbf{h}}_{E,i}, \hat{\mathbf{h}}_{E,j}\right) \\ &\le N_E K_2 \log_2 P_2 - N_E n' \log_2 P_2 + o(\log_2 P_2).\end{aligned} \quad (114)$$

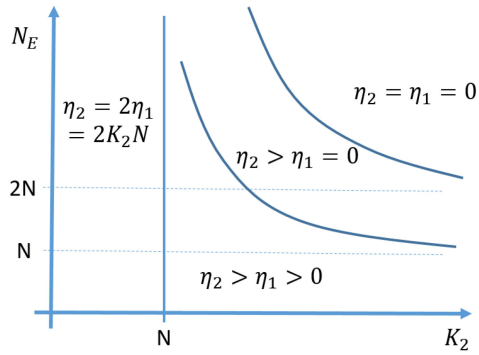


Fig. 5. Illustration of regions of the SDoF of information transmission for one-way and two-way schemes. The upper curve is governed by $N_E = \frac{2K_2N}{K_2-N}$ without integer constraint while the lower dashed curve is by $N_E = \frac{K_2N}{K_2-N}$ without integer constraint.

Therefore, the following property is ready to be verified.

Property 11: For two-way information transmission between a pair of users following all-user ANECE, the secrecy capacity $C_{trans}^{(two)}$ in bits per channel coherence period has the following property. For $E_1 = P_1K_1 \to \infty$ and $P_2 = \frac{E_2}{K_2} \to \infty$,

$$C_{trans}^{(two)} \ge \eta_2 \log_2 P_2 + o(\log_2 P_2) \quad (115)$$

with

$$\eta_2 = [2K_2N - N_E(K_2 - n')]^+ \quad (116)$$

and $n' = \min(N, K_2)$. For $K_2 \le N$, $\eta_2 = 2K_2N$ regardless of N_E. For $K_2 > N$, $\eta_2 = 2K_2N - N_E(K_2 - N)$ which is positive if and only if $N_E < \frac{2K_2N}{K_2-N}$.

Proof: Using (83) and (114) in (102) yields (115). The rest is also easy to verify.

Note that Property 11 is a correction of (74) in [8] where the cross-correlation between $\Delta\mathbf{h}_{E,i}$ and $\Delta\mathbf{h}_{E,j}$ was incorrectly treated as zero. Like the one-way case, the value of $M \ge 2$ in the M-user (all-user) ANECE does not affect $C_{trans}^{(two)}$ under $E_1 \to \infty$ and $P_2 \to \infty$.

C. Comparison Between One-Way and Two-Way

Comparing Properties 10 and 11, we see that the condition for a positive unconditional secrecy for both one-way and two-way cases is the same, i.e., $K_2 \le N$. But when positive, $C_{trans}^{(two)}$ is larger than $C_{trans}^{(one)}$ due to the factor two in the first term in (116). Also unlike the one-way case, to have a positive non-diminishing degree of freedom in the secrecy rate $\frac{C_{trans}^{(two)}}{K_2}$ in bits/s/Hz for large E_1, P_2 and K_2, we only need to have $N > \frac{N_E}{2}$ (instead of $N > N_E$) under which $\lim_{K_2\to\infty}\frac{C_{trans}^{(two)}}{K_2\log_2 P_2} \gtrsim 2N - N_E + \frac{N_E N}{K_2} > 0$.

The SDoF of $C_{trans}^{(one)}$ and $C_{trans}^{(two)}$ with respect to $\log_2 P_2$ in phase 2 are simply $\frac{\eta_1}{K_2}$ and $\frac{\eta_2}{K_2}$ respectively. Four regions of η_1 and η_2 in terms of K_2 and N_E with reference to N are illustrated in Fig. 5. Recall that K_2 is the length of transmission in phase 2, N_E is the number of antennas on Eve, and N is the number of antennas on each user. On and to the left of the vertical line

at $K_2 = N$, $\eta_2 = 2\eta_1 = 2K_2N$. Below the lower curve, $\eta_2 > \eta_1 > 0$. On and above the lower curve and below the upper curve, $\eta_2 > \eta_1 = 0$. On and above the upper curve, $\eta_2 = \eta_1 = 0$.

We see that the upper and lower curves in Fig. 5 converge towards $2N$ and N respectively as K_2 increases. But they are always above $2N$ and N respectively. Therefore, if $N_E \leq N$, then $\eta_1 > 0$ for all $K_2 \geq 1$. And if $N_E \leq 2N$, then $\eta_2 > 0$ for all $K_2 \geq 1$.

But due to integer constraint on N_E, there is a finite integer $V_1 = \min(K_2)$ subject to $\lfloor \frac{K_2N}{K_2-N} \rfloor = N$. For $K_2 \geq V_1$, $\eta_1 = 0$ if and only if $N_E \geq N+1$. Similarly, there is a finite integer $V_2 = \min(K_2)$ subject to $\lfloor \frac{2K_2N}{K_2-N} \rfloor = 2N$. For $K_2 \geq V_2$, $\eta_2 = 0$ if and only if $N_E \geq 2N+1$.

