entropy

MDPI

Article

Secure Retrospective Interference Alignment

Mohamed Seif, Ravi Tandon * and Ming Li

Department of Electrical and Computer Engineering, University of Arizona, Tucson, AZ 85721, USA; mseif@email.arizona.edu (M.S.); lim@email.arizona.edu (M.L.)

* Correspondence: tandonr@email.arizona.edu

Received: 21 September 2019; Accepted: 24 October 2019; Published: 7 November 2019

Abstract: In this paper, the K-user interference channel with secrecy constraints is considered with delayed channel state information at transmitters (CSIT). We propose a novel secure retrospective interference alignment scheme in which the transmitters carefully mix information symbols with artificial noises to ensure confidentiality. Achieving positive secure degrees of freedom (SDoF) is challenging due to the delayed nature of CSIT, and the distributed nature of the transmitters. Our scheme works over two phases: Phase one, in which each transmitter sends information symbols mixed with artificial noises, and repeats such transmission over multiple rounds. In the next phase, each transmitter uses the delayed CSIT of the previous phase and sends a function of the net interference and artificial noises (generated in previous phase), which is simultaneously useful for all receivers. These phases are designed to ensure the decodability of the desired messages while satisfying the secrecy constraints. We present our achievable scheme for three models, namely: (1) K-user interference channel with confidential messages (IC-CM), and we show that $\frac{1}{2}(\sqrt{K}-6)$ SDoF is achievable; (2) K-user interference channel with an external eavesdropper (IC-EE); and (3) K-user IC with confidential messages and an external eavesdropper (IC-CM-EE). We show that for the K-user IC-EE, $\frac{1}{2}(\sqrt{K}-3)$ SDoF is achievable, and for the K-user IC-CM-EE, $\frac{1}{2}(\sqrt{K}-6)$ is achievable. To the best of our knowledge, this is the first result on the K-user interference channel with secrecy constrained models and delayed CSIT that achieves an SDoF which scales with $\sqrt{K}$, square-root of number of users.

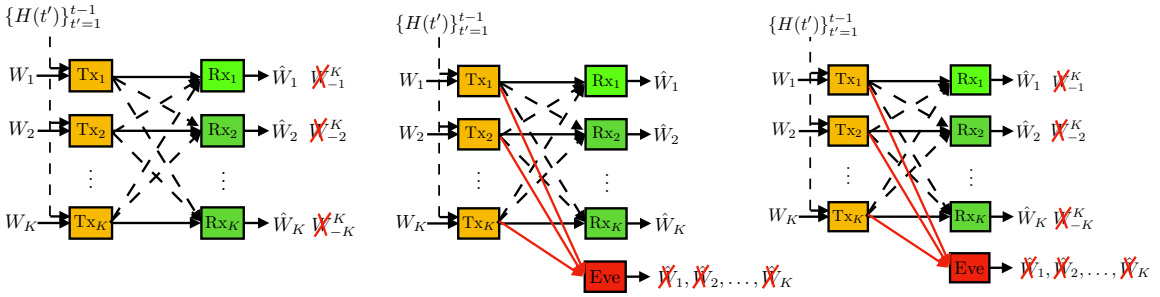
Keywords: interference channel; secure retrospective interference alignment; secure degrees of freedom (SDoF); delayed CSIT

1. Introduction

Delayed channel state information at transmitters (CSIT) can impact the spectral efficiency of wireless networks, and this problem has received significant recent attention. Maddah Ali and Tse in [1] studied the delayed CSIT model for the K-user multiple-input single-output (MISO) broadcast channel, and showed that the optimal sum degrees of freedom (DoF) is given by $K/(1+\frac{1}{2}+\cdots+\frac{1}{K})$ which is strictly greater than one DoF (with no CSIT) and less than K DoF (with perfect CSIT). For the K-user single-input single-output (SISO) X network, $\frac{K^2}{2K-1}$ is maximum DoF with perfect CSIT [2]. In [3], Ghasemi et al. devised a transmission scheme for the X channel with delayed CSIT, and showed that for the K-user SISO X channel under delayed CSIT, $\frac{4}{3}-\frac{2}{3(3K-1)}$ DoF are achievable. The problem of delayed CSIT for interference channels has been studied in several works [3–7]. The main drawback of these schemes is that the achievable DoF *does not* scale with the number of users. In a recent work [8], a novel transmission scheme for the K-user SISO interference channel is presented which achieves $\frac{\lfloor\sqrt{K}\rfloor}{2}$ DoF almost surely under delayed CSIT model. The result in [8] is particularly interesting, as it shows that the sum DoF for the K-user interference channel *does* scale with $\sqrt{K}$, even with delayed CSIT.

