Secure Degree of Freedom of Wireless Networks Using Collaborative Pilots

Yingbo Hua ⬡, *Fellow, IEEE*, Qingpeng Liang ⬡, *Member, IEEE*,
and Md Saydur Rahman ⬡, *Graduate Student Member, IEEE*

Abstract—A wireless network of full-duplex nodes/users, using anti-eavesdropping channel estimation (ANECE) based on collaborative pilots, can yield a positive secure degree-of-freedom (SDoF) regardless of the number of antennas an eavesdropper may have. This article presents novel results on SDoF of ANECE by analyzing secret-key capacity (SKC) of each pair of nodes in a network of multiple collaborative nodes per channel coherence period. Each transmission session of ANECE has two phases: phase 1 is used for pilots, and phase 2 is used for random symbols. This results in two parts of SDoF of ANECE. Both lower and upper bounds on the SDoF of ANECE for any number of users are shown, and the conditions for the two bounds to meet are given. This leads to important discoveries, including: a) The phase-1 SDoF is the same for both multi-user ANECE and pair-wise ANECE while the former may require only a fraction of the number of time slots needed by the latter; b) For a three-user network, the phase-2 SDoF of all-user ANECE is generally larger than that of pair-wise ANECE; c) For a two-user network, a modified ANECE deploying square-shaped nonsingular pilot matrices yields a higher total SDoF than the original ANECE. The multi-user ANECE and the modified two-user ANECE shown in this article appear to be the best full-duplex schemes known today in terms of SDoF subject to each node using a given number of antennas for both transmitting and receiving.

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

I. INTRODUCTION

FULL-DUPLEX radio wireless networks have attracted much attention in recent years. The feasibility of full-duplex has been enabled by both signal processing algorithms and microwave circuit implementations, e.g., see [1] and [2]. It is widely known that full-duplex can be used to enhance both the spectral efficiency of a radio network and the security of the radio network against eavesdropping. This article is concerned with the latter.

In the field of wireless security, physical layer security is an important topic because of its ability to provide information-theoretic security between a pair of nodes or users even in the absence of a pre-existing secret key [3]. (We will use "node" and "user" as interchangeable terms.) However, numerous studies show that with only half-duplex radio nodes, the secrecy between the nodes goes to zero if the channel between the nodes lacks any reciprocal property and the eavesdropping channel is stronger than the legitimate channel, e.g., see [4] and [5]. A strong eavesdropping channel may be due to an eavesdropper (Eve) having more antennas than either the legitimate transmitter (Alice) or the legitimate receiver (Bob), or simply due to Alice being closer to Eve than to Bob.

With full-duplex radio coming over the horizon of future wireless networks, many researchers have explored the advantages of full-duplex for wireless security in a variety of settings [6], [7], [8], [9], [10], [11], [12], [13], [14], [15]. For example, a full-duplex destination node can receive a signal from a source node, and at the same time and same frequency transmit a jamming noise to reduce the signal-to-noise ratio (SNR) at Eve. Also, for example, a full-duplex multi-antenna relay can receive a signal from a source node, relay another signal to a destination node and transmit a jamming noise to interfere Eve, all at the same time and same frequency. The literature shows various problem formulations and corresponding solutions for power optimization at source, destination and/or relay to maximize a secrecy capacity subject to some partial or full knowledge of Eve's receive channel.

Full-duplex radio is also considered in [16], [17], [18], where secure degrees of freedom (SDoF) of their proposed schemes were analyzed. The work in [17] only considers a single antenna on each user and Eve. The SDoF of the schemes in [16] and [18] reduce to zero when the number of antennas on Eve is no less than the sum of the transmit antennas at Alice and Bob. Those findings have significantly underestimated the potentials of full-duplex in terms of SDoF, which will become clear later.

We like to mention that an early study of SDoF was shown in [19] and later in [5] for a half-duplex MIMO channel-model (CM) based secrecy capacity, where SDoF is shown to be zero if and only if the number of antennas on either Alice or Bob is no larger than that on Eve. More recently, SDoF is analyzed in [33] for a half-duplex MIMO source-model (SM) based secrecy capacity (also known as secret-key capacity), where (assuming

imperfect channel reciprocity) SDoF is shown to be zero if and only if the number of antennas on each of Alice and Bob is no larger than that on Eve.

It appeared not until the work [20] that researchers started to pay much attention to the impact of Eve who may have not only unlimited computing power but also a large number of antennas. The work [20] discovered a method that allows transmission of secret information between two or more full-duplex nodes regardless of the number of antennas on Eve even if there is no reciprocal property in the channels between the legitimate nodes. The method shown in [20] is called anti-eavesdropping channel estimation (ANECE), which uses collaborative pilots to exploit the condition that at the beginning of every channel coherence period, no node knows its receive channel state, and channel estimation (including carrier synchronization) must be conducted for any reliable transmission between users. ANECE is based on full-duplex, which is unlike the discriminatory channel estimation (DCE) methods based on half-duplex [21], [22], [23], and [24]. Without a channel reciprocal property between Alice and Bob, DCE is unable to yield a positive secrecy if the number of antennas on Eve is relatively large.

Further developments and understandings of ANECE are available in [25], [26], and [27]. Lower bounds on the degree of freedom (DoF) of secrecy capacity for a two-user ANECE were first reported in [25]. Optimal designs of the pilots required by a multi-user ANECE were developed in [26]. A total secrecy of ANECE in both channel estimation phase (phase 1) and random symbol transmission phase (phase 2) was studied in [27]. It was shown in [27] that a lower bound on the SDoF of ANECE is positive even if there is no channel reciprocity and Eve has an unlimited number of antennas. Furthermore, a related study of SDoF of two full-duplex nodes without initial knowledge of channel state information was studied in [28]. Despite the prior works on ANECE, many important questions remained. For example, the work in [28] only handles two full-duplex nodes instead of a network of more than two full-duplex nodes in each channel coherence period. The SDoF results shown in [27] are based on the optimal pilots designed in [26], instead of the most basic constraints on the collaborative pilots needed for ANECE.

This article aims to reveal further insights into SDoF of ANECE. We will adopt secret-key capacity (SKC) as a generalized secrecy capacity for both phases of ANECE. SKC is also known as the SM based secrecy capacity. Unlike the CM based secrecy capacity which allows public communication (for sharing a coding scheme for example) only before the transmission of a secret, SKC allows public communication both before and after the transmission of a secret [30]. While the "secret" transmitted using CM and the "secret" extracted from SM generally differ in "content", the secrecy in both cases can be measured in the same way if we ignore the cost and timing of operations. For example, any secret key of n bits extracted using SM can be used to keep any information of n bits in total secrecy even via public communications. So, it is meaningful to treat SKC as a generalization of secrecy capacity. In fact, a similar notion was used for a half-duplex SISO channel in [32] and recently for a half-duplex MIMO channel in [33]. In this article, we will use SKC and secrecy capacity interchangeably unless mentioned otherwise.

More specifically, we consider M multi-antenna (full-duplex) nodes/users cooperatively performing the two-phased ANECE within each channel coherence period $\mathcal{P}$ against a multi-antenna Eve. This is also referred to as multi-user ANECE. We aim to determine its pair-wise SDoF defined as $\mathtt{SDoF}_{i,j} \doteq \mathtt{DoF}(C_{\mathtt{key},i,j}) \doteq \lim_{\sigma^2 \to \infty} \frac{C_{\mathtt{key},i,j}}{\log \sigma^2}$, where $C_{\mathtt{key},i,j}$ is the SKC (in bits per $\mathcal{P}$) between node i and node j based on the observations from the execution of multi-user ANECE, and σ^2 is a nominal transmission power consumed by each node for multi-user ANECE. It will become clear that $\mathtt{SDoF}_{i,j} = \mathtt{SDoF}_{i,j}^{(1)} + \mathtt{SDoF}_{i,j}^{(2)}$ with $\mathtt{SDoF}_{i,j}^{(1)}$ and $\mathtt{SDoF}_{i,j}^{(2)}$ corresponding to phases 1 and 2 of multi-user ANECE. We will also study $\mathtt{SDoF}_{i,j}$ for a pair-wise ANECE where each $\mathcal{P}$ is shared orthogonally by all pairs of users among M users and each pair performs a two-user ANECE. Furthermore, we will study a modified two-user ANECE for a two-user network (i.e., only two users are considered within each $\mathcal{P}$). Different coherence periods are assumed to be independent.

Major results in this article are summarized in the properties at the end of each relevant section. The most significant ones are shown in Properties 6–9. In particular, we will show:

1) Assuming reciprocal channels between users and a full rank condition on pilots, $\mathtt{SDoF}_{i,j}^{(1)} = N_i N_j$ for both multi-user ANECE and pair-wise ANECE with N_i denoting the number of antennas on user i. This holds for all $M \geq 2$ and $N_E \geq 0$ with N_E denoting the number of antennas on Eve. But the pair-wise ANECE requires $\frac{M}{2}$ times more time-slots for phase-1 transmissions.

2) For a three-user symmetric network where $M = 3$ and $N_i = N$ for all i, $\mathtt{SDoF}_{i,j}^{(2)} = 0$ for pair-wise ANECE if $N_E \geq 2N$ while $\mathtt{SDoF}_{i,j}^{(2)} > 0$ for multi-user ANECE if $1 \leq K_2 < 2N$ and (even) $N_E \geq 3N$. Here K_2 is the number of time-slots used for phase-2 transmissions.

3) For a two-user network where $M = 2$ and $N_1 \leq N_2$, $\mathtt{SDoF}_{1,2}^{(2)} = 2N_1^2$ for the original two-user ANECE (and also for the scheme in [28]) if $N_E \geq N_1 + N_2$ and $K \geq N_1 + N_2$, but $\mathtt{SDoF}_{1,2}^{(2)} = N_1(N_1 + N_2)$ for the modified two-user ANECE under the same conditions. Here K is the total number of time-slots used for both phase-1 and phase-2 transmissions.

These results are novel and significant. (The notations used later for SDoFs will be slightly different for better referencing.)

The rest of the article is organized as follows. In Section II, we review the meanings of DoF and SDoF, and introduce several lemmas that are important throughout this article. The notion of pair-wise SKC for multi-user ANECE is defined. In Section III, we show the key features of multi-user ANECE, including the structure of the collaborative pilots and random symbols used by multi-user ANECE. The effects of the collaborative pilots on channel estimation at users and Eve are highlighted. In Section IV, we show how the pair-wise SKC of multi-user ANECE can be decomposed into two components: one for phase 1 and the other for phase 2. These two components will be treated separately in the sequel

of the article. In Section V, we derive the pair-wise SDoF for phase 1 of multi-user ANECE subject to reciprocal random channel parameters between users, which leads to a result more general than one in [27]. In Section VI, the pair-wise SDoF for phase 2 of multi-user ANECE is shown in terms of its lower and upper bounds. The conditions for the bounds to meet are also given. In Section VII, we compare all-user ANECE with a pair-wise ANECE for a network of $M \geq 3$ users. We show that all-user ANECE has advantages over pair-wise ANECE in both phases 1 and 2. In Section VIII, we propose and analyse a modified two-user ANECE for a two-user network, for which each user applies a square-shaped pilot matrix. In Section IX, we show that the SDoF of the modified two-user ANECE is generally larger than that of the original two-user ANECE. Section X concludes the article.

Notations: The superscripts ^{T}, * and ^{H} denote transpose, conjugate and conjugate transpose respectively while $*$ as an entry in a matrix denotes a quantity whose specific form is not important for discussion. $\mathbb{C}^{m \times n}$ denotes the space of $m \times n$ complex matrices. $\mathbb{E}$, $\otimes$ and $\oplus$ are respectively expectation, Kronecker product and (element-wise) exclusive-OR. $(a)^+ = \max(a, 0)$. $|\mathbf{A}|$ is the determinant of the matrix $\mathbf{A}$. All matrices and vectors are represented respectively by boldface upper cases and boldface lower cases. $\doteq$ stresses "equal by definition". $\mathcal{CN}(\mathbf{m}, \mathbf{R})$ denotes the circular complex Gaussian probability-density-function (PDF) with mean vector $\mathbf{m}$ and covariance matrix $\mathbf{R}$. $\mathbb{I}(A; B|C)$ denotes the mutual information between A and B, conditional on C. $h(A|B)$ is the differential entropy of A conditional on B (or entropy of A conditional on B if A is discrete). And $(\cdot)$, $[\cdot]$ and $\{\cdot\}$ are used interchangeably.

II. PRELIMINARIES

A. Degree of Freedom and Secure Degree of Freedom

The degree of freedom (DoF) of a function $g(\sigma^2)$ of σ^2 is said to be d relative to $\log \sigma^2$ if for large σ^2, $g(\sigma^2) \approx d \log \sigma^2 + c$ where both d and c are invariant to σ^2. In this case, we also write $\text{DoF}(g(\sigma^2)) = d$. The approximation "$\approx$" we use in this article will be always such that it does not affect the DoF. If a scheme has the secrecy capacity C_s and $\text{DoF}(C_s) = d$, we say that this scheme has the secure degree of freedom (SDoF) equal to d.

The following two lemmas and corollary will be used frequently.

Lemma 1: Let $\mathbf{R}$ be an $n \times n$ complex positive-definite matrix dependent on σ^2. If for large σ^2 the ith eigenvalue λ_i of $\mathbf{R}$ can be written as $\eta_i \sigma^2 + a_i$ for $1 \leq i \leq m$, and as b_i for $m + 1 \leq i \leq n$, where $\eta_i > 0$, a_i and $b_i > 0$ are all invariant to σ^2, then $\text{DoF}(\log_2 |\mathbf{R}|) = m$.

Proof: It is known that $\log_2 |\mathbf{R}| = \log_2 \prod_{i=1}^n \lambda_i = \sum_{i=1}^n \log_2 \lambda_i$. As σ^2 increases, $\log_2 |\mathbf{R}|$ converges to $\sum_{i=1}^m \log_2(\eta_i \sigma^2 + a_i) + \sum_{i=m+1}^n \log b_i = m \log_2 \sigma^2 + \sum_{i=1}^m \log_2 (1 + \frac{a_i}{\eta_i \sigma^2}) + \sum_{i=1}^m \log_2 \eta_i + \sum_{i=m+1}^n \log_2 b_i$ which further converges to $m \log_2 \sigma^2 + \sum_{i=1}^m \log_2 \eta_i + \sum_{i=m+1}^n \log_2 b_i$. Hence $\text{DoF}(\log_2 |\mathbf{R}|) = m$. $\square$

Corollary 1: Let $\mathbf{R}$ be a (continuous) random covariance matrix. If with probability one $\mathbf{R}$ converges to $\sigma^2 \mathbf{R}_m + \mathbf{I}_n$ as σ^2 increases to ∞ where $\mathbf{R}_m$ has the rank $m \leq n$, then $\text{DoF}(\mathbb{E}\{\log_2 |\mathbf{R}|\}) = m$.

Proof: This follows from Lemma 1. $\square$

Lemma 2: Let $\mathbf{Y} = \sigma \mathbf{H} \mathbf{X} + \mathbf{W} \in \mathbb{C}^{m \times k}$ where $\mathbf{H} \in \mathbb{C}^{m \times n}$, $\mathbf{X} \in \mathbb{C}^{n \times k}$ and $\mathbf{W} \in \mathbb{C}^{m \times k}$ are independent of each other and consist independent and identically distributed (i.i.d.) $\mathcal{CN}(0, 1)$ entries. Then, relative to $\log \sigma^2$, $\text{DoF}(h(\mathbf{Y}|\mathbf{H})) = \min(m, n)k$ and $\text{DoF}(h(\mathbf{Y}|\mathbf{X})) = \min(n, k)m$.

Proof: Let $\mathbf{y} = vec(\mathbf{Y})$, $\mathbf{x} = vec(\mathbf{X})$ and $\mathbf{w} = vec(\mathbf{W})$. Then $\mathbf{y} = \sigma(\mathbf{I}_k \otimes \mathbf{H})\mathbf{x} + \mathbf{w}$, which has the conditional PDF $f(\mathbf{y}|\mathbf{H}) = \mathcal{CN}(0, \mathbf{R})$ with $\mathbf{R} = \sigma^2(\mathbf{I}_k \otimes \mathbf{H}\mathbf{H}^H) + \mathbf{I}_{mk}$. It follows that $h(\mathbf{Y}|\mathbf{H}) = \mathbb{E}\{\log \frac{1}{f(\mathbf{y}|\mathbf{H})}\} = \mathbb{E}\{\log(\pi^{mk}|\mathbf{R}| \exp(\mathbf{y}^H \mathbf{R}^{-1} \mathbf{y}))\} = \mathbb{E}\{\log(e^{mk} \pi^{mk} |\mathbf{R}|)\}$. It follows from Corollary 1 that $\text{DoF}(h(\mathbf{Y}|\mathbf{H})) = \min(m, n)k$. Also note $\text{DoF}(h(\mathbf{Y}|\mathbf{X})) = \text{DoF}(h(\mathbf{Y}^T|\mathbf{X})) = \min(k, n)m$. $\square$

The next lemma will be useful for cases where the Gaussian assumption in Lemma 2 is not available.

Lemma 3: Let $\mathbf{y} = \sigma \mathbf{x} + \mathbf{w} \in \mathbb{C}^{n \times 1}$ where $\mathbf{w}$ is noise and all entries in $\mathbf{x}$ can be asymptotically determined by minimal d entries of $\mathbf{y}$ as σ^2 increases. Then relative to $\log \sigma^2$, $\text{DoF}(h(\mathbf{y})) = d$ (regardless of the PDF of $\mathbf{y}$).

Proof: For any real-valued random variable X with a PDF $f_X(x)$, it is known that another random variable $Y = \sigma X$ has the PDF $f_Y(y) = \frac{1}{\sigma} f_X(\frac{y}{\sigma})$. Then $h(Y) = \int f_Y(y) \log \frac{1}{f_Y(y)} dy = \log \sigma + h(X)$. For a complex-valued vector $\mathbf{y}$, we now assume that its first d (complex) elements are the minimum number of elements needed to determine all elements in $\mathbf{x}$ at high power σ^2. Then we have $h(\mathbf{y}) = h(y_1) + h(y_2|y_1) + \cdots + h(y_d|y_{d-1}, \cdots, y_1) + c$ where y_i is the ith element of $\mathbf{y}$, and $c = h(y_{d+1}|y_d, \cdots, y_1) + \cdots + h(y_n|y_{n-1}, \cdots, y_1)$ which is invariant to σ^2 for large σ^2. Note that since x_{d+m} can be determined by $y_1, \cdots, y_d$ at large σ^2, then as σ^2 increases, $h(y_{d+m}|y_1, \cdots, y_d)$ converges to $h(w_{d+m}|y_1, \cdots, y_d)$ which is invariant to σ^2. On the other hand, for $i \leq d$, since no x_i (real or imaginary part) can be determined by $y_{i-1}, \cdots, y_1$ for large σ^2, it follows that $h(y_i|y_{i-1}, \cdots, y_1)$ converges to $h(\sigma e_i)$ for large σ^2 where e_i is the nonzero estimation error of x_i at $\sigma^2 = \infty$. Furthermore, $h(\sigma e_i) = 2 \log \sigma + h(\Re(e_i)) + h(\Im(e_i)|\Re(e_i))$. Hence, $\text{DoF}(h(\mathbf{y})) = d$. $\square$

B. Pair-Wise Secret-Key Capacities

Consider M users where user i has the random data set $\mathcal{Y}_i$ with $i = 1, \cdots, M$, and Eve who has the random data set $\mathcal{Y}_E$. We will be interested in the secret-key capacity (SKC) of each pair of users against Eve, which is also called pair-wise SKC.