In fact, V_1 is the smallest integer K_2 satisfying $\frac{K_2N}{K_2-N} < N+1$, which implies $K_2 > N^2+N$ and hence $V_1 = N^2+N+1$. Similarly, V_2 is the smallest integer K_2 satisfying $\frac{2K_2N}{K_2-N} < 2N+1$, which implies $K_2 > 2N^2+N$ and hence $V_2 = 2N^2+N+1$. Here $\frac{V_2}{V_1} = 1 + \frac{N^2}{N^2+N+1} \to 2$ as $N \to \infty$.

The above phenomena of positive or zero SDoF in terms of N_E are quite different from the previous results in [4] and [10]. In the case of [10] assuming N antennas on each of transmit and receive nodes and N_E antennas on Eve, there is zero SDoF if $N_E \geq N$ and $K \geq 2N$. Here K is the number of channel uses in each coherence period. In the case of [4], zero SDoF is also implied by $N_E \geq N$.

Of course, the system models in [4] and [10] are different from ours in phase 2. The model in [4] assumes that Eve knows her CSI as well as the CSI of the users. The model in [10] assumes that CSI anywhere is unknown everywhere initially. The properties of η_1 and η_2 shown above have benefited from ANECE in phase 1, which not only allowed users to obtain their CSI but also prevented Eve from obtaining any of her own CSI and users' CSI.

V. Monte Carlo Simulation

The results shown so far are completely based on mathematical analysis. In this section, we show a Monte Carlo simulation to validate the theoretical results of η_1 and η_2 shown in (98) and (116). To make the simulation feasible without unnecessary numerical issues, we need to reformulate the expressions of (77) and (102). We next provide some of the details leading up to an alternative expression of (102). Then, a similar expression of (77) is also provided.

We can rewrite (102) as

$$C_{trans}^{(two)} = \left[\mathcal{E}\left\{ \log_2 \frac{f(\mathbf{y}'_j|\mathbf{x}_i, \mathbf{h}_{j,i})}{f(\mathbf{y}'_j|\mathbf{h}_{j,i})} \right\} + \mathcal{E}\left\{ \log_2 \frac{f(\mathbf{y}'_i|\mathbf{x}_j, \mathbf{h}_{j,i})}{f(\mathbf{y}'_i|\mathbf{h}_{j,i})} \right\} - \mathcal{E}\left\{ \log_2 \frac{f(\mathbf{y}'_E|\mathbf{x}_i, \mathbf{x}_j, \hat{\mathbf{h}}_{E,i}, \hat{\mathbf{h}}_{E,j})}{f(\mathbf{y}'_E|\hat{\mathbf{h}}_{E,i}, \hat{\mathbf{h}}_{E,j})} \right\} \right]^+ \quad (117)$$

where the first two terms (after the expectations) are identical due to statistical symmetry between node i and node j.

Recall the following probability density functions (PDFs) $f(\mathbf{x}_i) = \mathcal{CN}(0, \frac{P_2}{N}\mathbf{I}_{NK_2})$, $f(\mathbf{x}_j) = \mathcal{CN}(0, \frac{P_2}{N}\mathbf{I}_{NK_2})$ $f(\mathbf{h}_{j,i}) = \mathcal{CN}(0, \mathbf{I}_{N^2})$, $f(\mathbf{h}_{E,i}) = \mathcal{CN}(0, \sigma_E^2\mathbf{I}_{N_EN})$, $f(\mathbf{h}_{E,j}) = \mathcal{CN}(0, \sigma_E^2\mathbf{I}_{N_EN})$, $f(\Delta\mathbf{h}_{E,i}) = \mathcal{CN}(0, \frac{1}{M}\sigma_E^2\mathbf{I}_{N_EN})$. Also note that $\mathbf{x}_i$, $\mathbf{x}_j$, $\mathbf{h}_{j,i}$, $\mathbf{h}_{E,i}$, $\mathbf{h}_{E,j}$ and $\Delta\mathbf{h}_{E,i}$ are independent of each other, and $\Delta\mathbf{h}_{E,i} = \Delta\mathbf{h}_{E,j}$ (due to Property 2).

It then follows from (71) that

$$f(\mathbf{y}'_j|\mathbf{x}_i, \mathbf{h}_{j,i}) = \mathcal{CN}(\mathbf{m}_1, \mathbf{R}_1) \quad (118)$$

with $\mathbf{m}_1 = (\mathbf{I}_{K_2} \otimes \mathbf{H}_{j,i})\mathbf{x}_i$ and $\mathbf{R}_1 = \mathbf{I}_{NK_2}$, and

$$f(\mathbf{y}'_j|\mathbf{h}_{j,i}) = \mathcal{CN}(\mathbf{m}_2, \mathbf{R}_2) \quad (119)$$

with $\mathbf{m}_2 = 0$ and $\mathbf{R}_2 = \mathbf{I}_{K_2} \otimes (\frac{P_2}{N}\mathbf{H}_{j,i}\mathbf{H}_{j,i}^H + \mathbf{I}_N)$.