Another important aspect in wireless networks is ensuring secure communication between transmitters and receivers. Many seminal works in the literature (see comprehensive surveys [9–11]) studied the secure capacity regions for multi-user settings such as wiretap channel, broadcast, and interference channels. Since the exact secure capacity regions for many multi-user networks are not known, secure degrees of freedom (SDoF) for a variety of models have been studied in [12–18]. More specifically, for the K-user MISO broadcast channel with confidential messages, the authors in [18] showed that the optimal sum SDoF with delayed CSIT is given by $K/(1+\frac{1}{2}+\cdots+\frac{1}{K}+\frac{K-1}{K})$. The achievability scheme is based on a modification of the (insecure) Maddah Ali and Tse's scheme in [1] along with a key generation and exploitation method which uses delayed CSIT. The expression of the sum SDoF in [18] is almost the same as in [1] except a penalty term due to secrecy constraints. For the K-user SISO interference channel with confidential messages under perfect CSIT, Xie and Ulukus showed in [19] that the optimal sum SDoF is $\frac{K(K-1)}{2K-1}$. In [20], for the K-user interference channel with an external eavesdropper, it has been showed $\frac{K-1}{2}$ SDoF is optimal for the interference channel with no eavesdropper CSIT. Also, there are various other works for different CSIT assumptions such as MIMO wiretap channel with no eavesdropper CSIT [21], broadcast channel with alternating CSIT [22].

In this work, we consider the K-user SISO interference channel with secrecy constraints and delayed CSIT. More specifically, we study three channel models (see Figure 1), namely: (1) K-user interference channel with confidential messages, (2) K-user interference channel with an external eavesdropper, and (3) K-user interference channel with confidential messages and an external eavesdropper. We focus on answering the following fundamental questions regarding these channel models: (a) is positive SDoF achievable for the interference channel with delayed CSIT?, and (b) if yes, then how does the SDoF scale with K, the number of users?



(a) K-user IC-CM. (b) K-user IC-EE. (c) K-user IC-CM-EE.

Figure 1. K-user interference channel (IC) with secrecy and delayed CSIT. The model in (**a**) has confidential message (CM) constraints, model in (**b**) assumes the presence of an external eavesdropper (EE), and the model in (**c**) has both confidentiality and secrecy constraints (CM and EE).

Contributions: We answer the above questions for all the three channel models in the affirmative by showing that positive SDoF is indeed achievable for all these models, for a large number of users, K. We show that for the K-user interference channel with confidential messages (IC-CM), $\frac{1}{2}(\sqrt{K}-6)$ SDoF is achievable. Also, we show that for the K-user interference channel with an external eavesdropper (IC-EE), $\frac{1}{2}(\sqrt{K}-3)$ SDoF is achievable, and for the K-user with confidential messages and an external eavesdropper (IC-CM-EE), $\frac{1}{2}(\sqrt{K}-6)$ is achievable. In Table 1, we summarize the main results for the K-user IC under various secrecy constraints, and different CSIT assumptions (i.e., perfect CSIT, delayed CSIT and no CSIT).

Table 1. Summary of results on the K-user interference channel with different secrecy constraints and channel state information at transmitters (CSIT) models. The highlighted results are from this paper. These results show that sum secure degrees of freedom (SDoF) scales with $\sqrt{K}$.

	Perfect CSIT	Delayed CSIT	No CSIT
No Secrecy	$\frac{K}{2}$	$> \frac{1}{2}(\sqrt{K}-1)$	1
Confidential Messages	$\frac{K(K-1)}{2K-1}$		0
External Eavesdropper	$\frac{K-1}{2}$		0
Confidential Messages and Eavesdropper	$\frac{K-1}{2}$	$> \frac{1}{2}(\sqrt{K}-6)$	0

These results highlight the fact that in presence of delayed CSIT, there is negligible DoF scaling loss due to the secrecy constraints in the network compared to the no secrecy case [8]. Our main contribution is a novel *secure* retrospective interference alignment scheme, that is specialized for the interference channel with delayed CSIT. Our transmission scheme is inspired by the work of [8] in terms of the organization of the transmission phases. One of the main differences is that the transmitters mix their information symbols with artificial noises so that the signals at each unintended receiver are completely immersed in the space spanned by artificial noise. However, this mixing must be done with only delayed CSIT, and it should also allow successful decoding at the respective receiver. Our scheme works over two phases: Phase one, in which each transmitter sends information symbols mixed with artificial noises, and repeats such transmission over multiple rounds. Subsequently, in the next phase, each transmitter carefully sends a function of the net interference and artificial noises (generated in previous phase), which is simultaneously useful to all receivers. The equivocation analysis of the proposed scheme is non-trivial due to the repetition and retransmission strategies employed by the transmitters.

Organization of the paper: The rest of the paper is organized as follows. Section 2 describes the system models. The main results and discussions are presented in Section 3. Section 4 provides the achievable scheme under delayed CSIT and confidential messages. Sections 5 and 6 discuss two other secrecy constraints: (1) K-user interference channel with an external eavesdropper (IC-EE), and (2) K-user interference channel with confidential messages and an external eavesdropper (IC-CM-EE), respectively. We conclude the paper and discuss the future directions in Section 7. Finally, the detailed proofs are deferred to the Appendices.