Lemma 4: The pair-wise SKC between user i and user j for all $i \neq j$ against Eve in bits per independent realization of $\{\mathcal{Y}_1, \cdots, \mathcal{Y}_M, \mathcal{Y}_E\}$, denoted by $C_{\text{key},i,j}$, satisfies

$$C_{\text{key},i,j} \geq \max(C_{\text{low},i,j}, C_{\text{low},j,i}) \tag{1}$$

with $C_{\text{low},i,j} = \mathbb{I}(\mathcal{Y}_i; \mathcal{Y}_j) - \mathbb{I}(\mathcal{Y}_i; \mathcal{Y}_E)$ and

$$C_{\text{key},i,j} \leq C_{\text{up},i,j} \doteq \mathbb{I}(\mathcal{Y}_i; \mathcal{Y}_j|\mathcal{Y}_E). \tag{2}$$

Proof: The following proof follows the same principle as in section 4.2.1 in [29]. First we assume that all entries in $\mathcal{Y}_1, \cdots, \mathcal{Y}_M$ and $\mathcal{Y}_E$ are discrete (after quantization), or equivalently each of the sets $\mathcal{Y}_1, \cdots, \mathcal{Y}_M$ and $\mathcal{Y}_E$ becomes a finite multidimensional discrete symbol belonging to a common set $\mathbb{Y}$. For each i, user i generates a uniformly random symbol $\mathcal{U}_i$ from $\mathbb{Y}$ and transmits $\bar{\mathcal{Y}}_i = \mathcal{Y}_i \oplus \mathcal{U}_i$ via public channel. These transmissions constitute M broadcast wiretap channels, i.e., from each user to all other users. It follows that the secrecy capacity from user i to user j against Eve in this case is

$$\mathbb{I}[\mathcal{U}_i; \mathcal{Y}_j, \bar{\mathcal{Y}}_1, \cdots, \bar{\mathcal{Y}}_M] - \mathbb{I}[\mathcal{U}_i; \mathcal{Y}_E, \bar{\mathcal{Y}}_1, \cdots, \bar{\mathcal{Y}}_M]$$
$$= h[\mathcal{U}_i|\mathcal{Y}_E, \bar{\mathcal{Y}}_1, \cdots, \bar{\mathcal{Y}}_M] - h[\mathcal{U}_i|\mathcal{Y}_j, \bar{\mathcal{Y}}_1, \cdots, \bar{\mathcal{Y}}_M] \quad (3)$$

Since the uniformly random $\mathcal{U}_i$ is independent of $\bar{\mathcal{y}}_j$ for all $j \neq i$, and $\bar{\mathcal{Y}}_j$ is independent of $\mathcal{Y}_E$, $\mathcal{Y}_i$ and $\bar{\mathcal{Y}}_i$ for all $i \neq j$, (3) is equivalent to

$$h[\mathcal{U}_i|\mathcal{Y}_E, \bar{\mathcal{Y}}_i] - h[\mathcal{U}_i|\mathcal{Y}_j, \bar{\mathcal{Y}}_i] = h[\mathcal{Y}_i|\mathcal{Y}_E] - h[\mathcal{Y}_i|\mathcal{Y}_j]$$
$$= \mathbb{I}(\mathcal{Y}_i; \mathcal{Y}_j) - \mathbb{I}(\mathcal{Y}_i; \mathcal{Y}_E) \quad (4)$$

where we have used $h[\mathcal{U}_i|\mathcal{Y}_E, \bar{\mathcal{Y}}_i] = h[\mathcal{U}_i, \bar{\mathcal{Y}}_i|\mathcal{Y}_E] - h[\bar{\mathcal{Y}}_i|\mathcal{Y}_E] = h[\mathcal{U}_i|\mathcal{Y}_E] + h[\bar{\mathcal{Y}}_i|\mathcal{U}_i, \mathcal{Y}_E] - h[\bar{\mathcal{Y}}_i|\mathcal{Y}_E] = h[\bar{\mathcal{Y}}_i|\mathcal{U}_i, \mathcal{Y}_E] = h[\mathcal{Y}_i|\mathcal{Y}_E]$. Since $C_{\text{key},i,j} = C_{\text{key},j,i}$, the above result leads to (1). For continuous $\mathcal{Y}_1, \cdots, \mathcal{Y}_M$ and $\mathcal{Y}_E$, (1) then follows from the generalized definition of mutual information as shown in [31]. Finally, (2) follows directly from the upper bound for the two-user case as shown in [30] and [29]. $\square$

Note that a key generated by a pair of users may be correlated with a key generated by another pair of users. Even $\mathcal{U}_i$ and $\mathcal{U}_j$ become correlated when conditioned on $\bar{\mathcal{Y}}_i$ and $\bar{\mathcal{Y}}_j$. For example, $\mathbb{I}(\mathcal{U}_i; \mathcal{U}_j|\bar{\mathcal{Y}}_i, \bar{\mathcal{Y}}_j) = h(\mathcal{U}_i|\bar{\mathcal{Y}}_i, \bar{\mathcal{Y}}_j) - h(\mathcal{U}_i|\mathcal{U}_j, \bar{\mathcal{Y}}_i, \bar{\mathcal{Y}}_j) = h(\mathcal{U}_i|\bar{\mathcal{Y}}_i) - h(\mathcal{U}_i|\mathcal{Y}_j, \bar{\mathcal{Y}}_i) = h(\mathcal{Y}_i) - h(\mathcal{Y}_i|\mathcal{Y}_j) = \mathbb{I}(\mathcal{Y}_i; \mathcal{Y}_j)$. However, as discussed later, some components in the pair-wise secret keys are independent of each other.

C. Secret-Key Capacity Versus Secrecy Capacity

For wiretap channel model (a notion for CM based secrecy) where a secret information is directly transmitted from a source to a destination, there is what is commonly called secrecy capacity. However, given a secret key $\mathcal{K}$ of n bits between the source and the destination, the source can send a secret information $\mathcal{I}$ of n bits by sending $\mathcal{I} \oplus \mathcal{K}$ via public channel, and this $\mathcal{I}$ is then received by the destination in complete secrecy. In other words, a secret-key capacity can be directly translated into a secrecy capacity if additional operation or communication is allowed. For this reason, we treat "secret-key capacity" and "secrecy capacity" interchangeably unless mentioned otherwise.

III. DATA MODELS OF MULTI-USER ANECE USING COLLABORATIVE PILOTS

We consider a network of M cooperative full-duplex nodes against Eve in *each channel coherence period*. All these full-duplex nodes are assumed to be ideal full-duplex. The power of the residual self-interference of a practical full-duplex radio is in general proportional to its transmitted power. A practical full-duplex radio can be modelled as an ideal full-duplex radio

as long as the transmitted power is no larger than a threshold. This threshold is governed by a target noise level and a residual self-interference power gain, which certainly affects the achievable distance of full-duplex communication. In order for the SDoF theory shown in this article to be applicable in a given application, that threshold needs to be relatively large. For implementation issues of full-duplex, see [1].

The numbers of antennas on the M full-duplex nodes and Eve are denoted by $N_1, N_2, \cdots, N_M$ and N_E respectively. To simplify many of the mathematical expressions, we assume that node i has the same number N_i of transmit antennas as receive antennas. Note that Eve here could represent multiple colluding eavesdroppers with their combined number of antennas equal to N_E.

The multi-user ANECE is such that every node transmits its packet (on the same carrier frequency) concurrently with all other nodes. The required precision for the concurrence is only at the symbol level. Each packet within a channel coherence period consists of a pilot matrix and a random symbol matrix, e.g., node i transmits the $N_i \times K_1$ pilot matrix $\mathbf{P}_i$ in phase 1 and the $N_i \times K_2$ symbol matrix $\mathbf{X}_i$ in phase 2. The pilot matrix and the symbol matrix transmitted from the same node can and should be transmitted sequentially in synch with each other up to the carrier frequency level. So, an estimated channel matrix based on a pilot matrix from node j can be used for coherent detection of the symbol matrix from node j. But for the purpose of this article, we do not need any explicit coherent detection at each node.

Define $\mathbf{P} = [\mathbf{P}_1^H, \cdots, \mathbf{P}_M^H]^H$ and $\mathbf{X} = [\mathbf{X}_1^H, \cdots, \mathbf{X}_M^H]^H$. Hence, the transmitted signals from all nodes can be expressed as

$$[\mathbf{P}, \mathbf{X}] = \begin{bmatrix} \mathbf{P}_1, & \mathbf{X}_1 \\ \cdots, & \cdots \\ \mathbf{P}_M, & \mathbf{X}_M \end{bmatrix} \in \mathbb{C}^{N_T \times (K_1 + K_2)} \quad (5)$$

where $N_T \doteq \sum_{i=1}^M N_i$, each row of (5) corresponds to an antenna among the nodes, and each column of (5) corresponds to a sampling interval. Furthermore, we let $\mathbf{P}_{(i)}$ be $\mathbf{P}$ without $\mathbf{P}_i$.

Assumption 1: The collaborative pilot matrices for all-user ANECE (with total M users) meet the following conditions: $\text{rank}(\mathbf{P}_i) = N_i$, $\text{rank}(\mathbf{P}_{(i)}) = N_T - N_i$ and $\text{rank}(\mathbf{P}) = \max_i \text{rank}(\mathbf{P}_{(i)}) = N_T - N_{\min}$ where $N_{\min} \doteq \min_i N_i$. Consequently, we need $K_1 \geq N_T - N_{\min}$.

A simple way to construct such a $\mathbf{P}$ is to first form a full-rank $N_T \times N_T$ matrix and then remove any $N_{\min}$ columns.

Assumption 2: All entries in $\mathbf{X}$ are i.i.d. $\mathcal{CN}(0, 1)$.

Assumption 1 is the most basic requirement on the collaborative pilots to achieve the maximum ambiguity of channel estimation at Eve, which was originally proposed in Appendix E in [20].

In phase 1, the signal received by node i is

$$\mathbf{Y}_i^{(1)} = \sigma \sum_{j \neq i} \mathbf{H}_{i,j} \mathbf{P}_j + \mathbf{W}_i^{(1)} = \sigma \mathbf{H}_i \mathbf{P}_{(i)} + \mathbf{W}_i^{(1)} \quad (6)$$

where $\mathbf{H}_{i,j}$ is the $N_i \times N_j$ channel matrix from node j to node i, and $\mathbf{H}_i$ is the $N_i \times (N_T - N_i)$ matrix from the horizontal stack of $\mathbf{H}_{i,j}$ for all $j \neq i$. At a high SNR or a high power σ^2 in phase 1, node i for every i can estimate $\mathbf{H}_i$ reliably. SNR in phase 1 is the ratio of the pilot power over the noise power.

Also in phase 1, Eve receives

$$\mathbf{Y}_E^{(1)} = \sigma \sum_{i=1}^{M} \mathbf{H}_{E,i} \mathbf{P}_i + \mathbf{W}_E^{(1)} = \sigma \mathbf{H}_E \mathbf{P} + \mathbf{W}_E^{(1)} \quad (7)$$

where $\mathbf{H}_{E,i}$ is the channel matrix from node i to Eve, and $\mathbf{H}_E = [\mathbf{H}_{E,1}, \cdots, \mathbf{H}_{E,M}]$. Since $\mathbf{P}$ does not have a full row rank, Eve is unable to obtain a consistent estimate of $\mathbf{H}_E$. But she is able to find a part of $\mathbf{H}_E$ as explain below.

Since the $N_T \times K_1$ matrix $\mathbf{P}$ has the rank $N_T - N_{\min}$, there exist a $N_T \times (N_T - N_{\min})$ orthonormal matrix $\mathbf{Q}_P$ and a $(N_T - N_{\min}) \times K_1$ matrix $\mathbf{R}_P$ (of rank $N_T - N_{\min}$) such that $\mathbf{P} = \mathbf{Q}_P \mathbf{R}_P$. This follows from the standard QR decomposition. There is also a $N_T \times N_{\min}$ orthonormal complement matrix $\mathbf{Q}_{P,\perp}$ such that $\mathbf{Q} \doteq [\mathbf{Q}_P, \mathbf{Q}_{P,\perp}]$ is unitary. Let $\mathbf{H}'_E \doteq \mathbf{H}_E \mathbf{Q} = [\mathbf{H}_{E,P}, \mathbf{H}_{E,P,\perp}]$ where $\mathbf{H}_{E,P} = \mathbf{H}_E \mathbf{Q}_P$ and $\mathbf{H}_{E,P,\perp} = \mathbf{H}_E \mathbf{Q}_{P,\perp}$. Then, (7) becomes

$$\mathbf{Y}_E^{(1)} = \sigma \mathbf{H}_E \mathbf{Q} \mathbf{Q}^H \mathbf{P} + \mathbf{W}_E^{(1)} = \sigma \mathbf{H}_{E,P} \mathbf{R}_P + \mathbf{W}_E^{(1)}. \quad (8)$$

So at a high SNR (or high power), Eve is able to estimate $\mathbf{H}_{E,P}$ reliably.

Note that given $\mathbf{Y}_E^{(1)}$ at high power, we know that $\mathbf{H}_E \approx \bar{\mathbf{H}}_E + \mathbf{T}_E \mathbf{N}_E$ where $\bar{\mathbf{H}}_E$ is the minimum-norm solution of $\mathbf{H}_E$ to $\mathbf{Y}_E^{(1)} = \sigma \mathbf{H}_E \mathbf{P}$, the row span of $\mathbf{N}_E$ is the left null space of $\mathbf{P}$, and $\mathbf{T}_E \in \mathbb{C}^{N_E \times N_{\min}}$ is "arbitrary". But $\mathbf{T}_E$ obtained from $\mathbf{Y}_E^{(1)}$ along with the prior i.i.d. statistics of $\mathbf{H}_E$ is not fully arbitrary. Yet, no entry of $\mathbf{T}_E$ can be consistently estimated from other entries of $\mathbf{T}_E$ when only $\mathbf{Y}_E^{(1)}$ is given. In this case, the entries of $\mathbf{T}_E$ are said to have full freedoms or simply called "arbitrary". The connection between $\mathbf{T}_E \in \mathbb{C}^{N_E \times N_{\min}}$ and $\mathbf{H}_{E,P,\perp} \in \mathbb{C}^{N_E \times N_{\min}}$ at high power is simply $\bar{\mathbf{H}}_E + \mathbf{T}_E \mathbf{N}_E = [\mathbf{H}_{E,P}, \mathbf{H}_{E,P,\perp}] \mathbf{Q}^H$. For this reason, $\mathbf{T}_E$ and $\mathbf{H}_{E,P,\perp}$ have the same impact on DoF despite the fact that the entries of $\mathbf{T}_E$ are not i.i.d. in general. This property will be used to simplify some of the analyses shown later.

Assumption 3: $\mathbf{H}_{i,j} = \mathbf{H}_{j,i}^T$, and all entries in $\mathbf{H}_{i,j}$ for all $i < j$ over different coherence periods are i.i.d. $\mathcal{CN}(0,1)$ and independent of $\mathbf{H}_E$.

Assumption 4: All entries of $\mathbf{H}_E$ are i.i.d. $\mathcal{CN}(0,1)$.

Assumption 5: The entries of all noise matrices at users are i.i.d. $\mathcal{CN}(0,1)$, and the entries of all noise matrices at Eve are i.i.d. $\mathcal{CN}(0,\omega^2)$.

In fact, the value of ω has zero effect on the SDoF analysis. We can also let $\omega = 1$ for convenience.

Because of Assumption 4, all entries in $\mathbf{H}'_E$ are also i.i.d. $\mathcal{CN}(0,1)$. In this case, all entries in the $N_E \times N_{\min}$ matrix $\mathbf{H}_{E,P,\perp}$ remain i.i.d. $\mathcal{CN}(0,1)$ even after $\mathbf{Y}_E^{(1)}$ is available.

In phase 2, node i for every i sends out the $N_i \times K_2$ symbol matrix $\mathbf{X}_i$. Then node i for every i receives

$$\mathbf{Y}_i^{(2)} = \sigma \sum_{j \neq i} \mathbf{H}_{i,j} \mathbf{X}_j + \mathbf{W}_i^{(2)} = \mathbf{H}_i \mathbf{X}_{(i)} + \mathbf{W}_i^{(2)} \quad (9)$$

with $\mathbf{X}_{(i)}$ being the vertical stack of $\mathbf{X}_j$ for $j \neq i$, and Eve receives

$$\mathbf{Y}_E^{(2)} = \sigma \sum_{j=1}^{M} \mathbf{H}_{E,j} \mathbf{X}_j + \mathbf{W}_E^{(2)} = \mathbf{H}_E \mathbf{X} + \mathbf{W}_E^{(2)} \quad (10)$$

where $\mathbf{X}$ is the $N_T \times K_2$ matrix from the vertical stack of $\mathbf{X}_j$ for all j. Let

$$\mathbf{X}' \doteq \mathbf{Q}^H \mathbf{X} = \begin{bmatrix} \mathbf{X}_P \\ \mathbf{X}_{P,\perp} \end{bmatrix} \quad (11)$$

where $\mathbf{X}_P = \mathbf{Q}_P^H \mathbf{X}$ and $\mathbf{X}_{P,\perp} = \mathbf{Q}_{P,\perp}^H \mathbf{X}$. Then (10) becomes

$$\mathbf{Y}_E^{(2)} = \sigma \mathbf{H}_{E,P} \mathbf{X}_P + \sigma \mathbf{H}_{E,P,\perp} \mathbf{X}_{P,\perp} + \mathbf{W}_E^{(2)}. \quad (12)$$

Because of Assumption 2, all entries in $\mathbf{X}'$ of (11) are also i.i.d. $\mathcal{CN}(0,1)$.

As a summary, we have:

Property 1: Under Assumptions 1 and 4 and a high SNR in phase 1, $\mathbf{Y}_i^{(1)}$ shown in (6) yields $\mathbf{H}_i$ uniquely, and $\mathbf{Y}_E^{(1)}$ shown in (7), or equivalently (8), implies a unique $\mathbf{H}_{E,P} \in \mathbb{C}^{N_E \times (N_T - N_{\min})}$ but leaves all entries of $\mathbf{H}_{E,P,\perp} \in \mathbb{C}^{N_E \times N_{\min}}$ as i.i.d. $\mathcal{CN}(0,1)$.

IV. DECOMPOSITION OF SECRET-KEY CAPACITY

Let $\mathcal{Y}_i \doteq \{\mathbf{X}_i, \mathbf{Y}_i^{(1)}, \mathbf{Y}_i^{(2)}\}$ and $\mathcal{Y}_E \doteq \{\mathbf{Y}_E^{(1)}, \mathbf{Y}_E^{(2)}\}$ be the data sets collected by user i and Eve after the two-phase processes of multi-user ANECE. We show next that each pair-wise secret-key capacity $C_{\text{key},i,j}$ based on the data sets at users i and j from multi-user ANECE can be decomposed into two components: one for phase 1 and the other for phase 2, i.e., $C_{\text{key},i,j} = C_{\text{key},i,j}^{(1)} + C_{\text{key},i,j}^{(2)}$.

A. Lower Bound

It follows from Lemma 4 that

$$C_{\text{key},i,j} \geq \mathbb{I}(\mathcal{Y}_i; \mathcal{Y}_j) - \min(C_{i,E}, C_{j,E}) \quad (13)$$

with $C_{i,E} = \mathbb{I}(\mathcal{Y}_i; \mathcal{Y}_E)$ and

$$\begin{aligned}
\mathbb{I}(\mathcal{Y}_i; \mathcal{Y}_j) &= \mathbb{I}(\mathbf{Y}_i^{(1)}; \mathcal{Y}_j) + \mathbb{I}(\mathbf{X}_i, \mathbf{Y}_i^{(2)}; \mathcal{Y}_j | \mathbf{Y}_i^{(1)}) \\
&= \mathbb{I}(\mathbf{Y}_i^{(1)}; \mathbf{Y}_j^{(1)}) + \mathbb{I}(\mathbf{Y}_i^{(1)}; \mathbf{X}_j, \mathbf{Y}_j^{(2)} | \mathbf{Y}_j^{(1)}) \\
&\quad + \mathbb{I}(\mathbf{X}_i, \mathbf{Y}_i^{(2)}; \mathbf{Y}_j^{(1)} | \mathbf{Y}_i^{(1)}) \\
&\quad + \mathbb{I}(\mathbf{X}_i, \mathbf{Y}_i^{(2)}; \mathbf{X}_j, \mathbf{Y}_j^{(2)} | \mathbf{Y}_i^{(1)}, \mathbf{Y}_j^{(1)}) \\
&\approx \mathbb{I}(\mathbf{Y}_i^{(1)}; \mathbf{Y}_j^{(1)}) + \mathbb{I}(\mathbf{X}_i, \mathbf{Y}_i^{(2)}; \mathbf{X}_j, \mathbf{Y}_j^{(2)} | \mathbf{H}_i, \mathbf{H}_j).
\end{aligned} \quad (14)$$

The above approximation holds under a high power in phase 1, which allows $\mathbf{Y}_i^{(1)}$ to uniquely determine $\mathbf{H}_i$ for all i and hence implies $\mathbb{I}(\mathbf{Y}_i^{(1)}; \mathbf{X}_j, \mathbf{Y}_j^{(2)} | \mathbf{Y}_j^{(1)}) \approx 0$ for all $i \neq j$. The above and all other similar approximations used in this article do not affect the DoF analyses. Because of $\mathbb{I}(\mathbf{Y}_i^{(1)}; \mathcal{Y}_E) = 0$ and $\mathbb{I}(\mathbf{X}_i, \mathbf{Y}_i^{(2)}; \mathbf{Y}_j^{(1)} | \mathbf{Y}_i^{(1)}) = 0$ (which is due to the independence between users' channel matrices and Eve's channel matrices), we have

$$\begin{aligned}
C_{i,E} &= \mathbb{I}(\mathbf{X}_i, \mathbf{Y}_i^{(2)}; \mathcal{Y}_E | \mathbf{Y}_i^{(1)}) = \mathbb{I}(\mathbf{X}_i, \mathbf{Y}_i^{(2)}; \mathbf{Y}_E^{(2)} | \mathbf{Y}_i^{(1)}, \mathbf{Y}_E^{(1)}) \\
&\approx \mathbb{I}(\mathbf{X}_i, \mathbf{Y}_i^{(2)}; \mathbf{Y}_E^{(2)} | \mathbf{H}_i, \mathbf{H}_{E,P})
\end{aligned} \quad (15)$$

where the approximation is due to Property 1.