It follows from (101) that

$$f(\mathbf{y}'_E|\mathbf{x}_i, \mathbf{x}_j, \hat{\mathbf{h}}_{E,i}, \hat{\mathbf{h}}_{E,j}) = \mathcal{CN}(\mathbf{m}_5, \mathbf{R}_5) \quad (120)$$

with $\mathbf{m}_5 = (\mathbf{I}_{K_2} \otimes \hat{\mathbf{H}}_{E,i})\mathbf{x}_i + (\mathbf{I}_{K_2} \otimes \hat{\mathbf{H}}_{E,j})\mathbf{x}_j$ and $\mathbf{R}_5 = (\frac{\sigma_E^2}{M}\mathbf{X}_i^T\mathbf{X}_i^* + \frac{\sigma_E^2}{M}\mathbf{X}_j^T\mathbf{X}_j^* + \frac{\sigma_E^2}{M}\mathbf{X}_i^T\mathbf{X}_j^* + \frac{\sigma_E^2}{M}\mathbf{X}_j^T\mathbf{X}_i^* + \mathbf{I}_{K_2}) \otimes \mathbf{I}_{N_E}$ and

$$f(\mathbf{y}'_E|\hat{\mathbf{h}}_{E,i}, \hat{\mathbf{h}}_{E,j}) = \mathcal{NN}(\mathbf{m}_6, \mathbf{R}_6) \quad (121)$$

with $\mathbf{m}_6 = 0$ and $\mathbf{R}_6 = \mathbf{I}_{K_2} \otimes (\frac{P_2}{N}\hat{\mathbf{H}}_{E,i}\hat{\mathbf{H}}_{E,i}^H + \frac{P_2}{N}\hat{\mathbf{H}}_{E,j}\hat{\mathbf{H}}_{E,j}^H + (2\frac{P_2\sigma_E^2}{M} + 1)\mathbf{I}_{N_E})$. Here $\mathcal{NN}(\mathbf{m}, \mathbf{R})$ denotes a non-Gaussian (non-Normal) PDF with mean $\mathbf{m}$ and covariance matrix $\mathbf{R}$. But there is no known closed form of $f(\mathbf{y}'_E|\hat{\mathbf{h}}_{E,i}, \hat{\mathbf{h}}_{E,j})$.

Note that for an n-dimensional circular complex Gaussian random vector $\mathbf{x}$ with mean $\mathbf{m}$ and covariance matrix $\mathbf{R}$ (i.e., $\mathcal{CN}(\mathbf{m}, \mathbf{R})$), its PDF has the form $f(\mathbf{x}) = \frac{1}{\pi^n|\mathbf{R}|}\exp(-(\mathbf{x}-\mathbf{m})^H\mathbf{R}^{-1}(\mathbf{x}-\mathbf{m}))$ and hence $\mathcal{E}\{-\ln f(\mathbf{x})\} = n\ln\pi + \ln|\mathbf{R}| + \mathcal{E}\{(\mathbf{x}-\mathbf{m})^H\mathbf{R}^{-1}(\mathbf{x}-\mathbf{m})\} = n\ln\pi + \ln|\mathbf{R}| + n$.

Using (118)-(121), one can verify that (117) is equivalent to

$$\begin{aligned} C_{trans}^{(two)} = (\log_2 e)[&\mathcal{E}\{-2\|\mathbf{y}'_j - \mathbf{m}_1\|^2 + 2\ln|\mathbf{R}_2| \\ &+ 2\mathbf{y}_j'^H\mathbf{R}_2^{-1}\mathbf{y}'_j + N_EK_2\ln\pi + \ln|\mathbf{R}_5| \\ &+ (\mathbf{y}'_E - \mathbf{m}_5)^H\mathbf{R}_5^{-1}(\mathbf{y}'_E - \mathbf{m}_5) \geq \\ &+\ln f(\mathbf{y}'_E|\hat{\mathbf{h}}_{E,i}, \hat{\mathbf{h}}_{E,j})\}]^+ (\log_2 e)[\mathcal{E}\{-2\|\mathbf{y}'_j - \mathbf{m}_1\|^2 \\ &+ 2\ln|\mathbf{R}_2| + 2\mathbf{y}_j'^H\mathbf{R}_2^{-1}\mathbf{y}'_j + \ln|\mathbf{R}_5| \\ &+ (\mathbf{y}'_E - \mathbf{m}_5)^H\mathbf{R}_5^{-1}(\mathbf{y}'_E - \mathbf{m}_5) \\ &- \ln|\mathbf{R}_6| - \mathbf{y}_E'^H\mathbf{R}_6^{-1}\mathbf{y}'_E\}]^+ \end{aligned} \quad (122)$$