Notations: Boldface uppercase letters denote matrices (e.g., $\mathbf{A}$), boldface lowercase letters are used for vectors (e.g., $\mathbf{a}$), we denote scalars by non-boldface lowercase letters (e.g., x), and sets by capital calligraphic letters (e.g., $\mathcal{X}$). The set of natural numbers, integer numbers, real numbers, and complex numbers are denoted by $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$, and $\mathbb{C}$, respectively. For a general matrix $\mathbf{A}$ with dimensions of $M \times N$, $\mathbf{A}^T$, and $\mathbf{A}^H$ denote the transpose and Hermitian transpose of $\mathbf{A}$, respectively. We denote the partitioned matrix of two matrices $\mathbf{A}_{L\times N}$ and $\mathbf{B}_{L\times M}$ as $(\mathbf{A}:\mathbf{B})$. We denote the identity matrix of the order M with $\mathbf{I}_M$. Let $h(\mathbf{x})$ denote the differential entropy of a random vector $\mathbf{x}$, and $I(\mathbf{x};\mathbf{y})$ denote the mutual information between two random vectors $\mathbf{x}$ and $\mathbf{y}$. We denote a complex-Gaussian distribution with a mean μ and a variance σ^2 by $\mathcal{CN}(\mu,\sigma^2)$. For rounding operations on a variable x, we use $\lfloor x \rfloor$ as the floor rounding operator on x and $\lceil x \rceil$ as the ceiling rounding operator on x.

2. System Model

We consider the K-user interference channel with secrecy constraints and delayed CSIT (shown in Figure 1). The input–output relationship at time slot t is

$$y_k(t) = h_{kk}(t)x_k(t) + \sum_{j=1, j\neq k}^{K} h_{kj}(t)x_j(t) + n_k(t), \tag{1}$$

$$z(t) = \sum_{j=1}^{K} g_j(t)x_j(t) + n_z(t), \tag{2}$$

where $y_k(t)$ is the signal received at receiver k at time t, $h_{kj}(t) \sim \mathcal{CN}(0,1)$ is the channel coefficient at time t between transmitter j and receiver k, and $x_k(t)$ is the transmitted signal from transmitter k at time t with an average power constraint $\mathbb{E}\{|x_k(t)|^2\} \leq P$. The additive noise $n_k(t) \sim \mathcal{CN}(0,1)$ at receiver k is independent and identically distributed (i.i.d.) across users and time. $z(t)$ is the received signal at the eavesdropper at time t, $g_j(t) \sim \mathcal{CN}(0,1)$ is the channel coefficient at time t between transmitter j and the external eavesdropper, and $n_z(t) \sim \mathcal{CN}(0,1)$ is the additive noise at the eavesdropper. The channel coefficients are assumed to be i.i.d. across time and users. We assume perfect CSI at all the receivers. We further assume that the CSIT is delayed, i.e., CSI is available at each transmitter after one time slot without error. Also, we assume that the external eavesdropper's CSI is not available at the transmitters (i.e., no eavesdropper CSIT).

Let $R_k = \frac{\log_2(|\mathcal{W}_k|)}{\tau}$ denote the rate of message W_k intended for receiver k, where $|\mathcal{W}_k|$ is the cardinality of the kth message. A $(2^{\tau R_1}, 2^{\tau R_2}, \ldots, 2^{\tau R_K}, \tau)$ code is described by the set of encoding and decoding functions as follows: the set of encoders at the transmitters are given as: $\{\psi_t^{(k)} : W_k \times \{H(t')\}_{t'=1}^{t-1} \to x_k(t)\}_{t=1}^{\tau}, \forall k = 1, \ldots, K$, where the message W_k is uniformly distributed over the set $\mathcal{W}_k$, and $H(t') \triangleq \{h_{kj}(t')\}_{k=1,j=1}^{K}$ is the set of all channel gains at time t'. The transmitted signal from transmitter k at time slot t is given as: $x_k(t) = \psi_t(W_k, \{H(t')\}_{t'=1}^{t-1})$. The decoding function at receiver k is given by the following mapping: $\phi^{(k)} : y_k^{(\tau)} \times \{H(t)\}_{t=1}^{\tau} \to W_k$, and the estimate of the message at receiver k is defined as: $\hat{W}_k = \phi^{(k)}(\{y_k(t), H(t)\}_{t=1}^{\tau})$. The rate tuple $(R_1, \ldots, R_K)$ is achievable if there exists a sequence of codes which satisfy the decodability constraints at the receivers, i.e.,

$$\lim_{\tau\to\infty} \sup \text{Prob}\left[\hat{W}_k \neq W_k\right] \leq \epsilon_\tau, \forall k = 1, \ldots, K \tag{3}$$

and the corresponding secrecy requirement is satisfied. We consider three different secrecy requirements:

1. IC-CM, Figure 1a, all unintended messages are kept secure against each receiver, i.e.,

$$\lim_{\tau\to\infty} \sup \frac{1}{\tau} I\left(W_{-k}^K; y_k^{(\tau)} | W_k, \Omega\right) \leq \epsilon_\tau, \forall k = 1, \ldots, K, \tag{4}$$

   where $\epsilon_\tau \to 0$ as $\tau \to \infty$, $W_{-k}^K \triangleq \{W_1, W_2, \ldots, W_K\} \backslash \{W_k\}$, and $\Omega \triangleq \{H(t), G(t)\}_{t=1}^{\tau}$ is the set of all channel gains (i.e., legitimate receivers and external eavesdropper) over the channel uses.
2. IC-EE, Figure 1b, all of the messages are kept secure against the external eavesdropper, i.e.,

$$\lim_{\tau\to\infty} \sup \frac{1}{\tau} I\left(W_1, W_2, \ldots, W_K; z^{(\tau)} | \Omega\right) \leq \epsilon_\tau, \forall k = 1, \ldots, K. \tag{5}$$

3. IC-CM-EE, Figure 1c, all of the messages are kept secure against both the $K-1$ unintended receivers and the external eavesdropper, i.e., we impose both secrecy constraints in Equations (4) and (5).