We can write from (13), (14) and (15) that

$$C_{\text{key},i,j} \geq C_{\text{key},i,j}^{(1)} + \max(C_{\text{key},i,j,\text{low}}^{(2)}, C_{\text{key},j,i,\text{low}}^{(2)}) \quad (16)$$

where the first term is the pair-wise SKC achievable based on the data collected in phase 1 of all-user ANECE and the second term is the pair-wise SKC achievable based on the data collected in phase 2 of all-user ANECE. Here $C_{\text{key},i,j}^{(1)} = \mathbb{I}(\mathbf{Y}_i^{(1)}; \mathbf{Y}_j^{(1)})$ and

$$C_{\text{key},i,j,\text{low}}^{(2)} = C_{i,j} - C_{i,E}, \quad (17)$$

with $C_{i,j} \approx \mathbb{I}(\mathbf{X}_i, \mathbf{Y}_i^{(2)}; \mathbf{X}_j, \mathbf{Y}_j^{(2)} | \mathbf{H}_i, \mathbf{H}_j)$.

B. Upper Bound

It also follows from Lemma 4 that with high power in phase 1,

$$\begin{aligned} C_{\text{key},i,j} &\leq \mathbb{I}(\mathcal{Y}_i; \mathcal{Y}_j | \mathcal{Y}_E) \\ &= \mathbb{I}(\mathbf{Y}_i^{(1)}; \mathcal{Y}_j | \mathcal{Y}_E) + \mathbb{I}(\mathbf{X}_i, \mathbf{Y}_i^{(2)}; \mathcal{Y}_j | \mathbf{Y}_i^{(1)}, \mathcal{Y}_E) \\ &\approx C_{\text{key},i,j}^{(1)} + C_{\text{key},i,j,\text{up}}^{(2)} \end{aligned} \quad (18)$$

with

$$C_{\text{key},i,j,\text{up}}^{(2)} = \mathbb{I}(\mathbf{X}_i, \mathbf{Y}_i^{(2)}; \mathbf{X}_j, \mathbf{Y}_j^{(2)} | \mathbf{Y}_i^{(1)}, \mathbf{Y}_j^{(1)}, \mathcal{Y}_E). \quad (19)$$

Note that the phase 1 component $C_{\text{key},i,j}^{(1)}$ of the SKC is exact while the phase 2 component of the SKC is lower bounded by $C_{\text{key},i,j,\text{low}}^{(2)}$ and upper bounded by $C_{\text{key},i,j,\text{up}}^{(2)}$.

C. Summary

Property 2: The pair-wise SKC in bits per channel coherence period based on the data collected from multi-user ANECE consists of two components. The first is $C_{\text{key},i,j}^{(1)} = \mathbb{I}(\mathbf{Y}_i^{(1)}; \mathbf{Y}_j^{(1)})$, and the second is lower bounded by $C_{\text{key},i,j,\text{low}}^{(2)}$ shown in (17) and upper bounded by $C_{\text{key},i,j,\text{up}}^{(2)}$ shown in (19).

We will next analyse the DoFs of $C_{\text{key},i,j}^{(1)}$ and $C_{\text{key},i,j,\text{low}}^{(2)}$ in two separate sections. The analysis of $C_{\text{key},i,j,\text{up}}^{(2)}$ is done in Appendix.

V. PHASE-1 SKC OF ALL-USER ANECE

In this section, we will derive the DoF of $C_{\text{key},i,j}^{(1)}$. First we can write

$$\begin{aligned} C_{\text{key},i,j}^{(1)} &\doteq \mathbb{I}(\mathbf{Y}_i^{(1)}; \mathbf{Y}_j^{(1)}) \\ &= h(\mathbf{Y}_i^{(1)}) + h(\mathbf{Y}_j^{(1)}) - h(\mathbf{Y}_i^{(1)}, \mathbf{Y}_j^{(1)}). \end{aligned} \quad (20)$$

It follows from (6) that

$$\mathbf{y}_i^{(1)} \doteq \text{vec}(\mathbf{Y}_i^{(1)}) = \sigma \sum_{l \neq i} (\mathbf{P}_l^T \otimes \mathbf{I}_{N_i}) \mathbf{h}_{i,l} + \mathbf{w}_i^{(1)}. \quad (21)$$

Then the PDF of $\mathbf{y}_i^{(1)}$ is $f(\mathbf{y}_i^{(1)}) = \mathcal{CN}(0, \mathbf{R}_{Y,i})$ with

$$\mathbf{R}_{Y,i} = \left(\sum_{l \neq i} \sigma^2 \mathbf{P}_l^T \mathbf{P}_l^* + \mathbf{I}_{K_1} \right) \otimes \mathbf{I}_{N_i}. \quad (22)$$

To determine the joint PDF $f(\mathbf{y}_{i,1}^{(1)}, \mathbf{y}_{j,1}^{(1)})$ of $\mathbf{y}_{i,1}^{(1)}$ and $\mathbf{y}_{j,1}^{(1)}$, consider

$$\mathbf{y}_j'^{(1)} \doteq \text{vec}(\mathbf{Y}_j^{T(1)}) = \sum_{l \neq j} \sigma \left(\mathbf{I}_{N_j} \otimes \mathbf{P}_l^T \right) \mathbf{h}_{l,j} + \mathbf{w}_j'^{(1)}. \quad (23)$$

There is a permutation matrix $\mathbf{\Pi}_j$ such that $\mathbf{y}_j^{(1)} = \text{vec}(\mathbf{Y}_j^{(1)}) = \mathbf{\Pi}_j \mathbf{y}_j'^{(1)}$. It follows that $f(\mathbf{y}_i^{(1)}, \mathbf{y}_j^{(1)}) = \mathcal{CN}(0, \mathbf{R}_{Y,(i,j)})$ where $\mathbf{R}_{Y,(i,j)}$ is the covariance matrix of $[\mathbf{y}_i^{(1)T}, \mathbf{y}_j^{(1)T}]^T$ which satisfies $|\mathbf{R}_{Y,(i,j)}| = |\mathbf{R}_{Y,(i,j)}'|$. Here

$$\begin{aligned} \mathbf{R}_{Y,(i,j)}' &\doteq \mathbb{E}\left\{ \begin{bmatrix} \mathbf{y}_i^{(1)} \\ \mathbf{y}_j'^{(1)} \end{bmatrix} \begin{bmatrix} \mathbf{y}_i^{(1)} \\ \mathbf{y}_j'^{(1)} \end{bmatrix}^H \right\} \\ &= \bar{\mathbf{R}}_{Y,(i,j)}' + \mathbf{I}_{(N_i+N_j)K_1}, \end{aligned} \quad (24)$$

with

$$\bar{\mathbf{R}}_{Y,(i,j)}' = \begin{bmatrix} \bar{\mathbf{R}}_{Y,i} & \sigma^2 \mathbf{P}_j^T \otimes \mathbf{P}_i^* \\ \sigma^2 \mathbf{P}_j^* \otimes \mathbf{P}_i^T & \bar{\mathbf{R}}_{Y,j}' \end{bmatrix}, \quad (25)$$

and $\mathbf{R}_{Y,j}' = \mathbf{I}_{N_j} \otimes (\sigma^2 \sum_{l \neq j} \mathbf{P}_l^T \mathbf{P}_l^* + \mathbf{I}_{K_1})$, $\bar{\mathbf{R}}_{Y,i} = (\sigma^2 \sum_{l \neq i} \mathbf{P}_l^T \mathbf{P}_l^*) \otimes \mathbf{I}_{N_i}$ and $\bar{\mathbf{R}}_{Y,j}' = \mathbf{I}_{N_j} \otimes (\sum_{l \neq j} \sigma^2 \mathbf{P}_l^T \mathbf{P}_l^*)$. Then, similar to an expression in the proof for Lemma 2, we have

$$C_{\text{key},i,j}^{(1)} = \log_2 |\mathbf{R}_{Y,i}| + \log_2 |\mathbf{R}_{Y,j}| - \log_2 |\mathbf{R}_{Y,(i,j)}|. \quad (26)$$

To study the DoF of $C_{\text{key},i,j}^{(1)}$, we need to remember Assumption 1. It follows from (22) that $\log_2 |\mathbf{R}_{Y,i}| = N_i \log_2 | \sum_{l \neq i} \sigma^2 \mathbf{P}_l^T \mathbf{P}_l^* + \mathbf{I}_{K_1}|$. It follows from Lemma 1 that with respect to $\log_2 \sigma^2$, $\text{DoF}(\log_2 |\mathbf{R}_{Y,i}|) = N_i \text{rank}(\sum_{l \neq i} \sigma^2 \mathbf{P}_l^T \mathbf{P}_l^*) = N_i \min(K_1, N_T - N_i) = N_i(N_T - N_i)$ where we used $K_1 \geq N_T - N_{\min}$.

To find the DoF of $\log_2 |\mathbf{R}_{Y,(i,j)}|$, we need to examine the rank of $\bar{\mathbf{R}}_{Y,(i,j)}'$ in (25), which is proportional to σ^2. We can rewrite (21) as $\mathbf{y}_i^{(1)} = \sigma \mathbf{A}_i \mathbf{h}_i + \mathbf{w}_i^{(1)}$, and (23) as $\mathbf{y}_j'^{(1)} = \sigma \mathbf{A}_j' \mathbf{h}_j + \mathbf{w}_i'^{(1)}$. Clearly, both $\mathbf{A}_i$ and $\mathbf{A}_j'$ have the full column ranks $N_i(N_T - N_i)$ and $N_j(N_T - N_j)$ respectively. Note that subject to Assumption 1, every node can consistently estimate its receive channel matrices relative to other nodes. Also, there are exactly $N_i N_j$ entries in $\mathbf{h}_i$ that are identical to a corresponding set of $N_i N_j$ entries in $\mathbf{h}_j$, which are the entries in $\mathbf{H}_{i,j} = \mathbf{H}_{j,i}^T$. Other than that, all entries in $\mathbf{h}_i$ and $\mathbf{h}_j$ are statistically independent. It follows that

$$\begin{aligned} \bar{\mathbf{R}}_{Y,(i,j)}' &= \sigma^2 \mathbb{E}\left\{ \begin{bmatrix} \mathbf{A}_i \mathbf{h}_i \\ \mathbf{A}_j' \mathbf{h}_j \end{bmatrix} \begin{bmatrix} \mathbf{A}_i \mathbf{h}_i \\ \mathbf{A}_j' \mathbf{h}_j \end{bmatrix}^H \right\} \\ &= \sigma^2 \text{diag}(\mathbf{A}_i, \mathbf{A}_j') \mathbf{R}_{h,i,j} \text{diag}^H(\mathbf{A}_i, \mathbf{A}_j') \end{aligned} \quad (27)$$

where

$$\mathbf{R}_{h,i,j} = \mathbb{E}\left\{ \begin{bmatrix} \mathbf{h}_i \\ \mathbf{h}_j \end{bmatrix} \begin{bmatrix} \mathbf{h}_i \\ \mathbf{h}_j \end{bmatrix}^H \right\}. \quad (28)$$

For each identical pair of entries in $\mathbf{h}_i$ and $\mathbf{h}_j$, there is a corresponding row vector $\mathbf{r}_l^H$ (with two nonzero entries equal to 1 and -1 and zeros elsewhere) such that $\mathbf{r}_l^H \mathbf{R}_{h,i,j} = 0$. Since there are total $N_i N_j$ such independent vectors, the rank

deficiency of $\mathbf{R}_{h,i,j}$ is $N_i N_j$, i.e., $\text{rank}(\mathbf{R}_{h,i,j}) = N_i(N_T - N_i) + N_j(N_T - N_j) - N_i N_j$. Since $\text{diag}(\mathbf{A}_i, \mathbf{A}_j')$ has the full column rank $N_i(N_T - N_i) + N_j(N_T - N_j)$, $\text{rank}(\bar{\mathbf{R}}_{Y,(i,j)}') = \text{rank}(\mathbf{R}_{h,i,j}) = N_i(N_T - N_i) + N_j(N_T - N_j) - N_i N_j$.

Then, it follows from (26) that for every $i \neq j$,

$$
\begin{aligned}
\text{DoF}(C_{\text{key},i,j}^{(1)}) &= \text{DoF}(\log_2 |\mathbf{R}_{Y,i}|) \\
&\quad + \text{DoF}(\log_2 |\mathbf{R}_{Y,j}|) - \text{DoF}(\log_2 |\mathbf{R}_{Y,(i,j)}'|) \\
&= N_i N_j.
\end{aligned}
\tag{29}
$$

Assumption 3 implies that the reciprocal channel matrices for different pairs of users are independent of each other for all $M \geq 2$. At high power in phase 1 of all-user ANECE, all users can obtain their reciprocal channel matrices, which can be then used to generate independent secret keys for all pairs. Hence, the pair-wise secret keys based on phase 1 of all-user ANECE are independent of each other.

As a summary, we have:

Property 3: Under Assumptions 1 and 3, the SDoF between user i and user j from phase 1 of all-user ANECE is $N_i N_j$. Furthermore, the pair-wise secret keys from phase 1 of all-user ANECE are independent of each other at high power for all $M \geq 2$.

This property is stronger than a related result in Property 5 in [27]. The latter is based on an optimal set of pilot matrices derived in [26] subject to Assumptions 1 and 3 and $N_i = N$ for all i. Property 3 shown above holds for all cases subject to Assumptions 1 and 3.

VI. PHASE-2 SKC OF ALL-USER ANECE

We now analyse $C_{\text{key},i,j,\text{low}}^{(2)}$ in (17), i.e., $C_{\text{key},i,j,\text{low}}^{(2)} = C_{i,j} - C_{i,E}$. Note that $C_{i,j}$ defined below (17) can be seen as encryption capacity between user i and user j via a public channel using data collected from phase 2 of all-user ANECE, and $C_{i,E}$ is the encryption leakage capacity from user i to Eve via the public channel. Also note that the phase-2 SKC of ANECE does not require channel reciprocity.

A. Encryption Capacity Between Users

It follows from the chain rule of mutual information that

$$
\begin{aligned}
C_{i,j} &\approx \mathbb{I}(\mathbf{X}_i, \mathbf{Y}_i^{(2)}; \mathbf{X}_j, \mathbf{Y}_j^{(2)} | \mathbf{H}_i, \mathbf{H}_j) \\
&= \mathbb{I}(\mathbf{X}_i; \mathbf{X}_j, \mathbf{Y}_j^{(2)} | \mathbf{H}_i, \mathbf{H}_j) + \mathbb{I}(\mathbf{Y}_i^{(2)}; \mathbf{X}_j, \mathbf{Y}_j^{(2)} | \mathbf{X}_i, \mathbf{H}_i, \mathbf{H}_j) \\
&= \mathbb{I}(\mathbf{X}_i; \mathbf{Y}_j^{(2)} | \mathbf{H}_i, \mathbf{H}_j) + \mathbb{I}(\mathbf{X}_i; \mathbf{X}_j | \mathbf{Y}_j^{(2)}, \mathbf{H}_i, \mathbf{H}_j) \\
&\quad + \mathbb{I}(\mathbf{Y}_i^{(2)}; \mathbf{X}_j | \mathbf{X}_i, \mathbf{H}_i, \mathbf{H}_j) \\
&\quad + \mathbb{I}(\mathbf{Y}_i^{(2)}; \mathbf{Y}_j^{(2)} | \mathbf{X}_i, \mathbf{X}_j, \mathbf{H}_i, \mathbf{H}_j).
\end{aligned}
\tag{30}
$$

Applying the independence between $\mathbf{X}_i$ and $\mathbf{X}_j$, the independence of $\mathbf{X}_i$ for any i on $\mathbf{H}_j$ for any j, the independence of $\mathbf{Y}_i^{(2)}$ on $\mathbf{X}_i$, and the independence of $\mathbf{Y}_j^{(2)}$ on $\mathbf{H}_i$ given $\mathbf{H}_j$, we then have

$$
\begin{aligned}
C_{i,j} &\approx \mathbb{I}(\mathbf{X}_i; \mathbf{Y}_j^{(2)} | \mathbf{H}_j) + \mathbb{I}(\mathbf{Y}_i^{(2)}; \mathbf{X}_j | \mathbf{H}_i) \\
&\quad + \mathbb{I}(\mathbf{Y}_i^{(2)}; \mathbf{Y}_j^{(2)} | \mathbf{X}_i, \mathbf{X}_j, \mathbf{H}_i, \mathbf{H}_j).
\end{aligned}
\tag{31}
$$

where the third term is zero for $M = 2$ but non-zero in general for $M > 2$. The first two terms are similar (or symmetric). The second term can be written as

$$
\mathbb{I}(\mathbf{Y}_i^{(2)}; \mathbf{X}_j | \mathbf{H}_i) = h(\mathbf{Y}_i^{(2)} | \mathbf{H}_i) - h(\mathbf{Y}_i^{(2)} | \mathbf{X}_j, \mathbf{H}_i).
\tag{32}
$$

The 3rd term in (31) is

$$
\begin{aligned}
&\mathbb{I}(\mathbf{Y}_i^{(2)}; \mathbf{Y}_j^{(2)} | \mathbf{X}_i, \mathbf{X}_j, \mathbf{H}_i, \mathbf{H}_j) \\
&= h(\mathbf{Y}_i^{(2)} | \mathbf{X}_i, \mathbf{X}_j, \mathbf{H}_i, \mathbf{H}_j) \\
&\quad - h(\mathbf{Y}_i^{(2)} | \mathbf{Y}_j^{(2)}, \mathbf{X}_i, \mathbf{X}_j, \mathbf{H}_i, \mathbf{H}_j) \\
&= h(\mathbf{Y}_i^{(2)} | \mathbf{X}_j, \mathbf{H}_i) + h(\mathbf{Y}_j^{(2)} | \mathbf{X}_i, \mathbf{X}_j, \mathbf{H}_i, \mathbf{H}_j) \\
&\quad - h(\mathbf{Y}_i^{(2)}, \mathbf{Y}_j^{(2)} | \mathbf{X}_i, \mathbf{X}_j, \mathbf{H}_i, \mathbf{H}_j) \\
&= h(\mathbf{Y}_i^{(2)} | \mathbf{X}_j, \mathbf{H}_i) + h(\mathbf{Y}_j^{(2)} | \mathbf{X}_i, \mathbf{H}_j) \\
&\quad - h(\mathbf{Y}_i^{(2)}, \mathbf{Y}_j^{(2)} | \mathbf{X}_i, \mathbf{X}_j, \mathbf{H}_i, \mathbf{H}_j).
\end{aligned}
\tag{33}
$$

Then it is easy to verify that (31) becomes

$$
\begin{aligned}
C_{i,j} &\approx h(\mathbf{Y}_i^{(2)} | \mathbf{H}_i) + h(\mathbf{Y}_j^{(2)} | \mathbf{H}_j) \\
&\quad - h(\mathbf{Y}_i^{(2)}, \mathbf{Y}_j^{(2)} | \mathbf{X}_i, \mathbf{X}_j, \mathbf{H}_i, \mathbf{H}_j).
\end{aligned}
\tag{34}
$$

While Lemma 2 readily applies to the first two terms in (34), the third term involves a complexity due to correlation between $\mathbf{H}_i$ and $\mathbf{H}_j$ for $M > 2$.