Here the inequality follows from $f(\mathbf{y}'_E|\hat{\mathbf{h}}_{E,i}, \hat{\mathbf{h}}_{E,j})$ being non-Gaussian. Since $\mathcal{E}\{\|\mathbf{y}'_j - \mathbf{m}_1\|^2\} = K_2N$, $\mathcal{E}\{\mathbf{y}_j'^H\mathbf{R}_2^{-1}\mathbf{y}'_j\} = K_2N$, $\mathcal{E}\{(\mathbf{y}'_E - \mathbf{m}_5)^H\mathbf{R}_5^{-1}(\mathbf{y}'_E - \mathbf{m}_5)\} = K_2N_E$ and $\mathcal{E}\{\mathbf{y}_E'^H\mathbf{R}_6^{-1}\mathbf{y}'_E\} = K_2N_E$, (122) is equivalent to

$$C_{trans}^{(two)} \geq (\log_2 e)[\mathcal{E}\{2\ln|\mathbf{R}_2| + \ln|\mathbf{R}_5| - \ln|\mathbf{R}_6|\}]^+ \quad (123)$$

Following a similar approach, we can express $C_{trans}^{(one)}$ in (77) as

$$\begin{aligned} C_{trans}^{(one)} = & \left[\mathcal{E}\left\{\log_2 \frac{f(\mathbf{y}'_j|\mathbf{x}_i,\mathbf{h}_{j,i})}{f(\mathbf{y}'_j|\mathbf{h}_{j,i})}\right\}\right. \\ & \left. - \mathcal{E}\left\{\log_2 \frac{f(\mathbf{y}'_E|\mathbf{x}_i,\hat{\mathbf{h}}_{E,i})}{f(\mathbf{y}'_E|\hat{\mathbf{h}}_{E,i})}\right\}\right]^+ \\ \geq & (\log_2 e)[\mathcal{E}\{-\|\mathbf{y}'_j-\mathbf{m}_1\|^2 + \ln|\mathbf{R}_2| + \mathbf{y}'^H_j\mathbf{R}_2^{-1}\mathbf{y}'_j \\ & + \ln|\mathbf{R}_3| + (\mathbf{y}'_E-\mathbf{m}_3)^H\mathbf{R}_3^{-1}(\mathbf{y}'_E-\mathbf{m}_3) \\ & - \ln|\mathbf{R}_4| - \mathbf{y}'^H_E\mathbf{R}_4^{-1}\mathbf{y}'_E\}]^+. \end{aligned} \quad (124)$$

where $\mathbf{m}_1$ and $\mathbf{R}_2$ are given before, and $\mathbf{m}_3 = (\mathbf{I}_{K_2} \otimes \hat{\mathbf{H}}_{E,i})\mathbf{x}_i$, $\mathbf{R}_3 = (\frac{\sigma_E^2}{M}\mathbf{X}_i^T\mathbf{X}_i^* + \mathbf{I}_{K_2}) \otimes \mathbf{I}_{N_E})$, and $\mathbf{R}_4 = \mathbf{I}_{K_2} \otimes (\frac{P_2}{N}\hat{\mathbf{H}}_{E,i}\hat{\mathbf{H}}_{E,i}^H + (\frac{P_2\sigma_E^2}{M}+1)\mathbf{I}_{N_E})$. Since $\mathcal{E}\{\|\mathbf{y}'_j-\mathbf{m}_1\|^2\} = \mathcal{E}\{\mathbf{y}'^H_j\mathbf{R}_2^{-1}\mathbf{y}'_j\} = K_2N$ and $\mathcal{E}\{(\mathbf{y}'_E-\mathbf{m}_3)^H\mathbf{R}_3^{-1}(\mathbf{y}'_E-\mathbf{m}_3)\} = \mathcal{E}\{\mathbf{y}'^H_E\mathbf{R}_4^{-1}\mathbf{y}'_E\} = K_2N_E$, (124) is equivalent to

$$C_{trans}^{(one)} \geq (\log_2 e)[\mathcal{E}\{\ln|\mathbf{R}_2| + \ln|\mathbf{R}_3| - \ln|\mathbf{R}_4|\}]^+. \quad (125)$$

The expressions of $C_{trans}^{(one)}$ and $C_{trans}^{(two)}$ shown in (125) and (123) are readily useful for simulation where the expectation can be replaced by averaging over many independent realizations of $\mathbf{x}_i$, $\mathbf{x}_j$, $\mathbf{h}_{j,i}$, $\mathbf{h}_{E,i}$, $\mathbf{h}_{E,j}$ and $\Delta\mathbf{h}_{E,i} = \Delta\mathbf{h}_{E,j}$ according to their distributions mentioned earlier.