The supremum of the achievable sum rate, $R_s \triangleq \sum_{k=1}^{K} R_k$, is defined as the secrecy sum capacity C_s. The optimal sum secure degrees of freedom (SDoF*) is then defined as follows:

$$\text{SDoF}^* \triangleq \lim_{P\to\infty} \frac{C_S}{\log(P)}. \tag{6}$$

SDoF* represents the optimal scaling of the secrecy capacity with $\log(P)$, where P is the transmitted power, i.e., it is the pre-log factor of the secrecy capacity at high SNR.

In the next Section, we present our main results on the achievable sum SDoF with the three different secrecy constraints and delayed CSIT.

3. Main Results

Theorem 1. *For the K-user IC-CM with delayed CSIT, the following secure sum degrees of freedom is achievable:*

$$\text{SDoF}^*_{\text{IC-CM}} \geq \text{SDoF}^{\text{ach.}}_{\text{IC-CM}} = \frac{KR(K-R-2)}{(K-1)\times[R(R+1)+K]}, \tag{7}$$

where,

$$R = \left\lfloor \frac{-K + K\times\sqrt{1+\frac{(K-1)(K-2)}{K}}}{K-1} \right\rfloor. \tag{8}$$

We next simplify the above expression and present a lower bound on the achievable SDoF. Using this lower bound, we observe that the achievable SDoF scales with $\sqrt{K}$, where K is the number of users.

Corollary 1. *For the K-user IC-CM with delayed CSIT, the achievable SDoF in Equation (7) is lower bounded as*

$$\text{SDoF}^{\text{ach.}}_{\text{IC-CM}} > \frac{1}{2}(\sqrt{K}-6)^+, \tag{9}$$

where $(x)^+ \triangleq \max(x,0)$.

We present the proof of Theorem 1 and Corollary 1 in Section 4.

Theorem 2. *For the K-user IC-EE with delayed CSIT, the following secure sum degrees of freedom is achievable:*

$$\text{SDoF}^*_{\text{IC-EE}} \geq \text{SDoF}^{\text{ach.}}_{\text{IC-EE}} = \frac{R(K-R-1)}{R(R+1)+K}, \tag{10}$$

where,

$$R = \lfloor\sqrt{K}\rfloor - 1. \tag{11}$$

We next simplify the above expression and present a lower bound on the $\text{SDoF}^{\text{ach.}}_{\text{sum}}$.

Corollary 2. *For the K-user IC-EE with delayed CSIT, the achievable SDoF in Equation (7) is lower bounded as*

$$\text{SDoF}^{\text{ach.}}_{\text{IC-EE}} > \frac{1}{2}(\sqrt{K}-3)^{+}. \tag{12}$$

We present the proof of Theorem 2 and Corollary 2 in Section 5.

Theorem 3. *For the K-user IC-CM-EE with delayed CSIT, the following secure sum degrees of freedom is achievable:*

$$\text{SDoF}^{*}_{\text{IC-CM-EE}} \geq \text{SDoF}^{\text{ach.}}_{\text{IC-CM-EE}} = \frac{KR(K-R-2)}{(K-1)\times[R(R+1)+K]}, \tag{13}$$

where,

$$R = \left\lfloor \frac{-K + K\times\sqrt{1+\frac{(K-1)(K-2)}{K}}}{K-1} \right\rfloor. \tag{14}$$

In the next Corollary, we simplify the above expression and present a lower bound on the $\text{SDoF}^{\text{ach.}}_{\text{sum}}$.

Corollary 3. *For the K-user IC-CM-EE with delayed CSIT, the achievable SDoF in Equation (7) is lower bounded as*

$$\text{SDoF}^{\text{ach.}}_{\text{IC-CM-EE}} > \frac{1}{2}(\sqrt{K}-6)^{+}. \tag{15}$$

We present the proof of Theorem 3 and Corollary 3 in Section 6.

Remark 1. *We next compare the secure sum DoF of the previous Theorems to that of [8] (i.e., without secrecy constraints). For the K-user interference channel without secrecy constraints, the achievable sum DoF in [8] is given as:*

$$\text{DoF}^{\text{ach.}}_{\text{sum}} = \frac{K}{\lfloor\sqrt{K}\rfloor - 1 + \frac{K}{\lfloor\sqrt{K}\rfloor}}, \tag{16}$$

$$\geq \frac{\lfloor\sqrt{K}\rfloor}{2}, \tag{17}$$

$$\overset{(a)}{>} \frac{1}{2}(\sqrt{K}-1)^{+}, \tag{18}$$

where (a) follows from the fact that $\lfloor x\rfloor > x-1$. Comparing these results together, we can conclude that the scaling behavior of the sum SDoF is still attainable and there is negligible scaling loss in sum SDoF compared to no secrecy case for sufficiently large K (see Figure 2). Also, we present numerical evaluations for the sum SDoF in Table 2.

Table 2. Numerical comparison between the sum SDoF with the sum DoF in [8].