It follows from (9) that

$$
\mathbf{y}_i^{(2)} \doteq \text{vec}(\mathbf{Y}_i^{(2)}) = \sigma \sum_{j \neq i} (\mathbf{I}_{K_2} \otimes \mathbf{H}_{i,j}) \mathbf{x}_j + \mathbf{w}_i.
\tag{35}
$$

One can then verify that the PDF of $\mathbf{y}_i^{(2)}$ given $\mathbf{H}_i$ is $f(\mathbf{y}_i^{(2)} | \mathbf{H}_i) = \mathcal{CN}(0, \mathbf{R}_i)$ with

$$
\mathbf{R}_i \doteq \mathbb{E}\{\mathbf{y}_i^{(2)} \mathbf{y}_i^{(2)H} | \mathbf{H}_i\} = \mathbf{I}_{K_2} \otimes (\sigma^2 \mathbf{R}_{H,i} + \mathbf{I}_{N_i})
\tag{36}
$$

with $\mathbf{R}_{H,i} = \sum_{l \neq i} \mathbf{H}_{i,l} \mathbf{H}_{i,l}^H$. Similarly, the PDF of $\mathbf{y}_{(i,j)}^{(2)}$ given $\mathbf{X}_i$, $\mathbf{X}_j$, $\mathbf{H}_i$ and $\mathbf{H}_j$ is $f(\mathbf{y}_{(i,j)}^{(2)} | \mathbf{X}_i, \mathbf{X}_j, \mathbf{H}_i, \mathbf{H}_j) = \mathcal{CN}(\mathbf{m}_{(i,j)}, \mathbf{R}_{(i,j)})$ with $\mathbf{y}_{(i,j)}^{(2)H} \doteq [\mathbf{y}_i^{(2)H}, \mathbf{y}_j^{(2)H}]$, $\mathbf{m}_{(i,j)} = [\mathbf{m}_i^H, \mathbf{m}_j^H]^H$, $\mathbf{m}_i = (\mathbf{I}_{K_2} \otimes \mathbf{H}_{i,j})\mathbf{x}_j$, $\mathbf{m}_j = (\mathbf{I}_{K_2} \otimes \mathbf{H}_{j,i})\mathbf{x}_i$, and

$$
\begin{aligned}
\mathbf{R}_{(i,j)} &\doteq \mathbb{E}\left\{(\mathbf{y}_{(i,j)}^{(2)} - \mathbf{m}_{(i,j)})(\mathbf{y}_{(i,j)}^{(2)} - \mathbf{m}_{(i,j)})^H | \mathbf{H}_i, \mathbf{H}_j\right\} \\
&= \begin{bmatrix} \mathbf{I}_{K_2} \otimes (\sigma^2 \mathbf{R}_{H,i} + \mathbf{I}_{N_i}) & \mathbf{I}_{K_2} \otimes \sigma^2 \mathbf{R}_{H,i,j} \\ \mathbf{I}_{K_2} \otimes \sigma^2 \mathbf{R}_{H,i,j} & \mathbf{I}_{K_2} \otimes (\sigma^2 \mathbf{R}_{H,j} + \mathbf{I}_{N_j}) \end{bmatrix} \\
&= \mathbf{\Pi}(\mathbf{I}_{K_2} \otimes (\sigma^2 \mathbf{R}_{H,(i,j)} + \mathbf{I}_{N_i + N_j})\mathbf{\Pi}^T
\end{aligned}
\tag{37}
$$

where $\mathbf{\Pi}$ is a permutation matrix, $\mathbf{R}_{H,i,j} = \sum_{l \notin \{i,j\}} \mathbf{H}_{i,l} \mathbf{H}_{i,l}^H$, $\mathbf{R}_{H,(i,j)} = \sum_{l \notin \{i,j\}} \mathbf{H}_{(i,j),l} \mathbf{H}_{(i,j),l}^H$, and $\mathbf{H}_{(i,j),l} = [\mathbf{H}_{i,l}^H, \mathbf{H}_{j,l}^H]^H$. It follows that

$$
\begin{aligned}
C_{i,j} &\approx \mathbb{E}\{\log_2 |\mathbf{R}_i|\} + \mathbb{E}\{\log_2 |\mathbf{R}_j|\} - \mathbb{E}\{\log_2 |\mathbf{R}_{(i,j)}|\} \\
&= K_2 \mathbb{E}\{\log_2 |\sigma^2 \mathbf{R}_{H,i} + \mathbf{I}_{N_i}|\} + K_2 \mathbb{E}\{\log_2 |\sigma^2 \mathbf{R}_{H,j} + \mathbf{I}_{N_j}|\} \\
&\quad - K_2 \mathbb{E}\{\log_2 |\sigma^2 \mathbf{R}_{H,(i,j)} + \mathbf{I}_{N_i + N_j}|\}.
\end{aligned}
\tag{38}
$$

It follows from Assumption 3 that with probability one, $\text{rank}(\mathbf{H}_{i,j}) = \min(N_i, N_j)$, $\text{rank}(\sum_{l \neq i} \mathbf{H}_{i,l} \mathbf{H}_{i,l}^H) = \min(N_i, \sum_{l \neq i} N_l)$, $\text{rank}(\mathbf{H}_{i,j,l}) = \min(N_i + N_j, N_l)$, and

$\text{rank}(\sum_{l \notin \{i,j\}} \mathbf{H}_{(i,j),l} \mathbf{H}_{(i,j),l}^H) = \min(N_i + N_j, \sum_{l \notin \{i,j\}} N_l)$ for all $i \neq j$ and all $l \notin \{i, j\}$. Therefore, it follows from a simple application of Corollary 1 to (38) that the DoF of $C_{i,j}$ relative to $\log_2 \sigma^2$ is

$$\text{DoF}(C_{i,j}) = K_2 \min(N_i, N_T - N_i) + K_2 \min(N_j, N_T - N_j)$$
$$- K_2 \min(N_i + N_j, N_T - N_i - N_j). \quad (39)$$

As a summary, we have:

Property 4: Under Assumptions 2, 3 and 5, the mutual information $C_{i,j}$ between $\{\mathbf{X}_i, \mathbf{Y}_i^{(2)}\}$ and $\{\mathbf{X}_j, \mathbf{Y}_j^{(2)}\}$ conditioned on $\{\mathbf{Y}_i^{(1)}, \mathbf{Y}_j^{(1)}\}$ at high power in phase 1 is given by (38) while the DoF of $C_{i,j}$ relative to $\log_2 \sigma^2$ is given by (39).

B. Encryption Leakage Capacity From User i to Eve

Now we consider $C_{i,E}$ shown in (15), from which we have

$$C_{i,E} \approx \mathbb{I}(\mathbf{X}_i, \mathbf{Y}_i^{(2)}; \mathbf{Y}_E^{(2)} | \mathbf{H}_i, \mathbf{H}_{E,P})$$
$$= h(\mathbf{X}_i, \mathbf{Y}_i^{(2)} | \mathbf{H}_i, \mathbf{H}_{E,P}) - h(\mathbf{X}_i, \mathbf{Y}_i^{(2)} | \mathbf{Y}_E^{(2)}, \mathbf{H}_i, \mathbf{H}_{E,P})$$
$$= h(\mathbf{X}_i | \mathbf{H}_i, \mathbf{H}_{E,P}) + h(\mathbf{Y}_i^{(2)} | \mathbf{X}_i, \mathbf{H}_i, \mathbf{H}_{E,P})$$
$$- h(\mathbf{X}_i, \mathbf{Y}_i^{(2)}, \mathbf{Y}_E^{(2)} | \mathbf{H}_i, \mathbf{H}_{E,P}) + h(\mathbf{Y}_E^{(2)} | \mathbf{H}_i, \mathbf{H}_{E,P})$$
$$= h(\mathbf{Y}_i^{(2)} | \mathbf{X}_i, \mathbf{H}_i, \mathbf{H}_{E,P}) + h(\mathbf{Y}_E^{(2)} | \mathbf{H}_i, \mathbf{H}_{E,P})$$
$$- h(\mathbf{Y}_i^{(2)}, \mathbf{Y}_E^{(2)} | \mathbf{X}_i, \mathbf{H}_i, \mathbf{H}_{E,P})$$
$$= h(\mathbf{Y}_i^{(2)} | \mathbf{H}_i) + h(\mathbf{Y}_E^{(2)} | \mathbf{H}_{E,P})$$
$$- h(\mathbf{Y}_i^{(2)}, \mathbf{Y}_E^{(2)} | \mathbf{X}_i, \mathbf{H}_i, \mathbf{H}_{E,P}). \quad (40)$$

The first term in (40) is the same as that in (34), for which we have a simple expression due to the Gaussian distribution of $\mathbf{Y}_i^{(2)}$ given $\mathbf{H}_i$, which (by Lemma 2) leads to

$$\text{DoF}(h(\mathbf{Y}_i^{(2)} | \mathbf{H}_i)) = \min(N_i, N_T - N_i) K_2. \quad (41)$$

But the other two terms in (40) are more involved because $\mathbf{Y}_E^{(2)}$ is not Gaussian distributed while conditioned on $\mathbf{H}_{E,P}$. Notice the second term in (12), which is a product of two Gaussian distributed matrices. While finding the simpler forms of $h(\mathbf{Y}_E^{(2)} | \mathbf{H}_{E,P})$ and $h(\mathbf{Y}_i^{(2)}, \mathbf{Y}_E^{(2)} | \mathbf{X}_i, \mathbf{H}_i, \mathbf{H}_{E,P})$ remains an open challenge, we focus next on the DoF of these two terms.

1) Analysis of $h(\mathbf{Y}_E^{(2)} | \mathbf{H}_{E,P})$: Recall (12) and define

$$\mathbf{Y}_E^{(2)} = [\mathbf{Y}_{E,\alpha}^{(2)}, \mathbf{Y}_{E,\beta}^{(2)}] = \begin{bmatrix} \mathbf{Y}_{E,\alpha,a}^{(2)}, \mathbf{Y}_{E,\beta,a}^{(2)} \\ \mathbf{Y}_{E,\alpha,a}^{(2)}, \mathbf{Y}_{E,\beta,b}^{(2)} \end{bmatrix} \quad (42)$$

where $\mathbf{Y}_{E,\alpha}^{(2)}$ has $\min(N_{\min}, K_2)$ columns, $\mathbf{Y}_{E,\beta}^{(2)}$ has $\Delta K_2 = (K_2 - N_{\min})^+$ columns, $[\mathbf{Y}_{E,\alpha,a}^{(2)}, \mathbf{Y}_{E,\beta,a}^{(2)}]$ has $\min(N_T, N_E)$ rows, and $[\mathbf{Y}_{E,\alpha,b}^{(2)}, \mathbf{Y}_{E,\beta,b}^{(2)}]$ has $\Delta N_E = (N_E - N_T)^+$ rows.

It follows that

$$\text{DoF}(h(\mathbf{Y}_{E,\alpha}^{(2)} | \mathbf{H}_{E,P})) \geq \text{DoF}(h(\mathbf{Y}_{E,\alpha}^{(2)} | \mathbf{H}_{E,P}, \mathbf{X}))$$
$$= N_E \min(N_{\min}, K_2) \quad (43)$$

where the last equality is due to the unknown $\mathbf{H}_{E,P,\perp}$. Also the equality in the above "$\geq$" holds since $\mathbf{Y}_{E,\alpha}^{(2)}$ has total $N_E \min(N_{\min}, K_2)$ elements.

If $K_2 > N_{\min}$, we can write

$$\mathbf{Y}_{E,T} \doteq [\mathbf{Y}_E^{(1)}, \mathbf{Y}_{E,\alpha}^{(2)}] = \sigma \mathbf{H}_E \mathbf{P}_T + \mathbf{W}_{E,T} \quad (44)$$

with $\mathbf{P}_T = [\mathbf{P}, \mathbf{X}_\alpha] \in \mathbb{C}^{N_T \times N_T}$ and $\mathbf{X}_\alpha$ being the first $N_{\min}$ columns of $\mathbf{X} = [\mathbf{X}_\alpha, \mathbf{X}_\beta]$. At high power, $\mathbf{H}_E \approx \frac{1}{\sigma} \mathbf{Y}_{E,T} \mathbf{P}_T^{-1}$. It follows that

$$\mathbf{Y}_{E,\beta,a}^{(2)} \approx \sigma \mathbf{H}_{E,a} \mathbf{X}_\beta = \mathbf{Y}_{E,T,a} \mathbf{P}_T^{-1} \mathbf{X}_\beta \quad (45)$$

where we have used the partitions $\mathbf{Y}_{E,T} = [\mathbf{Y}_{E,T,a}^T, \mathbf{Y}_{E,T,b}^T]^T$ and $\mathbf{H}_E = [\mathbf{H}_{E,a}^T, \mathbf{H}_{E,b}^T]^T$, which are compatible with $\mathbf{Y}_{E,\alpha} = [\mathbf{Y}_{E,\alpha,a}^T, \mathbf{Y}_{E,\alpha,b}^T]^T$. Furthermore,

$$\text{DoF}(h(\mathbf{Y}_{E,\beta,a}^{(2)} | \mathbf{Y}_{E,\alpha}^{(2)}, \mathbf{H}_{E,P})) = \text{DoF}(h(\mathbf{Y}_{E,\beta,a}^{(2)} | \mathbf{Y}_{E,\alpha}^{(2)}, \mathbf{Y}_E^{(1)}))$$
$$= \text{DoF}(h(\mathbf{Y}_{E,\beta,a}^{(2)} | \mathbf{Y}_{E,T})) \geq \text{DoF}(h(\mathbf{Y}_{E,\beta,a}^{(2)} | \mathbf{Y}_{E,T}, \mathbf{X}_\alpha))$$
$$= \text{DoF}(h(\mathbf{Y}_{E,\beta,a}^{(2)} | \mathbf{H}_E)) = \min(N_E, N_T) \Delta K_2 \quad (46)$$

where the last equality is due to the unknown $\mathbf{X}_\beta$. Also the equality in the above "$\geq$" holds since $\mathbf{Y}_{E,\beta,a}^{(2)}$ has $\min(N_E, N_T) \Delta K_2$ elements.

If $K_2 > N_{\min}$ and $N_E > N_T$, we can write $\mathbf{Y}_{E,\beta,b}^{(2)} \approx \mathbf{Y}_{E,T,b} \mathbf{P}_T^{-1} \mathbf{X}_\beta \approx \mathbf{Y}_{E,T,b} \mathbf{Y}_{E,T,a}^{-1} \mathbf{Y}_{E,\beta,a}^{(2)}$ and hence

$$\text{DoF}(h(\mathbf{Y}_{E,\beta,b}^{(2)} | \mathbf{Y}_{E,\beta,a}^{(2)}, \mathbf{Y}_{E,\alpha}^{(2)}, \mathbf{H}_{E,P})) = 0. \quad (47)$$

Adding up (43), (46) and (47) yields

$$\text{DoF}(h(\mathbf{Y}_E^{(2)} | \mathbf{H}_{E,P}))$$
$$= N_E \min(N_{\min}, K_2) + \min(N_E, N_T) \Delta K_2. \quad (48)$$

2) Analysis of $h(\mathbf{Y}_i^{(2)}, \mathbf{Y}_E^{(2)} | \mathbf{X}_i, \mathbf{H}_i, \mathbf{H}_{E,P})$: For convenience, let us consider $h(\mathbf{Y}_1^{(2)}, \mathbf{Y}_E^{(2)} | \mathcal{C}_1)$ without loss of generality. Here $\mathcal{C}_1 = \{\mathbf{X}_1, \mathbf{H}_{1,P}, \mathbf{H}_{1,P,\perp}, \mathbf{H}_{E,P}\}$. It follows that

$$\text{DoF}(h(\mathbf{Y}_1^{(2)} | \mathcal{C}_1)) = \min(N_1, N_{T,1}) K_2 \quad (49)$$

with $N_{T,1} = N_T - N_1$. Given high power, $\mathcal{C}_1$ and $\mathbf{Y}_1^{(2)} \approx \sigma \mathbf{H}_1 \mathbf{X}_{(1)}$, we have $\mathbf{X}_{(1)} \approx \mathbf{X}_{1,0} + \mathbf{N}_1 \mathbf{T}_1$, where $\mathbf{X}_{1,0}$ and $\mathbf{N}_1$ are given, the column span of $\mathbf{N}_1 \in \mathbb{C}^{N_{T,1} \times (N_{T,1} - N_1)^+}$ is the right null space of $\mathbf{H}_1 \in \mathbb{C}^{N_1 \times N_{T,1}}$, and $\mathbf{T}_1 \in \mathbb{C}^{(N_{T,1} - N_1)^+ \times K_2}$ is arbitrary. So,

$$\mathbf{Y}_E^{(2)} \approx \sigma \mathbf{H}_E \mathbf{X} \approx \sigma \mathbf{H}_E \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_{1,0} + \mathbf{N}_1 \mathbf{T}_1 \end{bmatrix}. \quad (50)$$

Recall the (α, β)-partitions $\mathbf{Y}_E^{(2)} = [\mathbf{Y}_{E,\alpha}^{(2)}, \mathbf{Y}_{E,\beta}^{(2)}]$ with $\mathbf{Y}_{E,\alpha}^{(2)} \in \mathbb{C}^{N_E \times \min(N_{\min}, K_2)}$ and $\mathbf{Y}_{E,\beta}^{(2)} \in \mathbb{C}^{N_E \times \Delta K_2}$ and $\Delta K_2 = (K_2 - N_{\min})^+$. Similar to (43),

$$\text{DoF}(h(\mathbf{Y}_{E,\alpha}^{(2)} | \mathbf{Y}_1^{(2)}, \mathcal{C}_1)) = N_E \min(N_{\min}, K_2). \quad (51)$$

For $K_2 > N_{\min}$, recall $\mathbf{Y}_{E,T} = [\mathbf{Y}_E^{(1)}, \mathbf{Y}_{E,\alpha}^{(2)}] \approx \mathbf{H}_E \mathbf{P}_T$ and $\mathbf{H}_E \approx \mathbf{Y}_{E,T} \mathbf{P}_T^{-1}$. Note that $\mathbf{P}_T$ is dependent on $\mathbf{X}_\alpha$ in $\mathbf{X} = [\mathbf{X}_\alpha, \mathbf{X}_\beta]$. More specifically, given $\{\mathbf{Y}_{E,\alpha}^{(2)}, \mathbf{Y}_1^{(2)}, \mathcal{C}_1\}$ and high power, $\mathbf{H}_E$ is a (uniquely valued) function of $\mathbf{T}_{1,\alpha} \in \mathbb{C}^{(N_{T,1} - N_1)^+ \times N_{\min}}$ in $\mathbf{T}_1 = [\mathbf{T}_{1,\alpha}, \mathbf{T}_{1,\beta}]$. We see that if $(N_{T,1} - N_1)^+ = 0$, there is no more freedom in $\mathbf{Y}_{E,\beta}^{(2)}$ since both $\mathbf{H}_E$ and $\mathbf{X}$ are given at high power.

If $(N_{T,1} - N_1)^+ > 0$, we now define the (a,b)-partitions of $\mathbf{Y}_{E,\beta}^{(2)}$ as $\mathbf{Y}_{E,\beta}^{(2)} = [\mathbf{Y}_{E,\beta,a}^{(2)T}, \mathbf{Y}_{E,\beta,b}^{(2)T}]^T$ with $\mathbf{Y}_{E,\beta,a}^{(2)} \in \mathbb{C}^{\min(N_E,(N_{T,1}-N_1)^+) \times \Delta K_2}$ and $\mathbf{Y}_{E,\beta,b}^{(2)} \in \mathbb{C}^{(N_E-(N_{T,1}-N_1)^+)^+ \times \Delta K_2}$. Then at high power, we can write

$$\mathbf{Y}_{E,\beta,a}^{(2)} \approx \mathbf{C}_{\beta,a} + \mathbf{B}_a \mathbf{T}_{1,\beta} \qquad (52)$$

where $\mathbf{C}_{\beta,a}$ and $\mathbf{B}_a$ are a function of $\mathbf{T}_{1,\alpha}$ but independent of $\mathbf{T}_{1,\beta}$. Note that $\mathbf{T}_1 = [\mathbf{T}_{1,\alpha}, \mathbf{T}_{1,\beta}]$. It follows that

$$\begin{aligned}
&\mathrm{DoF}(h(\mathbf{Y}_{E,\beta,a}^{(2)}|\mathbf{Y}_{E,\alpha}^{(2)}, \mathbf{Y}_1^{(2)}, \mathcal{C}_1)) \\
&\geq \mathrm{DoF}(h(\mathbf{Y}_{E,\beta,a}^{(2)}|\mathbf{T}_{1,\alpha}, \mathbf{Y}_{E,\alpha}^{(2)}, \mathbf{Y}_1^{(2)}, \mathcal{C}_1)) \\
&= \min(N_E, (N_{T,1}-N_1)^+)\Delta K_2 \qquad (53)
\end{aligned}$$

where the last equality is due to the unknown $\mathbf{T}_{1,\beta}$. Also the equality in the above "$\geq$" holds because of the fullness of DoF (i.e., it equals the number of entries in $\mathbf{Y}_{E,\beta,a}^{(2)}$).