For a large P_2, we expect both $C_{trans}^{(one)}$ and $C_{trans}^{(two)}$ to have the following structure

$$C_{trans} \geq \bar{C}_{trans} = \eta \log_2 P_2 + B \quad (126)$$

where B is virtually invariant to P_2. (Due to limited numerical range of computer, $\log_2 P_2$ may not be much larger than B even if P_2 is chosen to be 150 dB.) In simulation, we choose $P_2 = 10^5$ (i.e., 50 dB) and $P'_2 = 2 \times 10^5$, from which we compute the corresponding $\bar{C}_{trans}$ and $\bar{C}'_{trans}$ respectively. Then, we compute η by using $\eta = \bar{C}'_{trans} - \bar{C}_{trans}$.

In Fig. 6, we show the simulation results of η_1 and η_2 based on the expressions in (125) and (123), where $N = 2$, $M = 2$, $N_E = 3$, $\sigma_E^2 = 1$, and the number of independent realizations used for averaging is 10^4. The simulation results agree with the theoretical results very well. The small difference between simulation and theory is due to finite sample size and finite power used in simulation. We also observed that the small difference as a function of K_2 varied from one simulation run to another, which is expected.

We would like to add a remark. If one wants to avoid the use of the inequality in (122), one has to compute the non-Gaussian PDF $f(\mathbf{y}'_E|\hat{\mathbf{h}}_{E,i},\hat{\mathbf{h}}_{E,j})$ (without a closed form) for every given set of $\mathbf{y}'_E$, $\hat{\mathbf{h}}_{E,i}$ and $\hat{\mathbf{h}}_{E,j}$. One way to do so is the following

$$\begin{aligned} & f(\mathbf{y}'_E|\hat{\mathbf{h}}_{E,i},\hat{\mathbf{h}}_{E,j}) \\ & = \int f(\mathbf{y}'_E|\mathbf{x}'_i,\mathbf{x}'_j,\hat{\mathbf{h}}_{E,i},\hat{\mathbf{h}}_{E,j}) f(\mathbf{x}'_i,\mathbf{x}'_j) d\mathbf{x}'_i d\mathbf{x}'_j \\ & = \mathcal{E}_{\mathbf{x}'_i,\mathbf{x}'_j}\{f(\mathbf{y}'_E|\mathbf{x}'_i,\mathbf{x}'_j,\hat{\mathbf{h}}_{E,i},\hat{\mathbf{h}}_{E,j})\} \end{aligned} \quad (127)$$

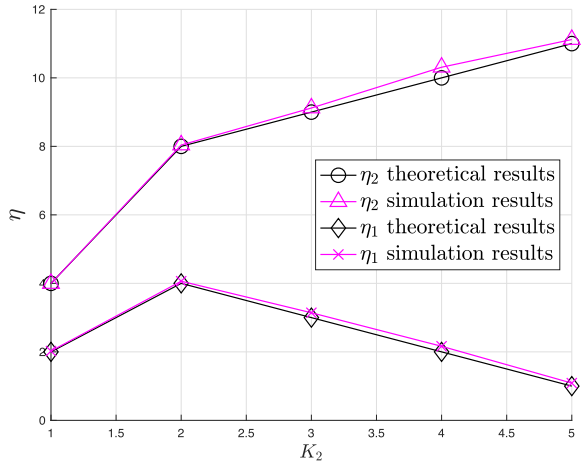


Fig. 6. Illustration of simulation results of η_1 and η_2 in comparison to their theoretical results shown in (98) and (116). The upper two curves are for η_2, and the lower two curves are for η_1.

where $\mathbf{x}'_i$ and $\mathbf{x}'_j$ have the same PDFs as $\mathbf{x}_i$ and $\mathbf{x}_j$, and $f(\mathbf{y}'_E|\mathbf{x}'_i,\mathbf{x}'_j,\hat{\mathbf{h}}_{E,i},\hat{\mathbf{h}}_{E,j})$ equals the Gaussian PDF $f(\mathbf{y}'_E|\mathbf{x}_i,\mathbf{x}_j,\hat{\mathbf{h}}_{E,i},\hat{\mathbf{h}}_{E,j})$ with $\mathbf{x}_i$ and $\mathbf{x}_j$ replaced by $\mathbf{x}'_i$ and $\mathbf{x}'_j$. But the PDF $f(\mathbf{y}'_E|\mathbf{x}'_i,\mathbf{x}'_j,\hat{\mathbf{h}}_{E,i},\hat{\mathbf{h}}_{E,j})$ is highly singular. Namely, for a given set of $\mathbf{y}'_E$ (*function* of $\mathbf{x}_i$ and $\mathbf{x}_j$), $\hat{\mathbf{h}}_{E,i}$ and $\hat{\mathbf{h}}_{E,j}$, $f(\mathbf{y}'_E|\mathbf{x}'_i,\mathbf{x}'_j,\hat{\mathbf{h}}_{E,i},\hat{\mathbf{h}}_{E,j})$ is near zero (especially at high power) for almost all $\mathbf{x}'_i$ and $\mathbf{x}'_j$. This singularity makes (127) difficult to compute reliably (if possible at all[2]).