K	5	6	7	8	9	10	11	12
$\mathrm{SDoF}^{\text{ach.}}_{\text{IC-CM}} = \mathrm{SDoF}^{\text{ach.}}_{\text{IC-CM-EE}}$	0.3571	0.4500	0.5185	0.5714	0.6136	0.8333	0.9059	0.9697
$\mathrm{SDoF}^{\text{ach.}}_{\text{IC-EE}}$	0.4286	0.5000	0.5556	0.6000	0.8000	0.8750	0.9412	1.0000
$\mathrm{DoF}^{\text{ach.}}$ [8]	1.4286	1.5	1.556	1.6	1.8	1.875	1.9412	2

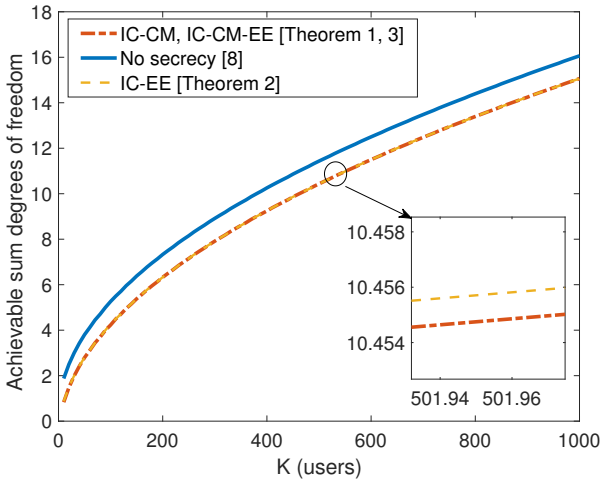


Figure 2. Comparison of achievable degrees of freedom (DoF) with delayed CSIT: with and without secrecy constraints.

4. Proof of Theorem 1

In this Section, we present the proof of Theorem 1. The transmission scheme consists of τ transmission blocks, where each block is of duration B. Across blocks, we employ stochastic wiretap coding (similar to the techniques employed in the literature on compound wiretap channels, see [19,23,24]). Within each block, the transmission is divided into two phases, which leverages delayed CSIT. In order to analyze the rate of the proposed scheme, we first take the limit of number of blocks $\tau \to \infty$, followed by the limit $B \to \infty$. For a given block, if we denote the (B-length) input of transmitter i as $\mathbf{x}_i$, and output of receiver i as $\mathbf{y}_i$, then the secure rate achievable by stochastic wiretap coding is given by:

$$R_i = \frac{I(\mathbf{s}_i; \mathbf{y}_i|\Omega) - \max_{j \neq i} I(\mathbf{s}_i; \mathbf{y}_j|\mathbf{s}^K_{-i}, \Omega)}{B}, \ i = 1, 2, \ldots, K, \tag{19}$$

where $\mathbf{s}_i$ is the vector of information symbols sent by transmitter i. Figure 3 gives an overview for these steps: stochastic encoding over blocks, and the two-phase scheme within each block that leverages delayed CSIT.

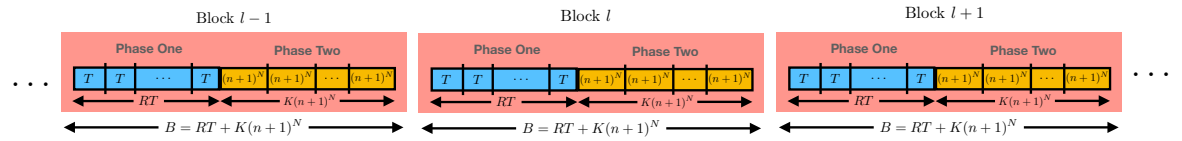


Figure 3. Stochastic encoding over transmission blocks for our proposed scheme.

4.1. Overview of the Achievability Scheme and SDoF Analysis

In this subsection, we present our secure transmission scheme. We consider a transmission block of length $RT + K(n+1)^N = R(Rn^N + (n+1)^N) + K(n+1)^N$, where R denotes the number of

transmission rounds and $N = RK(K-1)$, and n is an integer. The transmission scheme works over two phases. The goal of each transmitter is to securely send $T_1 = Rn^N + (n+1)^N - \lceil \frac{RT+(K-1)(n+1)^N}{K-1} \rceil$ information symbols to its corresponding receiver. In the first phase, each transmitter sends random linear combinations of the T_1 information symbols and the $T_2 = \lceil \frac{RT+(K-1)(n+1)^N}{K-1} \rceil$ artificial noise symbols in T time slots. Each transmitter repeats such transmission for R rounds, and hence, phase one spans RT time slots.

By the end of phase one, each receiver applies local interference alignment on its received signal to reduce the dimension of the aggregate interference. In the second phase, each transmitter knows the channel coefficients of phase one due to delayed CSIT. Subsequently, each transmitter sends a function of the net interference and artificial noises (generated in previous phase) which is simultaneously useful to all receivers. More specifically, each transmitter seperately sends $(n+1)^N$ linear equations of the past interference to all receivers. Therefore, phase 2 spans $K(n+1)^N$ time slots.

By the end of both phases, each receiver is able to decode its desired T_1 information symbols while satisfying the confidentiality constraints. The main aspect is that the parameters of the scheme (i.e., number of artificial noise symbols, number of repetition rounds, and durations of the phases) must be carefully selected to allow for reliable decoding of legitimate symbols, while satisfying the confidentiality constraints.