If $K_2 > N_{\min}$ and $N_E > (N_{T,1}-N_1)^+ > 0$, given $\{\mathbf{Y}_{E,\beta,a}^{(2)}, \mathbf{Y}_{E,\alpha}^{(2)}, \mathbf{Y}_1^{(2)}, \mathcal{C}_1\}$ implies that $\mathbf{H}_E$ is a function of $\mathbf{T}_{1,\alpha}$ but now conditioned by another random realization of $\mathbf{T}_{1,\beta}$. This makes $\mathbf{H}_E$ deterministic with probability one at high power. A constant $\mathbf{H}_E$ (or equivalently constant $\mathbf{C}_{\beta,a}$ and $\mathbf{B}_a \in \mathbb{C}^{(N_{T,1}-N_1) \times (N_{T,1}-N_1)}$ in (52)) combined with the knowledge of $\mathbf{Y}_{E,\beta,a}^{(2)}$ also makes $\mathbf{T}_{1,\beta}$ constant at high power. This in turn makes $\mathbf{Y}_{E,\beta,b}^{(2)}$ constant. Therefore,

$$\mathrm{DoF}(h(\mathbf{Y}_{E,\beta,b}^{(2)}|\mathbf{Y}_{E,\beta,a}^{(2)}, \mathbf{Y}_{E,\alpha}^{(2)}, \mathbf{Y}_1^{(2)}, \mathcal{C}_1)) = 0. \qquad (54)$$

Adding up (49), (51), (53) and (54) but with N_1 replaced by N_i yields

$$\begin{aligned}
&\mathrm{DoF}(h(\mathbf{Y}_i^{(2)}, \mathbf{Y}_E^{(2)}|\mathbf{X}_i, \mathbf{H}_i, \mathbf{H}_{E,P})) \\
&= K_2 \min(N_i, N_T - N_i) + N_E \min(N_{\min}, K_2) \\
&\quad + \Delta K_2 \min(N_E, (N_T - 2N_i)^+). \qquad (55)
\end{aligned}$$

3) Summary: Combining the above results, we have:

Property 5: The DoF of the encryption leakage capacity $C_{i,E}$ is given by Eq.(41)+Eq.(48)-Eq.(55).

C. Phase-2 SDoF of All-User ANECE

It follows from Properties 4 and 5 (i.e., (39) (40), (41), (48) and (55)) that the DoF of $C_{\mathrm{key},i,j,\mathrm{low}}^{(2)} = C_{i,j} - C_{i,E}$ is

$$\begin{aligned}
&\mathrm{DoF}(C_{\mathrm{key},i,j,\mathrm{low}}^{(2)}) = K_2 \min(N_j, N_T - N_j) \\
&\quad + K_2 \min(N_i, N_T - N_i) + \Delta K_2 \min(N_E, (N_T - 2N_i)^+) \\
&\quad - K_2 \min(N_i + N_j, N_T - N_i - N_j) \\
&\quad - \Delta K_2 \min(N_E, N_T) \qquad (56)
\end{aligned}$$

with $\Delta K_2 = (K_2 - N_{\min})^+$. The above is a non-increasing function of N_E. Since $\mathrm{DoF}(C_{\mathrm{key},i,j,\mathrm{low}}^{(2)})$ could be negative, we will also use $\mathrm{DoF}^+(C_{\mathrm{key},i,j,\mathrm{low}}^{(2)}) = \max(0, \mathrm{DoF}(C_{\mathrm{key},i,j,\mathrm{low}}^{(2)}))$ when it is needed.

It follows from the results shown in Appendix (i.e., (133), (48), (55) and (138)) that

$$\begin{aligned}
&\mathrm{DoF}(C_{\mathrm{key},i,j,\mathrm{up}}^{(2)}) \\
&= K_2 \min(N_i, N_T - N_i) + K_2 \min(N_j, N_T - N_j) \\
&\quad + \Delta K_2 \min(N_E, (N_T - 2N_i)^+) \\
&\quad + \Delta K_2 \min(N_E, (N_T - 2N_j)^+) \\
&\quad - \Delta K_2 \min(N_E, N_T) \\
&\quad - \Delta K_2 \min(N_E, (N_T - 2N_i - 2N_j)^+) \\
&\quad - K_2 \min(N_i, N_T - N_i - N_j) \\
&\quad - K_2 \min(N_j, (N_T - 2N_i - N_j)^+). \qquad (57)
\end{aligned}$$

Furthermore, one can verify that the gap between the upper and lower bounds is

$$\begin{aligned}
G_{i,j} &\doteq \mathrm{DoF}(C_{\mathrm{key},i,j,\mathrm{up}}^{(2)}) - \mathrm{DoF}(C_{\mathrm{key},i,j,\mathrm{low}}^{(2)}) \\
&= \Delta K_2 \min(N_E, (N_T - 2N_j)^+) \\
&\quad + K_2 \min(N_i + N_j, N_T - N_i - N_j) \\
&\quad - K_2 \min(N_i, N_T - N_i - N_j) \\
&\quad - K_2 \min(N_j, (N_T - 2N_i - N_j)^+) \\
&\quad - \Delta K_2 \min(N_E, (N_T - 2N_i - 2N_j)^+). \qquad (58)
\end{aligned}$$

Note that the upper bound is always symmetric, i.e., $C_{\mathrm{key},i,j,\mathrm{up}}^{(2)} = C_{\mathrm{key},j,i,up}^{(2)}$.

1) For a symmetric multi-user network: We now consider a multi-user symmetric network where $N_i = N$. One can verify that (58) reduces to

$$G_{i,j} = \begin{cases} 0, & M = 2; \\ \Delta K_2 \min(N_E, N), & M = 3; \\ \Delta K_2 \min(N_E, 2N), & M = 4; \\ 0, & M \geq \lceil 4 + \frac{N_E}{N} \rceil. \end{cases} \qquad (59)$$

with $\Delta K_2 = K_2 - \min(N, K_2) = (K_2 - N)^+$.

Furthermore, for $M \geq 4$, the gap is

$$\begin{aligned}
G_{i,j} &= \Delta K_2[\min(N_E, (M-2)N) \\
&\quad - \min(N_E, (M-4)N)] \geq 0 \qquad (60)
\end{aligned}$$

with equality if $M \geq \lceil 4 + \frac{N_E}{N} \rceil$. Also, under $M \geq \lceil 4 + \frac{N_E}{N} \rceil$, one can verify from (56) and (57) that

$$\mathrm{DoF}(C_{\mathrm{key},i,j,\mathrm{up}}^{(2)}) = \mathrm{DoF}(C_{\mathrm{key},i,j,\mathrm{low}}^{(2)}) = 0. \qquad (61)$$

One can also verify that for $N_E \geq MN$, (56) reduces to

$$\begin{aligned}
&\mathrm{DoF}^+(C_{\mathrm{key},i,j,\mathrm{low}}^{(2)}) \\
&= \begin{cases} 2N \min(N, K_2), & M = 2; \\ N(2\min(N, K_2) - K_2)^+, & M = 3; \\ (-2\Delta K_2 N)^+ = 0, & M \geq 4; \end{cases} \qquad (62)
\end{aligned}$$

and (57) reduces to $\mathrm{DoF}(C_{\mathrm{key},i,j,\mathrm{up}}^{(2)}) = 0$ for $M \geq 4$. The above (62) reaches its maximum when $K_2 = N$. We also see that if $N_E \geq MN$, the above (62) is positive for "$M = 2$ and $K_2 \geq 1$" and for "$M = 3$ and $1 \leq K_2 < 2N$".

It also follows that if $K_2 = N$, then for all $N_E \geq 0$,

$$\mathrm{DoF}(C_{\mathrm{key},i,j,\mathrm{low}}^{(2)}) = \mathrm{DoF}(C_{\mathrm{key},i,j,\mathrm{up}}^{(2)}) = \begin{cases} 2N^2, & M = 2; \\ N^2, & M = 3. \end{cases}$$

2) For asymmetric two-user network: For $M = 2$ and $N_1 \leq N_2$, one can verify that $G_{1,2} = 0$ and $G_{2,1} = \Delta K_2 \min(N_E, N_2 - N_1)$. This means that a public transmission from use 1 to user 2 using the data from phase 2 of two-user ANECE achieves the optimal SDoF. Furthermore, in this case, (57) reduces to

$$
\begin{aligned}
\mathrm{SDoF}^{(2)}_{\mathrm{original}} &\doteq \mathrm{DoF}(C^{(2)}_{s,1,2,\mathrm{up}}) = \mathrm{DoF}(C^{(2)}_{s,1,2,\mathrm{low}}) \\
&= \begin{cases} 2K_2 N_1, & \mathrm{C1}; \\ 2K_2 N_1 - \Delta K_2(N_E - \Delta N), & \mathrm{C2}; \\ 2\min(N_1, K_2)N_1, & \mathrm{C3}; \end{cases}
\end{aligned} \quad (63)
$$

with $\Delta N = N_2 - N_1$ and $\Delta K_2 = (K_2 - N_1)^+$. And "C1", "C2" and "C3" are respectively "$0 \leq N_E \leq \Delta N$", "$\Delta N \leq N_E \leq N_T = N_1 + N_2$" and "$N_E \geq N_T$", which are "small, medium and large" regions of N_E.

D. Summary

Property 6: The phase-2 SDoF of multi-user ANECE has the following properties:
1) The phase-2 SDoF of M-user ANECE is lower bounded by (56) and upper bounded by (57).
2) For a symmetric network with $N_i = N$ for all $i = 1, \cdots, M$, the gap between the upper and lower bounds is given by (58), which is zero if $K_2 \leq N$, or if $M = 2$, or if $M \geq 4 + \lceil \frac{N_E}{N} \rceil$.
3) For $M = 2$ and $N_1 \leq N_2$, the phase-2 SDoF is given by (63), which equals $2N_1^2$ if $N_E \geq N_T$ and $K_2 \geq N_1$.
4) For a symmetric network of $M \geq 3$ users, the phase-2 SDoF is zero if $M \geq 4 + \lceil \frac{N_E}{N} \rceil$, or if $N_E \geq MN$ and $M \geq 4$. But for $M = 3$, the phase-2 SDoF is positive subject to $1 \leq K_2 < 2N$ even if $N_E \geq 3N$. Furthermore, if using $M = 3$ and $K_2 = N$, the phase-2 SDoF equals N^2 for all $N_E \geq 0$.

VII. PAIR-WISE ANECE FOR MULTI-USER NETWORK

We now consider the use of pair-wise ANECE in a network of $M \geq 3$ nodes (or users) with N_i antennas on the ith node. We will show that the pair-wise ANECE has a serious disadvantage when compared with all-user ANECE.

Assume that within each coherence period, all pairs of users sequentially conduct the two-phase sessions of 2-user (i.e., pairwise) ANECE. The signal received by user i_p in phase 1 of the pth session is then

$$
\mathbf{Y}^{(1)}_{i_p} = \sigma \mathbf{H}_{i_p,j_p} \mathbf{P}_{j_p} + \mathbf{W}^{(1)}_{i_p} \quad (64)
$$

where $p = 1, 2, \cdots, P_0$, $P_0 = \frac{1}{2}M(M-1)$, $i_p \neq j_p$, and (i_p, j_p) denotes the pair of node i_p and node j_p. Assume that $\mathbf{P}_{j_p} \in \mathbb{C}^{N_{j_p} \times k_1}$ has the full row rank for all $j_p \in \{1, \cdots, M\}$. Then node i_p can obtain a consistent estimate of $\mathbf{H}_{i_p,j_p}$ from $\mathbf{Y}^{(1)}_{i_p}$. Here P_0 is the number of pairs among M users.

The signals received by Eve in phase 1 of all P_0 sessions can be now grouped into (up to a known permutation):

$$
\mathbf{Y}^{(1)}_{E,pair} = \sigma \mathbf{H}_E \mathbf{P}_{\mathrm{pair}} + \mathbf{W}^{(1)}_E \quad (65)
$$

where $\mathbf{Y}^{(1)}_E = [\mathbf{Y}^{(1)}_{E,1}, \cdots, \mathbf{Y}^{(1)}_{E,P_0}]$, $\mathbf{H}_E = [\mathbf{H}_{E,1}, \cdots, \mathbf{H}_{E,M}]$, $\mathbf{W}^{(1)}_E = [\mathbf{W}^{(1)}_{E,1}, \cdots, \mathbf{W}^{(1)}_{E,P_0}]$, and

$$
\mathbf{P}_{\mathrm{pair}} = \begin{bmatrix} \mathbf{P}_1 & \mathbf{P}_1 & 0 & \cdots \\ \mathbf{P}_2 & 0 & \mathbf{P}_2 & \cdots \\ 0 & \mathbf{P}_3 & \mathbf{P}_3 & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix}. \quad (66)
$$

The matrix $\mathbf{P}_{\mathrm{pair}}$ has $M \times P_0$ blocks as illustrated above and has the full-row rank N_T provided $M \geq 3$. Hence Eve can also obtain a consistent estimate of $\mathbf{H}_E$ in this case.

A. Analysis of Phase 1

Since phase 1 of each session requires k_1 sampling intervals (or time slots), the total sampling intervals required for all sessions within a coherence period is $P_0 k_1$. Provided that the full row-rank condition of $\mathbf{P}_{j_p}$ is met for all p, every pair of users can obtain a consistent estimate of their reciprocal channel matrix, and hence the DoF of the secret-key capacity from phase 1 of session p is given by $N_{i_p} N_{j_p}$, which is the same as (29).

For a further comparison, let $N_i = N$ for all i. Then the pair-wise ANECE requires a total of $P_0 N = \frac{1}{2}M(M-1)N$ time slots for phase-1 transmissions. But the multi-user ANECE requires only $(M-1)N$ time slots for phase-1 transmission.

B. Analysis of Phase 2

In phase 2 of the pth session, users i_p and j_p transmit $\mathbf{X}_{i_p} \in \mathbb{C}^{N_{i_p} \times k_2}$ and $\mathbf{X}_{j_p} \in \mathbb{C}^{N_{j_p} \times k_2}$ respectively, and they also receive respectively

$$
\mathbf{Y}^{(2)}_{i_p} = \sigma \mathbf{H}_{i_p,j_p} \mathbf{X}_{j_p} + \mathbf{W}^{(2)}_{i_p} \quad (67)
$$

and $\mathbf{Y}^{(2)}_{j_p} = \sigma \mathbf{H}^T_{i_p,j_p} \mathbf{X}_{i_p} + \mathbf{W}^{(2)}_{j_p}$. At the same time, Eve receives

$$
\mathbf{Y}^{(2)}_{E,p} = \sigma [\mathbf{H}_{E,i_p}, \mathbf{H}_{E,j_p}] \begin{bmatrix} \mathbf{X}_{i_p} \\ \mathbf{X}_{j_p} \end{bmatrix} + \mathbf{W}^{(2)}_{E,p} \quad (68)
$$

The secret-key capacity $C^{(2)}_{key,i_p,j_p}$ between user i_p and user j_p using measurements in phase 2 is lower bounded by

$$
C^{(2)}_{s,i_p,j_p,\mathrm{low}} = C_{i_p,j_p} - C_{i_p,E} \quad (69)
$$

where

$$
C_{i_p,j_p} \approx \mathbb{I}(\mathbf{X}_{i_p}; \mathbf{Y}^{(2)}_{j_p} | \mathbf{H}_{j_p,i_p}) + \mathbb{I}(\mathbf{X}_{j_p}; \mathbf{Y}^{(2)}_{i_p} | \mathbf{H}_{i_p,j_p}) \quad (70)
$$

which is similar to (31) where the last term is zero due to $M = 2$. It follows that $\mathrm{DoF}(\mathbb{I}(\mathbf{X}_{i_p}; \mathbf{Y}^{(2)}_{j_p} | \mathbf{H}_{j_p,i_p})) = \mathrm{DoF}(h(\mathbf{Y}^{(2)}_{j_p} | \mathbf{H}_{j_p,i_p})) = \min(N_{i_p}, N_{j_p}) k_2$. The second term is symmetric to the first term. Also

$$
\begin{aligned}
C_{i_p,E} &\approx h(\mathbf{Y}^{(2)}_{i_p} | \mathbf{H}_{i_p,j_p}) + h(\mathbf{Y}^{(2)}_{E,p} | \mathbf{H}_E) \\
&\quad - h(\mathbf{Y}^{(2)}_{i_p}, \mathbf{Y}^{(2)}_{E,p} | \mathbf{X}_{i_p}, \mathbf{H}_{i_p,j_p}, \mathbf{H}_E). \quad (71)
\end{aligned}
$$

which is similar to (40) but here the entire $\mathbf{H}_E$ is known to Eve. We know that $\mathrm{DoF}(h(\mathbf{Y}^{(2)}_{i_p} | \mathbf{H}_{i_p,j_p})) = \min(N_{i_p}, N_{j_p}) k_2$ and

$h(\mathbf{Y}_{E,p}^{(2)}|\mathbf{H}_E) = \min(N_E, N_{i_p} + N_{j_p})k_2$. For the third term in (71), we can write

$$\begin{bmatrix} \mathbf{Y}_{i_p}^{(2)} \\ \mathbf{Y}_{E,p}^{(2)} \end{bmatrix} = \sigma \begin{bmatrix} 0 \\ \mathbf{H}_{E,i_p} \end{bmatrix} \mathbf{X}_{i_p} + \sigma \begin{bmatrix} \mathbf{H}_{i_p,j_p} \\ \mathbf{H}_{E,j_p} \end{bmatrix} \mathbf{X}_{j_p}$$
$$+ \begin{bmatrix} \mathbf{W}_{i_p}^{(2)} \\ \mathbf{W}_{E,p}^{(2)} \end{bmatrix} \qquad (72)$$

where $rank \begin{bmatrix} \mathbf{H}_{i_p,j_p} \\ \mathbf{H}_{E,j_p} \end{bmatrix} = \min(N_E + N_{i_p}, N_{j_p})$. It follows that $\mathtt{DoF}(h(\mathbf{Y}_{i_p}^{(2)}, \mathbf{Y}_{E,p}^{(2)}|\mathbf{X}_{i_p}, \mathbf{H}_{i_p,j_p}, \mathbf{H}_E)) = \min(N_E + N_{i_p}, N_{j_p})k_2$.

Combing the above results, we have

$$\mathtt{DoF}(C_{s,i_p,j_p,\mathtt{low}}^{(2)}) = \min(N_{i_p}, N_{j_p})k_2$$
$$- \min(N_E, N_{i_p} + N_{j_p})k_2 + \min(N_E + N_{i_p}, N_{j_p})k_2 \qquad (73)$$

where k_2 is the number of time slots used for phase 2 of each session. For comparison with all-user ANECE, we will choose $k_2 = \frac{1}{P_0}K_2$.

The upper bound of $C_{\mathtt{key},i_p,j_p}^{(2)}$ is similar to (133) but with $M = 2$ and $\mathbf{H}_E$ known to Eve, i.e.,

$$C_{s,i_p,j_p,\mathtt{up}}^{(2)} \approx -h(\mathbf{Y}_E^{(2)}|\mathbf{H}_E) + h(\mathbf{Y}_{i_p}^{(2)}, \mathbf{Y}_{E,p}^{(2)}|\mathbf{X}_{i_p}, \mathbf{H}_{i_p,j_p}, \mathbf{H}_E)$$
$$+ h(\mathbf{Y}_{j_p}^{(2)}, \mathbf{Y}_{E,p}^{(2)}|\mathbf{X}_{j_p}, \mathbf{H}_{j_p,i_p}, \mathbf{H}_E)$$
$$- h(\mathbf{Y}_{i_p}^{(2)}, \mathbf{Y}_{j_p}^{(2)}, \mathbf{Y}_{E,p}^{(2)}|\mathbf{X}_{i_p}, \mathbf{X}_{j_p}, \mathbf{H}_{i_p,j_p}, \mathbf{H}_E) \qquad (74)$$

where $\mathbf{H}_{i_p,j_p} = \mathbf{H}_{j_p,i_p}^T$. We know the DoF of each of the first three terms in (74). The last term in (74) clearly has a zero DoF. One can verify that

$$\mathtt{DoF}(C_{s,i_p,j_p,\mathtt{up}}^{(2)})$$
$$= -\min(N_E, N_{i_p} + N_{j_p})k_2 + \min(N_E + N_{i_p}, N_{j_p})k_2$$
$$+ \min(N_E + N_{j_p}, N_{i_p})k_2. \qquad (75)$$

The gap between the upper bound in (75) and the lower bound in (73) is

$$G_{i_p,j_p} = \begin{cases} 0, & N_{i_p} \leq N_{j_p}; \\ (\min(N_E + N_{j_p}, N_{i_p}) - N_{j_p})k_2, & N_{i_p} > N_{j_p}. \end{cases} \qquad (76)$$

Note that the above implies that $\mathtt{DoF}(C_{s,i_p,j_p,up}^{(2)})$ shown in (75) is achievable by a public communication (for secret key generation) from the node with equal or smaller number of antennas to the other node.