VI. Further Discussions

Before phase 1 of our scheme, CSI anywhere is assumed to be unknown everywhere, which somehow resembles the case in [10]. For this reason, there should be an interest to look deeper into the similarities and differences between this work and [10]. Furthermore, the work in [11] is also related. Both [10] and [11] are discussed below.

The authors of [10] studied the SDoF for a one-way transmission scheme between two multi-antenna nodes against a multi-antenna Eve where CSI for every node is unknown everywhere. Under the assumption that all channel elements are block-wise i.i.d. Gaussian and all noises are i.i.d. Gaussian, the SDoF in bits/s/Hz(or more precisely in bits per channel use per doubling of large power) shown in [10] can be expressed by

$$d_{prior} = (\min(n_t,n_r) - n_e)^+(K - \min(n_t,n_r))\frac{1}{K} \quad (128)$$

where n_t is the number of antenna at the transmitter, n_r is the number of antennas at the receiver, n_e is the number of antennas at Eve, K is the number of sampling instants per coherence period, and $K \geq 2\min(n_t,n_r)$. This SDoF is zero if $n_e \geq \min(n_t,n_r)$.

The authors of [11] extended the work of [10] to a two-way scheme where two full-duplex multi-antenna nodes transmit information to each other concurrently against a multi-antenna

[2]Using an extremely large number of samples.

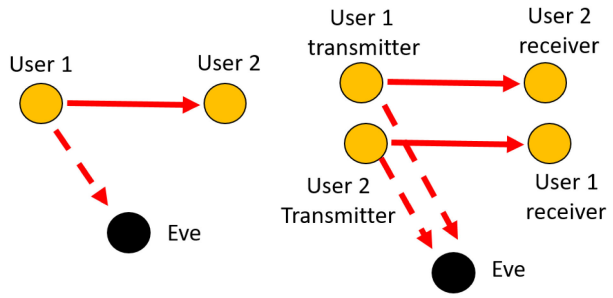


Fig. 7. Comparison of the model in [10] on the left with the (equivalent) model in [11] on the right.

Eve and CSI is unknown everywhere. Specifically, they assume that node 1 use M_1 antennas for transmitting and N_1 antennas for receiving, and node 2 uses M_2 antennas for transmitting and N_2 antennas for receiving. Furthermore, there is no interference between the transmitting antennas and the receiving antennas on each node, and there is Eve with N_E receiving antennas. Assuming i.i.d. channel elements and noises, their proposition 6 states that the maximum SDoF in bits per coherence period of T samples[3] of their scheme is $\max D_s = 2M_1M_2$ if $N_E \geq M_1 + M_2$, $M_1 \leq N_2$, $M_2 \leq N_1$ and $T \geq M_1 + M_2$. There is an agreement between proposition 6 of [11] and our η_2 in (116). Specifically, if $M_1 = M_2 = N_1 = N_2 = N$, then $\max D_s = \max_{K_2} \eta_2 = \eta_2|_{K_2=N} = 2N^2$. The model in [11] does not start with any structure for the transmitted signals while for our two-way scheme we apply ANECE in phase 1 of K_1 samples and two-way information transmission in phase 2 of K_2 samples. Here we achieve the same $\max D_s = 2N^2$ by using $K_1 = N$ and $K_2 = N$, which corresponds to the choice of $T = M_1 + M_2 = 2N$ in [11]. Differing from [11], our use of ANECE in phase 1 (subject to reciprocal channel between users) yields additional secrecy for secret key generation as discussed in section III.

The system model in [11] can be related to that in [10] by choosing $n_t = M_1 + M_2$, $n_r = N_1 + N_2$, $\mathbf{H} = diag(\mathbf{H}_{1,2}, \mathbf{H}_{2,1})$ and $\mathbf{G} = [\mathbf{H}_{1,e}, \mathbf{H}_{2,e}]$. See equations (2) and (3) in [10] and equations (1) and (2) in [11]. In other words, the model in [11] is a special case of that in [10]. See Fig. 7. The only difference is that the effective $(N_1 + N_2) \times (M_1 + M_2)$ channel matrix $\mathbf{H}$ (from an equivalent half-duplex node with $M_1 + M_2$ transmit antennas to another equivalent half-duplex node with $N_1 + N_2$ receive antennas) now has a block diagonal structure and there is no joint encoding between the group of M_1 transmit antennas and the other group of M_2 transmit antennas. We will next focus on a comparison of our results with [10].

To contrast the results shown in this paper against [10], it is important to notice that this paper assumes a reciprocal channel between every pair of full-duplex users/nodes. The reciprocal channel assumption allows a positive secret key capacity C_{key}, which is a bonus secrecy not available in the model considered in [10].

[3]The authors of [11] did not explicitly distinguish the two different units: "bits per coherence period" and "bits/s/Hz" for their choices of SDoF although they treated $\bar{D}_s$ differently from $D_s = T\bar{D}_s$.