Therefore, the transmission scheme spans $RT + K(n+1)^N$ time slots, this scheme leads to the following achievable SDoF:

$$\text{SDoF}^{\text{ach.}}_{\text{IC-CM}} = \frac{KR(K-R-2)}{(K-1) \times [R(R+1)+K]}. \tag{20}$$

We calculate the achievable sum SDoF of this scheme in full detail in Section 4.3. Before we present the details of the scheme, we first optimize the achievable SDoF in Equation (20) with respect to the number of rounds R and also simplify the above expression, which leads to the expression in Corollary 1.

Lemma 1. *The optimal value of R^* which maximizes Equation (20) is given by*

$$R^* = \left\lfloor \frac{-K + K \times \sqrt{1 + \frac{(K-1)(K-2)}{K}}}{K-1} \right\rfloor. \tag{21}$$

Now, in order simplify the obtained expression in Equation (20), we state the following Corollary.

Corollary 4. *The optimal value of number of rounds R^* is lower bounded by*

$$R^* > \sqrt{K} - 5 = R_{lb}. \tag{22}$$

We present the proof of Lemma 1 and Corollary 4 in Appendix A.

Figure 4 depicts a comparison of between the two values of R (i.e., optimal R^* and lower bound R_{lb}). By substituting R_{lb} in Equation (20) leads to a lower bound on $\text{SDoF}^{\text{ach.}}_{\text{IC-CM}}$ as follows:

$$\text{SDoF}^{\text{ach.}}_{\text{IC-CM}} = \frac{KR(K-R-2)}{(K-1)\times[R(R+1)+K]}, \tag{23}$$

$$> \frac{R(K-R-2)}{R(R+1)+K} = \frac{(\sqrt{K}-5)(K-\sqrt{K}+3)}{(\sqrt{K}-5)(\sqrt{K}-4)+K}, \tag{24}$$

$$\overset{(a)}{=} \frac{K\sqrt{K}-6K+8\sqrt{K}-15}{2K-9\sqrt{K}+20}, \tag{25}$$

$$> \frac{K\sqrt{K}-6K+8\sqrt{K}-15}{2K}, \tag{26}$$

$$= \frac{\sqrt{K}-6}{2} + \frac{8\sqrt{K}-15}{2K}, \tag{27}$$

$$\overset{(b)}{>} \frac{1}{2}(\sqrt{K}-6)^+, \tag{28}$$

where in (a), the term $-9\sqrt{K}+20$ in the denominator is negative, $\forall K \geq 5$, so neglecting this term gives us Equation (26). In step (b), since the term $8\sqrt{K}-15$ is positive, $\forall K \geq 4$, hence omitting this term gives Equation (28). To this end, we get Equation (28) which shows the scaling of the achievable SDoF with K, the number of users.

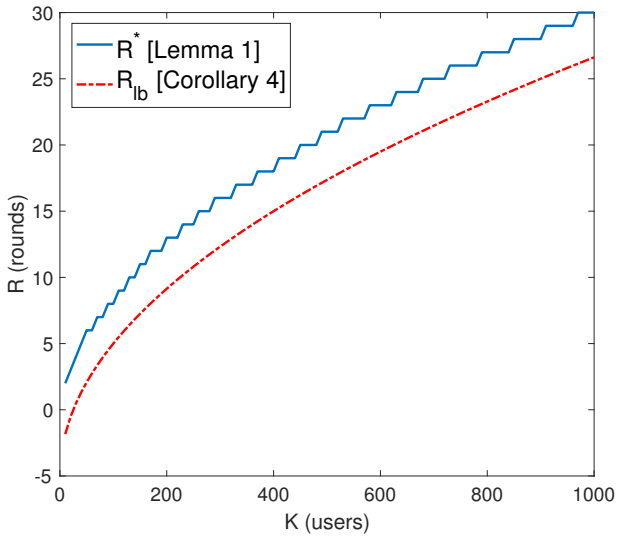


Figure 4. Comparison between the optimal value of R (number of rounds in phase one of the scheme) and its lower bound.

4.2. Detailed Description of the Achievability Scheme

Figure 5 depicts an overview of the two transmission phases. We now present the transmission scheme in full detail. For our scheme, we collectively denote the L symbols transmitted over L time slots as a super symbol and call this as the L symbol extension of the channel. With the extended channel, the signal vector at the kth receiver can be expressed as

$$\mathbf{y}_k = \sum_{j=1}^{K} \mathbf{H}_{kj}\mathbf{x}_j + \mathbf{n}_k, \tag{29}$$

where $\mathbf{x}_j$ is a $L \times 1$ column vector representing the L symbols transmitted by transmitter k in L time slots. $\mathbf{H}_{kj}$ is a $L \times L$ diagonal matrix representing the L symbol extension of the channel as follows:

$$\mathbf{H}_{kj} = \text{diag}\left(h_{kj}(1), h_{kj}(2), \ldots, h_{kj}(L)\right), \tag{30}$$

where $h_{kj}(t)$ is the channel coefficient between transmitter j and receiver k at time slot t. Now we proceed to the proposed scheme which works over two phases.