Furthermore, if $N_i = N$ for all $i = 1, \cdots, M$, then the pair-wise SDoF for phase 2 of the pair-wise ANECE with $M \geq 3$ is

$$\mathtt{DoF}(C_{s,i_p,j_p,\mathtt{up}}^{(2)}) = \mathtt{DoF}(C_{s,i_p,j_p,\mathtt{low}}^{(2)})$$
$$= (2N - \min(N_E, 2N))k_2 \qquad (77)$$

where $k_2 = \frac{K_2}{P_0} = \frac{2}{M(M-1)}K_2$. Here we see that the phase-2 pair-wise SDoF for the pair-wise ANECE decreases to zero rapidly as $\frac{1}{M^2}$ when the number M of users in the network increases. Also for $N_E \geq 2N$, $\mathtt{DoF}(C_{s,i_p,j_p,\mathtt{up}}^{(2)}) = 0$.

C. Summary

Property 7: The pair-wise ANECE is a scheme where two-user ANECE is applied to each pair of users in a network of $M \geq 3$ multi-antenna users within each coherence period, for which we have found:

1) For pair-wise ANECE, Eve of $N_E \geq 1$ antennas is able to determine her receive channel state information completely (at high power) relative to all users.
2) While the phase-1 SDoF ($\mathtt{SDoF}^{(1)}$) of pair-wise ANECE is the same as that for all-user ANECE, the pair-wise ANECE requires $\frac{M}{2}$ times more time-slots for phase-1 transmissions than all-user ANECE.
3) The phase-2 SDoF ($\mathtt{SDoF}^{(2)}$) of pair-wise ANECE is given by (75) regardless of $N_{i_p} \leq N_{j_p}$ or $N_{i_p} > N_{j_p}$.
4) For a symmetric network where every node has the same number N of antennas, $\mathtt{SDoF}^{(2)}$ of pair-wise ANECE decreases to zero at least like $\frac{1}{M^2}$ as M increases as shown in (77).
5) For $M = 3$ and $N_E \geq 2N$, $\mathtt{SDoF}^{(2)}$ of pair-wise ANECE (see (77)) is zero. This is in contrast to $\mathtt{SDoF}^{(2)}$ of all-user ANECE (see (62)) which is positive subject to $M = 3$, $1 \leq K_2 < 2N$ and (even) $N_E \geq 3N$.

VIII. A MODIFIED TWO-USER ANECE

We now focus on a network of two users with N_1 and N_2 antennas respectively against Eve with N_E antennas. Assuming $N_1 \leq N_2$, we consider what will be called a modified two-user ANECE where node 1 transmits $[\mathbf{P}_1, \mathbf{X}_1] \in \mathbb{C}^{N_1 \times K}$ while node 2 transmits $[\mathbf{P}_2, \mathbf{X}_2] \in \mathbb{C}^{N_2 \times K}$ over total K slots in each coherence period, and each of $\mathbf{P}_1$ and $\mathbf{P}_2$ is a full-rank square matrix.

Assumption 6: $\mathbf{P}_1$ has rank N_1 and dimension $N_1 \times N_1$, and $\mathbf{P}_2$ has rank N_2 and dimension $N_2 \times N_2$.

Notice that the two square-shaped pilot matrices $\mathbf{P}_1$ and $\mathbf{P}_2$ have their lengths (numbers of columns) equal to N_1 and N_2 respectively. The two transmission phases of node 1 do not completely align with those of node 2 unless $N_1 = N_2$.

We can ignore the propagation delays. Then the signals received by node 1 during the slots $1 \leq k \leq N_2$ can be described by

$$\mathbf{Y}_1^{(1)} = \sigma \mathbf{H}_{1,2}\mathbf{P}_2 + \mathbf{W}_1^{(1)}, \qquad (78)$$

and the signals received by node 2 during $1 \leq k \leq N_1$ are

$$\mathbf{Y}_2^{(1)} = \sigma \mathbf{H}_{2,1}\mathbf{P}_1 + \mathbf{W}_2^{(1)}. \qquad (79)$$

Similarly, the signals received by node 1 during $N_2 + 1 \leq k \leq K$ are

$$\mathbf{Y}_1^{(2)} = \sigma \mathbf{H}_{1,2}\mathbf{X}_2 + \mathbf{W}_1^{(2)}, \qquad (80)$$

and the signals received by node 2 during $N_1 + 1 \leq k \leq K$ are

$$\mathbf{Y}_2^{(2)} = \sigma \mathbf{H}_{2,1}\mathbf{X}_1 + \mathbf{W}_2^{(2)}. \qquad (81)$$

Subject to a high power in pilots, Node 1 can obtain its channel matrix $\mathbf{H}_{1,2}$ from $\mathbf{Y}_1^{(1)}$, and then receive the "information" in the $N_2 \times (K - N_2)$ matrix $\mathbf{X}_2$ from $\mathbf{Y}_1^{(2)}$. Similarly, node 2

can obtain its channel matrix $\mathbf{H}_{2,1}$ from $\mathbf{Y}_2^{(1)}$, and then receive the "information" in the $N_1 \times (K - N_1)$ matrix $\mathbf{X}_1$ from $\mathbf{Y}_2^{(2)}$. Note that we will still follow the generalized secrecy capacity. So the "information" above can be randomly generated by each node.

During $1 \leq k \leq N_1$, Eve receives

$$\mathbf{Y}_E^{(1)'} = \sigma \mathbf{H}_E \mathbf{P}' + \mathbf{W}_E^{(1)'} \tag{82}$$

with $\mathbf{H}_E = [\mathbf{H}_{E,1}, \mathbf{H}_{E,2}]$, $\mathbf{P}' = [\mathbf{P}_1^H, \mathbf{P}_{2,1}^H]^H$ and $\mathbf{P}_{2,1}$ being the first N_1 columns of $\mathbf{P}_2$. Let $\mathbf{Q}_{P'}$ be a $(N_1 + N_2) \times N_1$ orthonormal matrix spanning the range of $\mathbf{P}'$, and $\mathbf{Q}_{P',\perp}$ be a $(N_1 + N_2) \times N_2$ orthogonal complement of $\mathbf{Q}_{P'}$. It follows from an analysis similar to (8) that at high power Eve can uniquely determine $\mathbf{H}_{E,P'} \doteq \mathbf{H}_E \mathbf{Q}_{P'}$ from $\mathbf{Y}_E^{(1)'}$. But all entries of $\mathbf{H}_{E,P',\perp} \doteq \mathbf{H}_E \mathbf{Q}_{P',\perp}$ remain i.i.d. $\mathcal{CN}(0,1)$. Note that the dimensions of $\mathbf{H}_{E,P'}$ and $\mathbf{H}_{E,P',\perp}$ are $N_E \times N_1$ and $N_E \times N_2$ respectively.

During $N_1 + 1 \leq k \leq K$, Eve receives

$$\mathbf{Y}_E^{(2)'} = \sigma \mathbf{H}_E \begin{bmatrix} \mathbf{X}_{1,1} & \mathbf{X}_{1,2} \\ \mathbf{P}_{2,2} & \mathbf{X}_2 \end{bmatrix} + \mathbf{W}_E^{(2)'}$$
$$= \sigma \mathbf{H}_E \mathbf{X}' + \mathbf{W}_E^{(2)'} \tag{83}$$

where $\mathbf{P}_{2,2}$ consists of the last $N_2 - N_1$ columns of $\mathbf{P}_2$, and $[\mathbf{X}_{1,1}, \mathbf{X}_{1,2}] = \mathbf{X}_1$. Equivalently,

$$[\mathbf{Y}_E^{(1)'}, \mathbf{Y}_E^{(2)'}] = \sigma \mathbf{H}_E \begin{bmatrix} \mathbf{P}_1 & \mathbf{X}_{1,1} & \mathbf{X}_{1,2} \\ \mathbf{P}_{2,1} & \mathbf{P}_{2,2} & \mathbf{X}_2 \end{bmatrix}$$
$$+ [\mathbf{W}_E^{(1)'}, \mathbf{W}_E^{(2)'}]. \tag{84}$$

Following a similar analysis shown in Section V, one can verify that the DoF of $\mathbb{I}(\mathbf{Y}_1^{(1)}; \mathbf{Y}_2^{(1)})$ is $N_1 N_2$ subject to $\mathbf{H}_{1,2} = \mathbf{H}_{2,1}^T$, which is the SDoF between the users based on phase 2 of the modified two-user ANECE. This SDoF for phase 1 is no different from the pair-wise SDoF for phase 1 of the multi-user ANECE shown before.

The secrecy capacity for phase 2 of the modified two-user ANECE is now denoted by $C_{\text{key}}^{(2)}$, which satisfies

$$C_{\text{key,low}}^{(2)} \leq C_{\text{key}}^{(2)} \leq C_{\text{key,up}}^{(2)} \tag{85}$$

where

$$C_{\text{key,low}}^{(2)} = \max_{i \neq j}(C_{\text{key},i,j,\text{low}}^{(2)}) = C_{\text{key},0}^{(2)} - \min_{i \neq j} C_{i,E}' \tag{86}$$

with $i = 1, 2$ and $j = 1, 2$. Furthermore,

$$C_{\text{key},0}^{(2)} \approx \mathbb{I}(\mathbf{X}_1; \mathbf{Y}_2^{(2)}|\mathbf{H}_{2,1}) + \mathbb{I}(\mathbf{X}_2; \mathbf{Y}_1^{(2)}|\mathbf{H}_{1,2}), \tag{87}$$

$$C_{i,E}' \approx \mathbb{I}(\mathbf{X}_i, \mathbf{Y}_i^{(2)}; \mathbf{Y}_E^{(2)'}|\mathbf{H}_{i,j}, \mathbf{H}_{E,P'}), \tag{88}$$

$$C_{\text{key,up}}^{(2)} \approx \mathbb{I}(\mathbf{X}_1, \mathbf{Y}_1^{(2)}; \mathbf{X}_2, \mathbf{Y}_2^{(2)}|\mathcal{C}') \tag{89}$$

where $\mathcal{C}' = \{\mathbf{H}_{1,2}, \mathbf{H}_{2,1}, \mathbf{H}_{E,P'}, \mathbf{Y}_E^{(2)'}\}$.

A. A Remark on Channel-Model Based Secrecy Capacity

From the perspective of wiretap channel model, another achievable secrecy capacity for phase 2 of the modified two-user ANECE (assuming a high power in phase 1) is

$$C_{\text{WTC,low}}^{(2)} = \mathbb{I}(\mathbf{X}_1; \mathbf{Y}_2^{(2)}|\mathbf{H}_{2,1}) + \mathbb{I}(\mathbf{X}_2; \mathbf{Y}_1^{(2)}|\mathbf{H}_{1,2}) - C_E' \tag{90}$$

where

$$C_E' = \mathbb{I}(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_E^{(1)'}, \mathbf{Y}_E^{(2)'}) = \mathbb{I}(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_E^{(2)'}|\mathbf{Y}_E^{(1)'})$$
$$= \mathbb{I}(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_E^{(2)'}|\mathbf{H}_{E,P'}). \tag{91}$$

To compare $C_{\text{key,low}}^{(2)}$ with $C_{\text{WTC,low}}^{(2)}$, we only need to compare $C_{1,E}'$ and $C_{2,E}'$ defined in (88) with C_E'. For DoF analysis, we can replace the approximation at high power by equality. Let us first consider $C_{2,E}'$ for which we can write

$$C_{2,E}' = \mathbb{I}(\mathbf{X}_2, \mathbf{Y}_2^{(2)}; \mathbf{Y}_E^{(2)'}|\mathbf{H}_{2,1}, \mathbf{H}_{E,P'})$$
$$= \mathbb{I}(\mathbf{X}_2, \mathbf{X}_1; \mathbf{Y}_E^{(2)'}|\mathbf{H}_{E,P'}) = C_E' \tag{92}$$

where we have used the fact that $\mathbf{Y}_2^{(2)} = \sigma \mathbf{H}_{2,1} \mathbf{X}_1$ at high power and in this case $\mathbf{Y}_2^{(2)}$ and $\mathbf{X}_1$ imply each other when the $N_2 \times N_1$ full column rank matrix $\mathbf{H}_{2,1}$ is given (here $N_1 \leq N_2$).

Next we consider $C_{1,E}'$ which is

$$C_{1,E}' = \mathbb{I}(\mathbf{X}_1, \mathbf{Y}_1^{(2)}; \mathbf{Y}_E^{(2)'}|\mathbf{H}_{1,2}, \mathbf{H}_{E,P'})$$
$$\leq \mathbb{I}(\mathbf{X}_1, \mathbf{X}_2; \mathbf{Y}_E^{(2)'}|\mathbf{H}_{E,P'}) = C_E' \tag{93}$$

where the inequality is due to the fact that at high power $\mathbf{Y}_1^{(2)} \approx \sigma \mathbf{H}_{1,2} \mathbf{X}_2$, and $\mathbf{Y}_1^{(2)}$ and $\mathbf{H}_{1,2}$ do not imply a unique $\mathbf{X}_2$ due to a right null space of the $N_1 \times N_2$ matrix $\mathbf{H}_{1,2}$ for $N_1 < N_2$.

Therefore, it follows from (86) and (90) that for $N_1 \leq N_2$, $\text{DoF}(C_{\text{key,low}}^{(2)}) \geq \text{DoF}(C_{\text{WTC,low}}^{(2)})$ where the strict inequality is expected if $N_1 < N_2$.

B. Analysis of the Lower Bound $\max_{i \neq j} C_{\text{key},i,j,\text{low}}^{(2)}$

We will find next the DoF of $C_{\text{key},i,j,\text{low}}^{(2)} = C_{\text{key},0}^{(2)} - C_{i,E}'$. The DoF of the first term is simply

$$\text{DoF}(C_{\text{key},0}^{(2)}) = \text{DoF}(\mathbb{I}(\mathbf{X}_1; \mathbf{Y}_2^{(2)}|\mathbf{H}_{2,1}))$$
$$+ \text{DoF}(\mathbb{I}(\mathbf{X}_2; \mathbf{Y}_1^{(2)}|\mathbf{H}_{1,2}))$$
$$= N_1(K - N_1) + N_1(K - N_2). \tag{94}$$

More effort is needed to find the DoF of $C_{i,E}'$. Similar to (40),

$$C_{i,E}' \approx h(\mathbf{Y}_i^{(2)}|\mathbf{H}_{i,j}) + h(\mathbf{Y}_E^{(2)'}|\mathbf{H}_{E,P'})$$
$$- h(\mathbf{Y}_i^{(2)}, \mathbf{Y}_E^{(2)'}|\mathbf{X}_i, \mathbf{H}_{i,j}, \mathbf{H}_{E,P'}). \tag{95}$$

It is clear that $\text{DoF}(h(\mathbf{Y}_1^{(2)}|\mathbf{H}_{1,2})) = N_1(K - N_2)$ and $\text{DoF}(h(\mathbf{Y}_2^{(2)}|\mathbf{H}_{2,1})) = N_1(K - N_1)$. The analyses of the other two terms in $C_{i,E}'$ are given below.

C. Analysis of $h(\mathbf{Y}_E^{(2)'}|\mathbf{H}_{E,P'})$ in the Lower Bound

Refer to (83) and define the following (α, β)-partitions

$$\mathbf{Y}_E^{(2)'} = [\mathbf{Y}_{E,\alpha}, \mathbf{Y}_{E,\beta}] \qquad (96)$$

where $\mathbf{Y}_{E,\alpha} \in \mathbb{C}^{N_E \times \min(N_2, K-N_1)}$, $\mathbf{Y}_{E,\beta} \in \mathbb{C}^{N_E \times (K-N_T)^+}$. Namely,

$$\mathbf{Y}_{E,\alpha} = \sigma \mathbf{H}_E \mathbf{X}_\alpha + \mathbf{W}_{E,\alpha}, \qquad (97)$$

$$\mathbf{Y}_{E,\beta} = \sigma \mathbf{H}_E \mathbf{X}_\beta + \mathbf{W}_{E,\beta}. \qquad (98)$$

where $[\mathbf{X}_\alpha, \mathbf{X}_\beta] = \mathbf{X}'$, $\mathbf{X}_\alpha \in \mathbb{C}^{N_T \times \min(N_2, K-N_1)}$ and $\mathbf{X}_\beta \in \mathbb{C}^{N_T \times (K-N_T)^+}$.

It follows that due to the unknown $\mathbf{H}_{E,P',\perp} \in \mathbb{C}^{N_E \times N_2}$,

$$\begin{aligned}
\mathrm{DoF}(h(\mathbf{Y}_{E,\alpha}|\mathbf{H}_{E,P'})) &\geq \mathrm{DoF}(h(\mathbf{Y}_{E,\alpha}|\mathbf{X}_\alpha, \mathbf{H}_{E,P'})) \\
&= N_E \min(N_2, K-N_1). \quad (99)
\end{aligned}$$

Since this is the total number of entries in $\mathbf{Y}_{E,\alpha}$, the equality in the above "$\geq$" holds.

If $K > N_T$, we define

$$\mathbf{Y}_{E,P} \doteq [\mathbf{Y}_E^{(1)'}, \mathbf{Y}_{E,\alpha}] = \sigma \mathbf{H}_E \mathbf{P}_T + \mathbf{W}_{E,P} \qquad (100)$$

with $\mathbf{P}_T = [\mathbf{P}', \mathbf{X}_\alpha] \in \mathbb{C}^{N_T \times N_T}$. At high power, $\mathbf{H}_E \approx \frac{1}{\sigma}\mathbf{Y}_{E,P}\mathbf{P}_T^{-1}$, which however depends on $\mathbf{X}_\alpha$.

Let the (a, b)-partitions of $\mathbf{Y}_{E,\beta}$ be

$$\mathbf{Y}_{E,\beta} = [\mathbf{Y}_{E,\beta,a}^T, \mathbf{Y}_{E,\beta,b}^T]^T \qquad (101)$$

with $\mathbf{Y}_{E,\beta,a} \in \mathbb{C}^{\min(N_E, N_T) \times (K-N_T)^+}$ and $\mathbf{Y}_{E,\beta,b} \in \mathbb{C}^{(N_E-N_T)^+ \times (K-N_T)^+}$.

It follows that due to the unknown $\mathbf{X}_\beta$ (consisting of i.i.d. entries),

$$\begin{aligned}
&\mathrm{DoF}(h(\mathbf{Y}_{E,\beta,a}|\mathbf{Y}_{E,\alpha}, \mathbf{H}_{E,P'})) \\
&\geq \mathrm{DoF}(h(\mathbf{Y}_{E,\beta,a}|\mathbf{Y}_{E,\alpha}, \mathbf{H}_{E,P'}, \mathbf{X}_\alpha)) \\
&= \mathrm{DoF}(h(\mathbf{Y}_{E,\beta,a}|\mathbf{Y}_{E,P}, \mathbf{P}_T)) \\
&= \mathrm{DoF}(h(\mathbf{Y}_{E,\beta,a}|\mathbf{H}_E)) \\
&= \min(N_E, N_T)(K-N_T)^+. \quad (102)
\end{aligned}$$

Note that conditioning on $\mathbf{H}_{E,P'}$ is equivalent at high power to conditioning on $\mathbf{Y}_E^{(1)'}$ in terms of DoF. Also, the equality in the above "$\geq$" holds since $\mathbf{Y}_{E,\beta,a}$ has total $\min(N_E, N_T)(K-N_T)^+$ elements.