So, d_{prior} in (128) is already the TSDoF of the model in [10]. In order to compare with d_{prior}, we define the TSDoF of our scheme/model as follows:

$$d_{new} = \lim_{P\to\infty} \frac{1}{K_1 + K_2} \frac{C_{key} + C_{trans}}{\log_2 P}. \tag{129}$$

Note that C_{key} is due to CSI between users while C_{trans} is independent of CSI between users. See Assumptions A and B stated in section II and section IV respectively. For comparison, we also need to choose $n_t = n_r = N$, $n_e = N_E$ and $K = K_1 + K_2$. Both d_{prior} and d_{new} have the same unit as desired. By choosing $K = K_1 + K_2$, we have ignored the processing time between phase 1 and phase 2 in our scheme. Furthermore, we assume $P_1 = P_2 = P$. For any fixed K_1 and any of the channel training schemes considered earlier (with $M \geq 2$), $\lim_{E_1\to\infty} \frac{C_{key}}{\log_2 E_1} = \lim_{P_1\to\infty} \frac{C_{key}}{\log_2 P_1} = N^2$ due to $E_1 = P_1K_1$. See (42), (50) and (63). Applying Properties 10 and 11, d_{new} becomes

$$d_{new} = \frac{N^2 + \eta}{K_1 + K_2} \tag{130}$$

where $\eta = \eta_1$ as in (98) for one-way transmission and $\eta = \eta_2$ as in (116) for two-way transmission.

Note that from Property 9, ANECE requires $K_1 \geq N$ if $M = 2$. The following property compares d_{new} with d_{prior}.

Property 12: Let $n_t = n_r = N$, $n_e = N_E$, $K = K_1 + K_2$ and $K_1 = N$. For $N_E \geq N$ and $K_2 \geq N$, $d_{prior} = 0$ while $d_{new} = \frac{N^2+\eta}{N+K_2} \geq \frac{N^2}{N+K_2}$. For $N_E < N$ and $K_2 \geq N$,

$$\frac{d_{prior}}{d_{new}} = \frac{(N - N_E)K_2}{N^2 + NN_E + (\xi N - N_E)K_2} < \frac{1}{\xi} \tag{131}$$

where $\xi = 1$ for one-way transmission following ANECE, and $\xi = 2$ for two-way transmission following ANECE. And $\frac{d_{prior}}{d_{new}}$ is an increasing function of K_2 with

$$\left.\frac{d_{prior}}{d_{new}}\right|_{K_2=N} = \frac{N - N_E}{N + \xi N} < \frac{1}{1+\xi}, \tag{132}$$

$$\left.\frac{d_{prior}}{d_{new}}\right|_{K_2=\infty} = \frac{N - N_E}{\xi N - N_E} < \frac{1}{\xi}. \tag{133}$$

Proof: These results follow from the previous discussions of d_{prior} in (128) and d_{new} in (130).

We see that the TSDoF of the ANECE based scheme is in general significantly larger than that shown in [10]. The former exploits full-duplex and reciprocal channels via ANECE while the latter does not. Both schemes assume that CSI is initially unknown everywhere except for their i.i.d. statistical properties.

VII. Conclusion

In this paper, we have shown further insights into the method called anti-eavesdropping channel estimation (ANECE) applicable for a network of cooperative full-duplex radio devices/users. Assuming a symmetric network and a large energy for channel training, we analyzed and compared the capacity C_{key} of secret key generated by each pair of users following pair-wise ANECE, all-user ANECE or the conventional method for channel training. The results show that the secure degree of

freedom (SDoF) in C_{key} is the same for all the three schemes. We also analyzed and compared the secrecy capacity C_{trans} of information transmission between a pair of users using the ANECE-assisted channel estimates. For C_{trans}, we considered a one-way transmission and a two-way transmission subject to a large energy in phase 1 and a large power in phase 2. The results on C_{key} are entirely new. The results on C_{trans} are also significant additions to the previous understandings shown in [7] and [8]. Under the conventional channel training, Eve is allowed to obtain her CSI, which in turn destroys the SDoF of C_{trans} when Eve has a sufficiently large number of antennas. But under ANECE, Eve is not able to do so and hence the SDoF of C_{trans} can be a positive constant invariant to the number of antennas on Eve subject to a limited transmission window in each coherence period. Consequently, there is a net increase of total secure degree of freedom (TSDoF) by using ANECE over the conventional channel training.

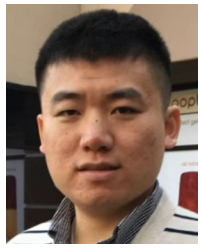
A number of important insights into ANECE are highlighted in Properties 1-12. In particular, Property 12 compares the TSDoF of the ANECE based scheme with that of a conventional scheme in [10], the latter of which does not exploit full-duplex radios or reciprocal channels. The comparison shows a substantial gain of TSDoF by using ANECE.