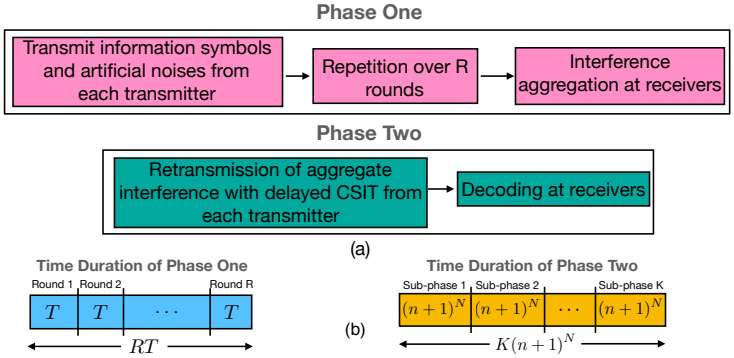


Figure 5. (**a**) Block diagram for the transmission scheme and (**b**) time duration of the phases.

4.2.1. Phase 1: Interference Creation with Information Symbols and Artificial Noises

Recall that the goal of each transmitter is to send T_1 information symbols securely to its respective receiver. This phase is comprised of R rounds, where, in each round, every transmitter j sends linear combinations of the T_1 information symbols $\mathbf{s}_j \in \mathbb{C}^{T_1 \times 1}$, mixed with T_2 artificial noises $\mathbf{u}_j \in \mathbb{C}^{T_2 \times 1}$, where the elements of $\mathbf{u}_j$ are drawn from complex-Gaussian distribution with average power P. Hence, the signal sent by transmitter j in each round r can be written as

$$\mathbf{x}_j = \mathbf{V}_j \begin{bmatrix} \mathbf{s}_j \\ \mathbf{u}_j \end{bmatrix}, \ \forall j = 1, 2, \ldots, K, \tag{31}$$

where $\mathbf{V}_j, \forall j = 1, 2, \ldots, K$ is a random *mixing* matrix of dimension $T \times T$ whose elements are drawn from complex-Gaussian distribution with zero mean and unit variance at transmitter j. $\mathbf{V}_j, \forall j = 1, 2, \ldots, K$ is known at all terminals (all transmitters and receivers). The received signal at receiver k for round $r \in \{1, 2, \ldots, R\}$ is given by

$$\mathbf{y}_k^r = \sum_{j=1}^{K} \mathbf{H}_{kj}^r \mathbf{x}_j + \mathbf{n}_k^r, \tag{32}$$

where $\mathbf{x}_j$ is the $T \times 1$ column vector representing the T symbol extension of the transmitted symbols from transmitter j, and $\mathbf{n}_k^r$ represents the receiver noise in round r at receiver k. This phase spans RT time slots where $R \in \mathbb{N}$ is the number of transmission rounds and $T = Rn^N + (n+1)^N$ time slots where $n \in \mathbb{N}$ and $N = RK(K-1)$.

◇ Interference Aggregation at Receivers

At the end of phase 1, each receiver k has the signals $\mathbf{y}_k = \{\mathbf{y}_k^r\}_{r=1}^R$, over R rounds. Each receiver performs a linear post-processing of its received signals in order to align the aggregate interference (generated from symbols and artificial noises) from the $(K-1)$ unintended transmitters. In particular,

each receiver multiplies its received signals in the rth block with a matrix $\mathbf{W}$ (of dimension $T \times n^N$) as follows:

$$\tilde{\mathbf{y}}_k^r = \mathbf{W}^H \mathbf{y}_k^r = \mathbf{W}^H \left(\sum_{j=1}^{K} \mathbf{H}_{kj}^r \mathbf{x}_j + \mathbf{n}_k^r \right), \tag{33}$$

$$= \mathbf{W}^H \mathbf{H}_{kk}^r \mathbf{x}_k + \sum_{j \neq k} \mathbf{W}^H \mathbf{H}_{kj}^r \mathbf{x}_j + \mathbf{W}^H \mathbf{n}_k^r, \tag{34}$$

$$= \mathbf{W}^H \mathbf{H}_{kk}^r \mathbf{x}_k + \sum_{j \neq k} \mathbf{W}^H \mathbf{H}_{kj}^r \mathbf{x}_j + \tilde{\mathbf{n}}_k^r. \tag{35}$$

The goal is to design the matrices $\mathbf{W}$ and $\mathbf{X}$ such that

$$\mathbf{W}^H \mathbf{H}_{kj}^r \prec \mathbf{X}, \forall k = 1, 2, \ldots, K, k \neq j, \forall r = 1, 2, \ldots, R, \tag{36}$$

where $\mathbf{X} \in \mathbb{C}^{(n+1)^N \times T}$. Here the notation $\mathbf{A} \prec \mathbf{B}$ means that the set of row vectors of matrix $\mathbf{A}$ is a subset of the row vectors of matrix $\mathbf{B}$. To this end, we choose $\mathbf{W}$ and $\mathbf{X}$ as follows:

$$\mathbf{W} = \left[\prod_{(r,m,i) \in \mathcal{S}} (\mathbf{H}_{mi}^{r^{(n_{mi}^r)}})^H \mathbb{1} : 0 \leq n_{mi}^r \leq n-1 \right], \tag{37}$$

$$\mathbf{X} = \left[\prod_{(r,m,i) \in \mathcal{S}} (\mathbf{H}_{mi}^{r^{(n_{mi}^r)}})^H \mathbb{1} : 0 \leq n_{mi}^r \leq n \right]^H, \tag{38}$$