If $K > N_T$ and $N_E > N_T$, then $\mathbf{Y}_{E,\beta,b}$ exists. In this case, we have at high power that $\mathbf{H}_{E,a} \approx \frac{1}{\sigma}\mathbf{Y}_{E,P,a}\mathbf{P}_T^{-1}$ and $\mathbf{H}_{E,b} \approx \frac{1}{\sigma}\mathbf{Y}_{E,P,b}\mathbf{P}_T^{-1}$. Also $\mathbf{Y}_{E,\beta,a} \approx \sigma \mathbf{H}_{E,a}\mathbf{X}_\beta$ implies $\mathbf{X}_\beta = \frac{1}{\sigma}\mathbf{H}_{E,a}^{-1}\mathbf{Y}_{E,\beta,a}$ where $\mathbf{H}_{E,a}^{-1} \in \mathbb{C}^{N_T \times N_T}$ exists with probability one. Finally,

$$\begin{aligned}
\mathbf{Y}_{E,\beta,b} &\approx \sigma \mathbf{H}_{E,b}\mathbf{X}_\beta \approx \mathbf{H}_{E,b}\mathbf{H}_{E,a}^{-1}\mathbf{Y}_{E,\beta,a} \\
&\approx \mathbf{Y}_{E,P,b}\mathbf{Y}_{E,P,a}^{-1}\mathbf{Y}_{E,\beta,a}. \quad (103)
\end{aligned}$$

Therefore,

$$\mathrm{DoF}(h(\mathbf{Y}_{E,\beta,b}|\mathbf{Y}_{E,\beta,a}, \mathbf{Y}_{E,\alpha}, \mathbf{H}_{E,P'})) = 0. \qquad (104)$$

Adding (99), (103) and (104) yields

$$\begin{aligned}
\mathrm{DoF}(h(\mathbf{Y}_E^{(2)'}|\mathbf{H}_{E,P'})) &= N_E \min(N_2, K-N_1) \\
&\quad + \min(N_E, N_T)(K-N_T)^+. \quad (105)
\end{aligned}$$

D. Analysis of $h(\mathbf{Y}_i^{(2)}, \mathbf{Y}_E^{(2)'}|\mathbf{X}_i, \mathbf{H}_{i,j}, \mathbf{H}_{E,P'})$ in the Lower Bound

1) For $h(\mathbf{Y}_1^{(2)}, \mathbf{Y}_E^{(2)'}|\mathbf{X}_1, \mathbf{H}_{1,2}, \mathbf{H}_{E,P'})$: We will write it as $h(\mathbf{Y}_1^{(2)}, \mathbf{Y}_E^{(2)'}|\mathcal{C}_{1,2})$ with $\mathcal{C}_{1,2} = \{\mathbf{X}_1, \mathbf{H}_{1,2}, \mathbf{H}_{E,P'}\}$. Here $\mathbf{X}_1$ is among the conditions. It follows from (80) that $h(\mathbf{Y}_1^{(2)}|\mathcal{C}_{1,2}) = h(\mathbf{Y}_1^{(2)}|\mathbf{H}_{1,2})$ and hence

$$\mathrm{DoF}(h(\mathbf{Y}_1^{(2)}|\mathcal{C}_{1,2})) = N_1(K-N_2). \qquad (106)$$

Next we need to consider the second component $h(\mathbf{Y}_E^{(2)'}|\mathbf{Y}_1^{(2)}, \mathcal{C}_{1,2})$. Conditioned on $\mathbf{Y}_1^{(2)}$ and $\mathcal{C}_{1,2}$ at high power, we know from (80) that $\mathbf{X}_2 \approx \mathbf{X}_{2,0} + \mathbf{N}_{x2}\mathbf{T}_{x2}$ where $\mathbf{X}_{2,0}$ is the minimum norm solution of $\mathbf{X}_2$ to $\mathbf{Y}_1^{(2)} = \sigma \mathbf{H}_{1,2}\mathbf{X}_2$, the columns of $\mathbf{N}_{x2} \in \mathbb{C}^{N_2 \times \Delta N}$ span the right null space of $\mathbf{H}_{1,2}$, and $\mathbf{T}_{x2} \in \mathbb{C}^{\Delta N \times (K-N_2)}$ is arbitrary. Then, (83) becomes

$$\mathbf{Y}_E^{(2)'} = \sigma \mathbf{H}_E \begin{bmatrix} \mathbf{X}_{1,1} & \mathbf{X}_{1,2} \\ \mathbf{P}_{2,2} & \mathbf{X}_{2,0} + \mathbf{N}_{x2}\mathbf{T}_{x2} \end{bmatrix} + \mathbf{W}_E^{(2)'} \qquad (107)$$

where only $\mathbf{T}_{x2}$ and $\mathbf{H}_{E,P',\perp}$ (inside $\mathbf{H}_E = [\mathbf{H}_{E,P'}, \mathbf{H}_{E,P',\perp}]\mathbf{Q}^H$) are unknown subject to given $\{\mathbf{Y}_1^{(2)}, \mathcal{C}_{1,2}\}$ and high power.

Recall the (α, β)-partitions $\mathbf{Y}_E^{(2)'} = [\mathbf{Y}_{E,\alpha}, \mathbf{Y}_{E,\beta}]$ with $\mathbf{Y}_{E,\alpha} \in \mathbb{C}^{N_E \times \min(N_2, K-N_1)}$ and $\mathbf{Y}_{E,\beta} \in \mathbb{C}^{N_E \times (K-N_T)^+}$. Also let $\mathbf{T}_{x2} = [\mathbf{T}_{x2,\alpha'}, \mathbf{T}_{x2,\beta}]$ with $\mathbf{T}_{x2,\alpha'} \in \mathbb{C}^{\Delta N \times \min(N_1, K-N_2)}$ and $\mathbf{T}_{x2,\beta} \in \mathbb{C}^{\Delta N \times (K-N_T)^+}$. Note that $\mathbf{Y}_{E,\alpha}$ depends on $\mathbf{T}_{x2,\alpha'}$, and $\mathbf{Y}_{E,\beta}$ depends on $\mathbf{T}_{x2,\beta}$.

Following a similar analysis leading to (99), we have that due to unknown $\mathbf{H}_{E,P',\perp} \in \mathbb{C}^{N_E \times N_2}$,

$$\begin{aligned}
\mathrm{DoF}(h(\mathbf{Y}_{E,\alpha}|\mathbf{Y}_1^{(2)}, \mathcal{C}_{1,2})) &= \mathrm{DoF}(h(\mathbf{Y}_{E,\alpha}|\mathcal{C}_{1,2})) \\
&= N_E \min(N_2, K-N_1) \quad (108)
\end{aligned}$$

which holds with or without the knowledge of $\mathbf{T}_{x2,\alpha'}$ due to the "fullness" of DoF (i.e., the above DoF equals to the number of entries in $\mathbf{Y}_{E,\alpha}$ even when conditioned on $\mathbf{T}_{x2,\alpha'}$).

For $K > N_T$, given $\mathbf{Y}_{E,\alpha}$ along with $\{\mathbf{Y}_1^{(2)}, \mathcal{C}_{1,2}\}$, we know $\mathbf{H}_E$ at high power if $\mathbf{T}_{x2,\alpha'}$ is also given.

Also for $K > N_T$, we now use the following (a', b')-partitions

$$\mathbf{Y}_{E,\beta} = [\mathbf{Y}_{E,\beta,a'}^T, \mathbf{Y}_{E,\beta,b'}^T]^T \qquad (109)$$

with $\mathbf{Y}_{E,\beta,a'} \in \mathbb{C}^{\min(N_E, \Delta N) \times (K-N_T)^+}$ and $\mathbf{Y}_{E,\beta,b'} \in \mathbb{C}^{(N_E-\Delta N)^+ \times (K-N_T)^+}$. It follows that due to unknown $\mathbf{T}_{x2,\beta} \in \mathbb{C}^{\Delta N \times (K-N_T)^+}$,

$$\begin{aligned}
&\mathrm{DoF}(h(\mathbf{Y}_{E,\beta,a'}|\mathbf{Y}_{E,\alpha}, \mathbf{Y}_1^{(2)}, \mathcal{C}_{1,2})) \\
&\geq \mathrm{DoF}(h(\mathbf{Y}_{E,\beta,a'}|\mathbf{T}_{x2,\alpha'}, \mathbf{Y}_{E,\alpha}, \mathbf{Y}_1^{(2)}, \mathcal{C}_{1,2})) \\
&= \mathrm{DoF}(h(\mathbf{Y}_{E,\beta,a'}|\mathbf{H}_E, \mathbf{X}_1, \mathbf{X}_{2,0}, \mathbf{N}_{x2})) \\
&= \min(N_E, \Delta N)(K-N_T)^+ \quad (110)
\end{aligned}$$

where the equality in "$\geq$" actually holds due to the fullness of DoF.

Next we show that for $K > N_T$ and $N_E > \Delta N$, $\mathbf{Y}_{E,\beta,b'}$ is determined with probability one at high power by $\{\mathbf{Y}_{E,\beta,a'}, \mathbf{Y}_{E,\alpha}, \mathbf{Y}_1^{(2)}, \mathcal{C}_{1,2}\}$.

With given $\mathbf{H}_{E,P'}$ (or equivalently $\mathbf{Y}_E^{(1)'}$) and the first ΔN columns of $\mathbf{Y}_{E,\alpha}$, we know that $\mathbf{H}_E \approx \bar{\mathbf{H}}_E + \mathbf{T}_E \mathbf{N}_E$ where only $\mathbf{T}_E \in \mathbb{C}^{N_E \times N_1}$ is unknown. Note that the row span of $\mathbf{N}_E$ is the left null space of $\mathbf{P}'' \doteq \begin{bmatrix} \mathbf{P}_1 & \mathbf{X}_{1,1} \\ \mathbf{P}_{2,1} & \mathbf{P}_{2,2} \end{bmatrix}$. Let $\mathbf{Y}_{E,\alpha'}$ be the last N_1 columns of $\mathbf{Y}_{E,\alpha}$. It follows that

$$\mathbf{Y}_{E,\alpha'} \approx \sigma(\bar{\mathbf{H}}_E + \mathbf{T}_E \mathbf{N}_E) \begin{bmatrix} \mathbf{X}_{1,2,\alpha'} \\ \mathbf{X}_{2,0,\alpha'} + \mathbf{N}_{x2}\mathbf{T}_{x2,\alpha'} \end{bmatrix} \quad (111)$$

where $\mathbf{T}_E \in \mathbb{C}^{N_E \times N_1}$ and $\mathbf{T}_{x2,\alpha'} \in \mathbb{C}^{\Delta N \times N_1}$ are the only (inter-dependent) unknowns. Specifically, $\mathbf{T}_E$ is a (uniquely valued) function of the unknown $\mathbf{T}_{x2,\alpha'}$ subject to given $\{\mathbf{Y}_{E,\alpha}, \mathbf{Y}_1^{(2)}, \mathcal{C}_{1,2}\}$. This is because of a property similar to that of $\mathbf{H}_E$ in (100).

Furthermore, we can write the (a', b')-partitions of $\mathbf{Y}_{E,\beta}$ as follows:

$$\mathbf{Y}_{E,\beta,a'} \approx \sigma(\bar{\mathbf{H}}_{E,a'} + \mathbf{T}_{E,a'}\mathbf{N}_E)\mathbf{X}'_{2,\beta}, \quad (112)$$

$$\mathbf{Y}_{E,\beta,b'} \approx \sigma(\bar{\mathbf{H}}_{E,b'} + \mathbf{T}_{E,b'}\mathbf{N}_E)\mathbf{X}'_{2,\beta}, \quad (113)$$

where

$$\mathbf{X}'_{2,\beta} = \begin{bmatrix} \mathbf{X}_{1,2,\beta} \\ \mathbf{X}_{2,0,\beta} + \mathbf{N}_{x2}\mathbf{T}_{x2,\beta} \end{bmatrix}. \quad (114)$$

We see that subject to $\mathbf{Y}_{E,\beta,a'}$ in addition to $\{\mathbf{Y}_{E,\alpha}, \mathbf{Y}_1^{(2)}, \mathcal{C}_{1,2}\}$, $\mathbf{T}_{E,a'}$ (and hence $\mathbf{T}_E$) is also a function of the unknown $\mathbf{T}_{x2,\beta}$. This makes $\mathbf{T}_E$ deterministic with probability one after $\{\mathbf{Y}_{E,\beta,a'}, \mathbf{Y}_{E,\alpha}, \mathbf{Y}_1^{(2)}, \mathcal{C}_{1,2}\}$ is given at high power. The above condition hence determines a unique $\mathbf{H}_E$, hence a unique $\mathbf{T}_{x2,\beta}$ and hence a unique $\mathbf{Y}_{E,\beta,b'}$. Namely,

$$\text{DoF}(h(\mathbf{Y}_{E,\beta,b'}|\mathbf{Y}_{E,\beta,a'}, \mathbf{Y}_{E,\alpha}, \mathbf{Y}_1^{(2)}, \mathcal{C}_{1,2})) = 0. \quad (115)$$

Therefore, adding (106), (108), (110) and (115) yields

$$\text{DoF}(h(\mathbf{Y}_1^{(2)}, \mathbf{Y}_E^{(2)'}|\mathcal{C}_{1,2})) = N_1(K - N_2) \\ + N_E \min(N_2, K - N_1) + \min(N_E, \Delta N)(K - N_T)^+. \quad (116)$$

2) For $h(\mathbf{Y}_2^{(2)}, \mathbf{Y}_E^{(2)'}|\mathbf{X}_2, \mathbf{H}_{2,1}, \mathbf{H}_{E,P'})$*:* Let $\mathcal{C}_{2,1} = \{\mathbf{X}_2, \mathbf{H}_{2,1}, \mathbf{H}_{E,P'}\}$. It follows from (80) that

$$\text{DoF}(h(\mathbf{Y}_2^{(2)}|\mathcal{C}_{2,1})) = \text{DoF}(h(\mathbf{Y}_2^{(2)}|\mathbf{H}_{2,1})) \\ = N_1(K - N_1). \quad (117)$$

Given $\mathbf{Y}_2^{(2)}$ and $\mathbf{H}_{2,1}$, $\mathbf{X}_1$ is unique (because of $N_2 \geq N_1$) at high power, and in this case, only $\mathbf{H}_{E,P',\perp}$ inside $\mathbf{H}_E$ shown in (83) is unknown subject to given $\mathcal{C}_{2,1}$ and high power. Recalling $\mathbf{Y}_E^{(2)'} = [\mathbf{Y}_{E,\alpha}, \mathbf{Y}_{E,\beta}]$, it follows that

$$\text{DoF}(h(\mathbf{Y}_{E,\alpha}|\mathbf{Y}_2^{(2)}, \mathcal{C}_{2,1})) = \text{DoF}(h(\mathbf{Y}_{E,\alpha}|\mathbf{X}_1, \mathbf{X}_2, \mathbf{H}_{E,P'})) \\ = N_E \min(N_2, K - N_1), \quad (118)$$

and for $K > N_T$,

$$\text{DoF}(h(\mathbf{Y}_{E,\beta}|\mathbf{Y}_{E,\alpha}, \mathbf{Y}_2^{(2)}, \mathcal{C}_{2,1})) \\ = \text{DoF}(h(\mathbf{Y}_{E,\beta}|\mathbf{X}_1, \mathbf{X}_2, \mathbf{H}_E)) = 0. \quad (119)$$

Therefore,

$$\text{DoF}(h(\mathbf{Y}_2^{(2)}, \mathbf{Y}_E^{(2)'}|\mathcal{C}_{2,1})) \\ = N_1(K - N_1) + N_E \min(N_2, K - N_1). \quad (120)$$

E. Summary of the Lower Bound

Combining the expressions in (94), (95), (105) and (116), we have

$$\text{SDoF}_{\text{new}}^{(2)} \doteq \text{DoF}(C_{\text{key},1,2,\text{low}}^{(2)}) \\ = N_1(2K - N_T) + \min(N_E, \Delta N)(K - N_T)^+ \\ - \min(N_E, N_T)(K - N_T)^+ \\ = \begin{cases} N_1(2K - N_T), & \text{C1;} \\ N_1(2K - N_T) - (N_E - \Delta N)(K - N_T)^+, & \text{C2;} \\ N_1[2K - N_T - (2K - 2N_T)^+], & \text{C3;} \end{cases} \quad (121)$$

where $K \geq N_2 \geq N_1$, $N_T = N_1 + N_2$, $\Delta N = N_2 - N_1$, and "C1, C2, C3" specify three (small, medium and large) regions of N_E as defined below (63).

We see that $\text{SDoF}_{\text{new}}^{(2)}$ is non-increasing function of N_E and non-decreasing function of K, which is consistent with intuitions. In the small and medium regions C1 and C2 of N_E (i.e., $0 \leq N_E < N_T$ specifically), $\text{SDoF}_{\text{new}}^{(2)}$ increases with K. But for the large region C3 of N_E (i.e., $N_E \geq N_T$), $\text{SDoF}_{\text{new}}^{(2)}$ increases with K for $N_2 \leq K \leq N_T$ and stays constant of K for $K \geq N_T$. For large N_E and K (i.e., $N_E \geq N_T$ and $K \geq N_T$), $\text{SDoF}_{\text{new}}^{(2)} = N_1 N_T$.

Combining the expressions in (94), (95), (105) and (120), we have

$$\text{DoF}(C_{\text{key},2,1,\text{low}}^{(2)}) \\ = N_1(2K - N_T) - \min(N_E, N_T)(K - N_T)^+. \quad (122)$$

It is clear that

$$\text{DoF}(C_{\text{key},1,2,\text{low}}^{(2)}) - \text{DoF}(C_{\text{key},2,1,\text{low}}^{(2)}) \\ \geq \min(N_E, \Delta N)(K - N_T)^+ \geq 0. \quad (123)$$

F. Analysis of the Upper Bound

We now analyze $C_{\text{key,up}}^{(2)}$ in (89). Similar to (127) and (133), we can write

$$C_{\text{key,up}}^{(2)} \approx -h(\mathbf{Y}_E^{(2)'}|\mathbf{H}_{E,P'}) \\ + h(\mathbf{Y}_1^{(2)}, \mathbf{Y}_E^{(2)'}|\mathbf{X}_1, \mathbf{H}_{1,2}, \mathbf{H}_{E,P'}) \\ + h(\mathbf{Y}_2^{(2)}, \mathbf{Y}_E^{(2)'}|\mathbf{X}_2, \mathbf{H}_{2,1}, \mathbf{H}_{E,P'}) \\ - h(\mathbf{Y}_1^{(2)}, \mathbf{Y}_2^{(2)}, \mathbf{Y}_E^{(2)'}|\mathbf{X}_1, \mathbf{X}_2, \mathbf{H}_{1,2}, \mathbf{H}_{2,1}, \mathbf{H}_{E,P'}). \quad (124)$$

The DoF of the first three terms are given by (105), (116) and (120). For the last term, we let $\mathcal{C}_T = \{\mathbf{X}_1, \mathbf{X}_2, \mathbf{H}_{1,2}, \mathbf{H}_{2,1}, \mathbf{H}_{E,P'}\}$. Then, we see $\text{DoF}(h(\mathbf{Y}_1^{(2)}|\mathcal{C}_T)) = 0$, $\text{DoF}(h(\mathbf{Y}_2^{(2)}|\mathbf{Y}_1^{(2)}, \mathcal{C}_T)) = 0$, and hence $\text{DoF}(h(\mathbf{Y}_1^{(2)}, \mathbf{Y}_2^{(2)}, \mathbf{Y}_E^{(2)'}|\mathcal{C}_T)) = \text{DoF}(h(\mathbf{Y}_E^{(2)'}|\mathcal{C}_T)) = \text{DoF}(h(\mathbf{Y}_{E,\alpha}|\mathcal{C}_T)) + \text{DoF}(h(\mathbf{Y}_{E,\beta}|\mathbf{Y}_{E,\alpha}, \mathcal{C}_T)) = N_E \Delta N$ where we have used $\text{DoF}(h(\mathbf{Y}_{E,\beta}|\mathbf{Y}_{E,\alpha}, \mathcal{C}_T)) = \text{DoF}(h(\mathbf{Y}_{E,\beta}|\mathbf{H}_E, \mathbf{X}_1, \mathbf{X}_2)) = 0$.

Combining the above results, one can verify that

$$\text{DoF}(C_{\text{key,up}}^{(2)}) = \text{DoF}(C_{\text{key,1,2,low}}^{(2)}) \tag{125}$$

which is given in (121). Therefore, $\text{DoF}(C_{\text{key}}^{(2)})$ is also given by (121).