ACKNOWLEDGMENT

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

REFERENCES

[1] H. V. Poor and R. F. Schaefer, "Wireless physical layer security," *Proc. Nat. Acad. Sci.*, vol. 114, no. 1, pp. 19–26, 2017.
[2] Y.-K. Chia and A. E. Gamal, "Wiretap channel with causal state information," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 2838–2849, May 2012.
[3] 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.
[4] X. He and A. Yener, "MIMO wiretap channels with unknown and varying eavesdropper channel states," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 6844–6869, Nov. 2014.
[5] 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.
[6] 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.
[7] Y. Hua, "Advanced properties of full-duplex radio for securing wireless network," *IEEE Trans. Signal Process.*, vol. 67, no. 1, pp. 120–135, Jan. 2019.
[8] R. Sohrabi, Q. Zhu, and Y. Hua, "Secrecy analyses of a full-duplex MIMOME network," *IEEE Trans. Signal Process.*, vol. 67, no. 23, pp. 5968–5982, Dec. 2019.
[9] Q. Zhu, S. Wu, and Y. Hua, "Optimal pilots for anti-eavesdropping channel estimation," *IEEE Trans. Signal Process.*, vol. 68, pp. 2629–2644, Apr., 2020.
[10] T.-Y. Liu, P. Mukherjee, S. Ulukus, S.-C. Lin, and Y.-W. P. Hong, "Secure degrees of freedom of MIMO Rayleigh block fading wiretap channels with no CSI anywhere," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2655–2669, May 2015.
[11] Q. Liang, D. Liu, and J. Hu, "Secure degrees of freedom of MIMO two-way wiretap channel with no CSI anywhere," *IEEE Trans. Wireless Commun.*, vol. 19, no. 12, pp. 7917–7931, Dec. 2020.
[12] U. M. Maurer, "Secret key agreement by public discussion from common information," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 733–742, May 1993.
[13] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge, U.K.: Cambridge Univ. Press, 2011.
[14] L. Lai, Y. Liang, and H. V. Poor, "A unified framework for key agreement over wireless fading channels," *IEEE Trans. Inf. Forensics Secur.*, vol. 7, no. 2, pp. 480–490, Apr. 2012.
[15] A. Khisti, "Secret-key agreement over non-coherent block-fading channels with public discussion," *IEEE Trans. Inf. Theory*, vol. 62, no. 12, pp. 7164–7178, Dec. 2016.
[16] B. T. Quist and M. A. Jensen, "Maximization of the channel-based key establishment rate in MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 10, pp. 5565–5573, Oct. 2015.
[17] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2011.
[18] A. Grant, "Rayleigh fading multi-antenna channels," *EURASIP J. Adv. Signal Process.*, vol. 2002, no. 3, 2002, Art. no. 260208.
[19] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 2012.
[20] T. M. Cover, *Elements of Information Theory*. Wiley, 1999.
[21] O. Oyman, R. U. Nabar, H. Bolcskei, and A. J. Paulraj, "Characterizing the statistical properties of mutual information in MIMO channels," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2784–2795, Nov. 2003.



Shuo Wu (Member, IEEE) received the M.E. degree in control science and engineering from the Harbin Institute of Technology, Harbin, China, in 2011, the M.A. degree in mathematics from San Francisco State University, San Francisco, CA, USA, in 2016, and the Ph.D. degree in electrical engineering from the University of California, Riverside, CA, USA, in 2021. Since 2021, he has been with Goldman Sachs, USA. His research interests include wireless channel estimation, MIMO channel capacity analysis, and resource allocation.

Yingbo Hua (Fellow, IEEE) received the bachelor's degree from Southeast University, Nanjing, China, in 1982, the M.S. and Ph.D. degrees from Syracuse University, NY, USA, in 1983 and 1988, respectively. He joined the Faculty of the University of Melbourne, Melbourne, VIC, Australia, in 1990, and moved to the University of California, Riverside, CA, USA, in 2001, where he has been a Senior Full Professor since 2009. He has authored or coauthored more than 130 journal articles, 200 conference papers, 7 book chapters, 3 volumes of edited books and 3 patents in the field of signal processing and wireless communications. He was an Associate Editor, the Editor, the Guest Editor and/or a Senior Area Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE SIGNAL PROCESSING LETTERS, *Signal Processing*, IEEE SIGNAL PROCESSING MAGAZINE, IEEE JOURNAL OF SELECTED AREAS IN COMMUNICATIONS, and IEEE TRANSACTIONS ON SIGNAL AND INFORMATION PROCESSING OVER NETWORKS. He is a Former Member of the IEEE SPS Technical Committees on Signal Processing for Communications and Networking, and Sensor Array and Multichannel Processing. He was a General Co-Chair for the IEEE China SIP2015, Chair for IEEE Globe SIP 2018 Symposium on Signal Processing for Wireless Network Security, and Chair of the Steering Committee for IEEE WIRELESS COMMUNICATIONS LETTERS during 2020–2021. Dr. Hua is a Fellow of AAAS.