G. Summary

Property 8: For a network of two users with $N_1 \leq N_2$, $N_T = N_1 + N_2$, and $\Delta N = N_2 - N_1$, we have found:

1) The lower and upper bounds on the phase-2 SDoF of the modified two-user ANECE coincide.
2) The total SDoF of the modified two-user ANECE equals $N_1 N_2 + \text{DoF}(C_{\text{key}}^{(2)})$ where $\text{DoF}(C_{\text{key}}^{(2)})$ is given by (125) or equivalently (121).
3) For $N_E < N_T$, $\text{SDoF}_{\text{new}}^{(2)}$ increases linearly with K.
4) For $N_E \geq N_T$, $\text{SDoF}_{\text{new}}^{(2)}$ increases linearly with K for $N_2 \leq K \leq N_T$, and equals $N_1 N_T$ for $K \geq N_T$.

IX. COMPARISON OF THE ORIGINAL TWO-USER ANECE AND THE MODIFIED TWO-USER ANECE

The phase-1 SDoF of the two versions is the same, i.e., $N_1 N_2$. We will only need to consider the phase-2 SDoF while assuming $N_1 < N_2$. (The two versions would be identical if $N_1 = N_2$.)

The phase-2 SDoF of the original version is given by (63) while that of the modified version is given by (121). For a fair comparison, we will let $K_2 = K - N_2 > 0$. It is easy to verify that

$$\text{SDoF}_{\text{new}}^{(2)} - \text{SDoF}_{\text{original}}^{(2)} = N_1(N_2 - N_1) \tag{126}$$

for all cases of $N_E \geq 0$ and $K \geq N_2$. It is interesting to observe that this gap equals the number of entries in $\mathbf{X}_{1,1}$ in the modified two-user ANECE. See $\mathbf{X}_{1,1}$ in (84) for example.

Property 9: The phase-2 SDoF of the modified two-user ANECE (see (121)) is $N_1 \Delta N$ larger than that (see (63)) of the original two-user ANECE, subject to $K_2 = K - N_2 \geq 0$.

X. CONCLUSION

We have shown novel insights into the (pair-wise) SDoF of a multi-user multi-antenna full-duplex wireless network using multi-user multi-antenna ANECE against Eve with any number of antennas. These insights are based on the DoF of lower and upper bounds on the pair-wise SKC between users using the observations from the multi-user ANECE. In many cases, the lower and upper bounds have the same DoF, which hence reveal the exact SDoF. In particular, the obtained lower and upper bounds on SDoF for two-user ANECE coincide, regardless of

the numbers of antennas on users and Eve, and regardless of the number of time slots used for phase 2 in each coherence period. These SDoF bounds also coincide for multi-user ANECE when every user has the same number of antennas and the number of time slots used for phase 2 is no larger than the number of antennas. And for a three-user network, all-user ANECE has been shown to have advantages over pair-wise ANECE in both phases 1 and 2.

We have also shown that when two users have different numbers of antennas, a modified two-user ANECE, for which each user applies a square-shaped nonsingular pilot matrix, has a larger SDoF than the original two-user ANECE for which the pilot matrices from both users consume the same number of time-slots. The proven result is also explainable intuitively.

The lower and upper bounds on the pair-wise SKC used in this article are the same as those established by Maurer [30] and others [29]. The matching DoF of the lower and upper bounds in many cases as shown in this article is a pleasant discovery. A similar phenomenon was also found recently in [33] for two-user half-duplex schemes.

This article has given a more complete picture of the pair-wise SDoF of multi-user ANECE. But many questions are still open, which include how much correlation there is between the keys generated by different pairs of users for phase 2, how undesirable this potential correlation could be in practice, and whether there is another multi-user scheme that can yield a higher SDoF. Even for the two-user case, it is not yet formally proven whether the modified two-user ANECE is already optimal among all possible schemes in terms of SDoF. We hope that good answers to the above questions will be found in the near future.

APPENDIX

A. Analysis of $C_{\text{key},i,j,\text{up}}^{(2)}$ in (19)

In this section, we analyse the DoF of $C_{\text{key},i,j,\text{up}}^{(2)}$ shown in (19). It follows from (19) that with a high power in phase 1,

$$C_{\text{key},i,j,\text{up}}^{(2)} \approx \mathbb{I}(\mathbf{X}_i, \mathbf{Y}_i^{(2)}; \mathbf{X}_j, \mathbf{Y}_j^{(2)}|\mathcal{C}) \tag{127}$$

with $\mathcal{C} = \{\mathbf{H}_i, \mathbf{H}_j, \mathbf{H}_{E,P}, \mathbf{Y}_E^{(2)}\}$. Furthermore,

$$\begin{aligned}
C_{\text{key},i,j,\text{up}}^{(2)} &\approx h(\mathbf{X}_i, \mathbf{Y}_i^{(2)}|\mathcal{C}) - h(\mathbf{X}_i, \mathbf{Y}_i^{(2)}|\mathbf{X}_j, \mathbf{Y}_j^{(2)}, \mathcal{C}) \\
&= h(\mathbf{X}_i|\mathcal{C}) + h(\mathbf{Y}_i^{(2)}|\mathbf{X}_i, \mathcal{C}) \\
&\quad - h(\mathbf{X}_i|\mathbf{X}_j, \mathbf{Y}_j^{(2)}, \mathcal{C}) - h(\mathbf{Y}_i^{(2)}|\mathbf{X}_i, \mathbf{X}_j, \mathbf{Y}_j^{(2)}, \mathcal{C}).
\end{aligned} \tag{128}$$

Using $h(A|B,C) = h(A,B|C) - h(B|C) = h(B|A,C) + h(A|C) - h(B|C)$, each of the four terms in the above can be rewritten as follows:

$$\begin{aligned}
h(\mathbf{X}_i|\mathcal{C}) &= h(\mathbf{X}_i|\mathbf{H}_{E,P}, \mathbf{Y}_E^{(2)}) = h(\mathbf{Y}_E^{(2)}|\mathbf{X}_i, \mathbf{H}_{E,P}) \\
&\quad + h(\mathbf{X}_i|\mathbf{H}_{E,P}) - h(\mathbf{Y}_E^{(2)}|\mathbf{H}_{E,P}) \\
&= h(\mathbf{Y}_E^{(2)}|\mathbf{X}_i, \mathbf{H}_{E,P}) + h(\mathbf{X}_i) - h(\mathbf{Y}_E^{(2)}|\mathbf{H}_{E,P}),
\end{aligned} \tag{129}$$

$$h(\mathbf{Y}_i^{(2)}|\mathbf{X}_i, \mathcal{C}) = h(\mathbf{Y}_i^{(2)}|\mathbf{X}_i, \mathbf{H}_i, \mathbf{H}_{E,P}, \mathbf{Y}_E^{(2)})$$
$$= h(\mathbf{Y}_i^{(2)}, \mathbf{Y}_E^{(2)}|\mathbf{X}_i, \mathbf{H}_i, \mathbf{H}_{E,P}) - h(\mathbf{Y}_E^{(2)}|\mathbf{X}_i, \mathbf{H}_i, \mathbf{H}_{E,P})$$
$$= h(\mathbf{Y}_i^{(2)}, \mathbf{Y}_E^{(2)}|\mathbf{X}_i, \mathbf{H}_i, \mathbf{H}_{E,P}) - h(\mathbf{Y}_E^{(2)}|\mathbf{X}_i, \mathbf{H}_{E,P}), \tag{130}$$

$$h(\mathbf{X}_i|\mathbf{X}_j, \mathbf{Y}_j^{(2)}, \mathcal{C}) = h(\mathbf{X}_i|\mathbf{X}_j, \mathbf{Y}_j^{(2)}, \mathbf{H}_j, \mathbf{H}_{E,P}, \mathbf{Y}_E^{(2)})$$
$$= h(\mathbf{Y}_j^{(2)}, \mathbf{Y}_E^{(2)}|\mathbf{X}_i, \mathbf{X}_j, \mathbf{H}_j, \mathbf{H}_{E,P}) + h(\mathbf{X}_i|\mathbf{X}_j, \mathbf{H}_j, \mathbf{H}_{E,P})$$
$$- h(\mathbf{Y}_j^{(2)}, \mathbf{Y}_E^{(2)}|\mathbf{X}_j, \mathbf{H}_j, \mathbf{H}_{E,P})$$
$$= h(\mathbf{Y}_j^{(2)}, \mathbf{Y}_E^{(2)}|\mathbf{X}_i, \mathbf{X}_j, \mathbf{H}_j, \mathbf{H}_{E,P}) + h(\mathbf{X}_i)$$
$$- h(\mathbf{Y}_j^{(2)}, \mathbf{Y}_E^{(2)}|\mathbf{X}_j, \mathbf{H}_j, \mathbf{H}_{E,P}) \tag{131}$$

$$h(\mathbf{Y}_i^{(2)}|\mathbf{X}_i, \mathbf{X}_j, \mathbf{Y}_j^{(2)}, \mathcal{C})$$
$$= h(\mathbf{Y}_i^{(2)}|\mathbf{X}_i, \mathbf{X}_j, \mathbf{Y}_j^{(2)}, \mathbf{H}_i, \mathbf{H}_j, \mathbf{H}_{E,P}, \mathbf{Y}_E^{(2)})$$
$$= h(\mathbf{Y}_i^{(2)}, \mathbf{Y}_j^{(2)}, \mathbf{Y}_E^{(2)}|\mathbf{X}_i, \mathbf{X}_j, \mathbf{H}_i, \mathbf{H}_j, \mathbf{H}_{E,P})$$
$$- h(\mathbf{Y}_j^{(2)}, \mathbf{Y}_E^{(2)}|\mathbf{X}_i, \mathbf{X}_j, \mathbf{H}_i, \mathbf{H}_j, \mathbf{H}_{E,P})$$
$$= h(\mathbf{Y}_i^{(2)}, \mathbf{Y}_j^{(2)}, \mathbf{Y}_E^{(2)}|\mathbf{X}_i, \mathbf{X}_j, \mathbf{H}_i, \mathbf{H}_j, \mathbf{H}_{E,P})$$
$$- h(\mathbf{Y}_j^{(2)}, \mathbf{Y}_E^{(2)}|\mathbf{X}_i, \mathbf{X}_j, \mathbf{H}_j, \mathbf{H}_{E,P}). \tag{132}$$

Using (129)–(132) into (128) yields

$$C_{\text{key},i,j,\text{up}}^{(2)} \approx -h(\mathbf{Y}_E^{(2)}|\mathbf{H}_{E,P})$$
$$+ h(\mathbf{Y}_i^{(2)}, \mathbf{Y}_E^{(2)}|\mathbf{X}_i, \mathbf{H}_i, \mathbf{H}_{E,P})$$
$$+ h(\mathbf{Y}_j^{(2)}, \mathbf{Y}_E^{(2)}|\mathbf{X}_j, \mathbf{H}_j, \mathbf{H}_{E,P})$$
$$- h(\mathbf{Y}_i^{(2)}, \mathbf{Y}_j^{(2)}, \mathbf{Y}_E^{(2)}|\mathbf{X}_i, \mathbf{X}_j, \mathbf{H}_i, \mathbf{H}_j, \mathbf{H}_{E,P}). \tag{133}$$

The DoF of the first term in (133) is given by (48) in Section VI.B.1, and the DoFs of the second and third terms follow (55) in Section VI.B.2. Next we only need to analyze the fourth term in (133).

1) Analysis of $h(\mathbf{Y}_i^{(2)}, \mathbf{Y}_j^{(2)}, \mathbf{Y}_E^{(2)}|\mathbf{X}_i, \mathbf{X}_j, \mathbf{H}_i, \mathbf{H}_j, \mathbf{H}_{E,P})$: Without loss of generality, let us consider $h(\mathbf{Y}_1^{(2)}, \mathbf{Y}_2^{(2)}, \mathbf{Y}_E^{(2)}|\mathcal{C}_2')$ with $\mathcal{C}_2' = \{\mathbf{X}_1, \mathbf{X}_2, \mathbf{H}_1, \mathbf{H}_2, \mathbf{H}_{E,P}\}$.

Note that $\mathbf{Y}_1^{(2)} = \sigma(\mathbf{H}_{1,2}\mathbf{X}_2 + \mathbf{H}_{1,3}\mathbf{X}_3 + \cdots + \mathbf{H}_{1,M}\mathbf{X}_M) + \mathbf{W}_1^{(2)}$ where the first term $\mathbf{H}_{1,2}\mathbf{X}_2$ is constant given $\mathcal{C}_2'$. It follows that

$$\text{DoF}(h(\mathbf{Y}_1^{(2)}|\mathcal{C}_2')) = \min(N_1, N_T - N_1 - N_2)K_2. \tag{134}$$

Let $\mathbf{X}_{(1,2)}$ be the vertical stack of $\mathbf{X}_3, \cdots, \mathbf{X}_M$. Then given a high power, $\mathbf{Y}_1^{(2)}$ and $\mathcal{C}_2'$, we have $\mathbf{X}_{(1,2)} \approx \mathbf{X}_{1,2,0} + \mathbf{N}_{1,2}\mathbf{T}_{1,2}$ where $\mathbf{X}_{1,2,0} \in \mathbb{C}^{(N_T - N_1 - N_2) \times K_2}$ and $\mathbf{N}_{1,2} \in \mathbb{C}^{(N_T - N_1 - N_2) \times (N_T - 2N_1 - 2N_2)^+}$ are constant while $\mathbf{T}_{1,2} \in \mathbb{C}^{(N_T - 2N_1 - 2N_2)^+ \times K_2}$ is arbitrary. The column span of $\mathbf{N}_{1,2}$ is the right null space of $[\mathbf{H}_{1,3}, \cdots, \mathbf{H}_{1,M}]$. It follows that

$$\text{DoF}(h(\mathbf{Y}_2^{(2)}|\mathbf{Y}_1^{(2)}, \mathcal{C}_2'))$$
$$= \min(N_2, (N_T - 2N_1 - N_2)^+)K_2. \tag{135}$$

Given a high power, $\mathbf{Y}_1^{(2)}$, $\mathbf{Y}_2^{(2)}$ and $\mathcal{C}_2'$, we have $\mathbf{X}_{(1,2)} \approx \bar{\mathbf{X}}_{1,2,0} + \bar{\mathbf{N}}_{1,2}\bar{\mathbf{T}}_{1,2}$ where $\bar{\mathbf{X}}_{1,2,0} \in \mathbb{C}^{(N_T - N_1 - N_2) \times K_2}$, $\bar{\mathbf{N}}_{1,2} \in$

$\mathbb{C}^{(N_T - N_1 - N_2) \times (N_T - 2N_1 - 2N_2)^+}$ are constant while $\bar{\mathbf{T}}_{1,2} \in \mathbb{C}^{(N_T - 2N_1 - 2N_2)^+ \times K_2}$ is arbitrary. The column span of $\bar{\mathbf{N}}_{1,2}$ is the right null space of $\begin{bmatrix} \mathbf{H}_{1,3} & \cdots & \mathbf{H}_{1,M} \\ \mathbf{H}_{2,3} & \cdots & \mathbf{H}_{2,M} \end{bmatrix}$.

Recall the (α, β)-partitions $\mathbf{Y}_E^{(2)} = [\mathbf{Y}_{E,\alpha}^{(2)}, \mathbf{Y}_{E,\beta}^{(2)}]$ as in (42). We know that due to the unknown $\mathbf{H}_{E,P,\perp}$,

$$\text{DoF}(h(\mathbf{Y}_{E,\alpha}^{(2)}|\mathbf{Y}_1^{(2)}, \mathbf{Y}_2^{(2)}, \mathcal{C}_2')) = N_E \min(N_{\min}, K_2). \tag{136}$$

For $K_2 > N_{\min}$, following a similar analysis leading to (53) and (54), we have

$$\text{DoF}(h(\mathbf{Y}_{E,\beta}^{(2)}|\mathbf{Y}_{E,\alpha}^{(2)}, \mathbf{Y}_1^{(2)}, \mathbf{Y}_2^{(2)}, \mathcal{C}_2'))$$
$$= \min(N_E, (N_T - 2N_1 - 2N_2)^+)\Delta K_2. \tag{137}$$

Adding up (134), (135), (136) and (137), but with N_1 and N_2 replaced by N_i and N_j, yields

$$\text{DoF}(h(\mathbf{Y}_i^{(2)}, \mathbf{Y}_j^{(2)}, \mathbf{Y}_E^{(2)}|\mathbf{X}_i, \mathbf{X}_j, \mathbf{H}_i, \mathbf{H}_j, \mathbf{H}_{E,P}))$$
$$= K_2 \min(N_i, N_T - N_i - N_j)$$
$$+ K_2 \min(N_j, (N_T - 2N_i - N_j)^+)$$
$$+ N_E \min(N_{\min}, K_2)$$
$$+ \Delta K_2 \min(N_E, (N_T - 2N_i - 2N_j)^+). \tag{138}$$

ACKNOWLEDGMENT

REFERENCES

[1] K. Kolodziej, B. Perry, and J. Herd, "In-band full-duplex technology: Techniques and systems survey," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 7, pp. 3025–3041, Jul. 2019.

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

[3] H. Poor and R. Schaefer, "Wireless physical layer security," *Proc. Natl. Acad. Sci.*, vol. 114, no. 1, pp. 19–26, 2017.

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

[5] T.-Y. Liu, P. Mukherjee, S. Ulukus, S.-C. Lin, and Y.-W. 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.

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

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

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

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

[10] M. Abedi, N. Mokari, H. Saeedi, and H. Yanikomeroglu, "Robust resource allocation to enhance physical layer security in systems with

full-duplex receivers: Active adversary," *IEEE Trans. Wireless Commun.*, vol. 16, no. 2, pp. 885–899, Feb. 2017.

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

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

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

[14] Y. Hua, Q. Zhu, and R. Sohrabi, "Fundamental properties of full-duplex radio for secure wireless communications," Accessed: Nov. 27, 2017. [Online]. Available: http://arxiv.org/abs/1711.10001

[15] I. Zabir, A. Maksud, G. Chen, B. Sadler, and Y. Hua, "Secrecy of multi-antenna transmission with full-duplex user in the presence of randomly located eavesdroppers," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 2060–2075, 2020.

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

[17] N. Mahmood, I. Ansari, P. Popovski, P. Mogensen, and K. Qaraqe, "Physical-layer security with full-duplex transceivers and multiuser receiver at eve," *IEEE Trans. Commun.*, vol. 65, no. 10, pp. 4392–4405, Oct. 2017.

[18] L. Li, Z. Chen, A. Petropulu, and J. Fang, "Linear precoder design for an MIMO Gaussian wiretap channel with full-duplex source and destination nodes," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 2, pp. 421–436, Feb. 2018.

[19] 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.

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

[21] 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.

[22] C.-W. Huang, T.-H. Chang, X. Zhou, and Y.-W. Hong, "Two-way training for discriminatory channel estimation in wireless MIMO systems," *IEEE Trans. Signal Process.*, vol. 61, no. 10, pp. 2724–2738, May 2013.

[23] J. Yang, S. Xie, X. Zhou, R. Yu, and Y. Zhang, "A semiblind two-way training method for discriminatory channel estimation in MIMO systems," *IEEE Trans. Commun.*, vol. 62, no. 7, pp. 2400–2410, 2014.

[24] 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.

[25] 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.

[26] Q. Zhu, S. Wu, and Y. Hua, "Optimal pilots for anti-eavesdropping channel estimation," *IEEE Trans. Signal Process.*, vol. 68, pp. 2629–2644, 2020.

[27] S. Wu and Y. Hua, "Total secrecy from anti-eavesdropping channel estimation," *IEEE Trans. Signal Process.*, vol. 70, pp. 1088–1103, 2022.

[28] 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.

[29] M. Bloch and J. Barros, *Physical-Layer Security*. Cambridge, UK: Cambridge Univ. Press, 2011.

[30] U. Maurer, "Secret key agreement by public discussion from common information," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 733–742, May 1993.

[31] T. Cover and J. Thomas, *Elements of Information Theory*, vol. 2. Hoboken, NJ, USA: John Wiley & Sons, 2006.

[32] L. Y. Liang and H. Poor, "A unified framework for key agreement over wireless fading channel," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 2, pp. 480–490, Apr. 2012.

[33] Y. Hua, "Generalized channel probing and generalized pre-processing for secret key generation," *IEEE Trans. Signal Process.*, vol. 71, pp. 1067–1082, 2023.

Yingbo Hua (Fellow, IEEE) is a Distinguished Professor with the University of California at Riverside, Riverside, CA, USA.

Qingpeng Liang (Member, IEEE) is an Assistant Professor with Southwest Jiaotong University, Sichuan, China.

Md Saydur Rahman (Graduate Student Member, IEEE) is a Graduate Student with the University of California at Riverside, Riverside, CA, USA.