where $\mathbb{1}$ is the all ones column vector and the set $\mathcal{S} = \{(r, m, i) : \forall r \in \{1, \ldots, R\}, \forall m \neq i \in \{1, \ldots, K\}\}$. Note that the set $\mathcal{S}$ does not contain the channel matrix $\mathbf{H}_{kk}^r$ that carries the information symbols intended to receiver k. However, multiplying with any channel gain that appears in $\mathbf{W}$ results in aligning this signal within $\mathbf{X}$ asymptotically. It is worth noting that, $\mathbf{X}$ defines all the possible interference generated by the transmitters at all receivers. Hence, this choice of $\mathbf{X}$ and $\mathbf{W}$ guarantees that the alignment condition Equation (36) is satisfied. Therefore, the received signal after post-processing in round r at receiver k can be written as

$$\tilde{\mathbf{y}}_k^r = \mathbf{W}^H \mathbf{H}_{kk}^r \mathbf{x}_k + \sum_{j \neq k} \mathbf{W}^H \mathbf{H}_{kj}^r \mathbf{x}_j + \mathbf{W}^H \mathbf{n}_k^r, \tag{39}$$

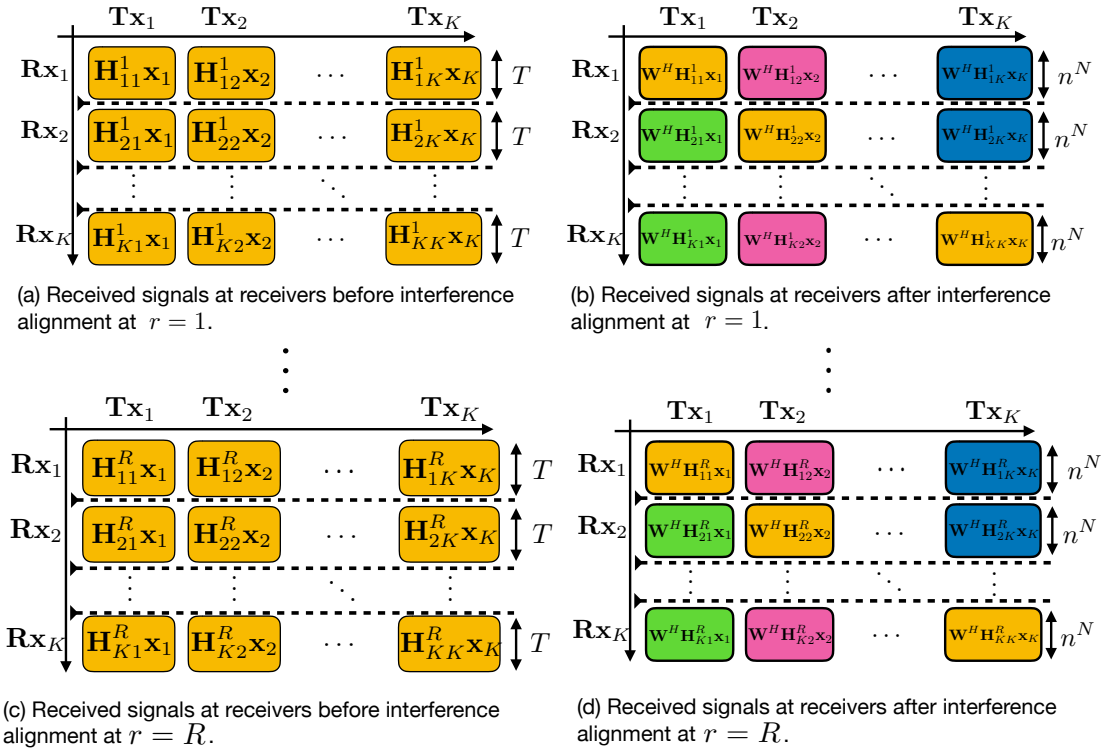
$$= \mathbf{W}^H \mathbf{H}_{kk}^r \mathbf{x}_k + \sum_{j \neq k} \Pi_{kj}^r \mathbf{X} \mathbf{x}_j + \mathbf{W}^H \mathbf{n}_k^r, \tag{40}$$

where $\Pi_{kj}^r \in \mathbb{C}^{n^N \times (n+1)^N}$ is a selection and permutation matrix. Now after the end of phase 1, receiver k has Rn^N equations of T desired symbols (which are composed of T_1 information symbols and T_2 artificial noises generated by the transmitter k) plus $(K-1)$ interference terms, which are of dimension $(n+1)^N$. Figure 6 gives a detailed structure for the first phase of the transmission scheme.

4.2.2. Phase 2: Re-Transmission of Aggregate Interference with Delayed CSIT

For the second phase, each transmitter k uses $(n+1)^N$ time slots to re-transmit the aggregated interference $(\mathbf{X}\mathbf{x}_k)$ generated in the first phase at the receivers, which is sufficient to cancel out the interference term at receiver $j \neq k$, and to provide additional $(n+1)^N$ equations of the desired symbols to receiver k. Then, this phase spans $K(n+1)^N$ time slots. The transmitted signal from transmitter k is as follows:

$$\mathbf{z}_k = \mathbf{X}\ \mathbf{x}_k, \forall k = 1, 2, \ldots, K. \tag{41}$$



(a) Received signals at receivers before interference alignment at $r = 1$.

(b) Received signals at receivers after interference alignment at $r = 1$.

(c) Received signals at receivers before interference alignment at $r = R$.

(d) Received signals at receivers after interference alignment at $r = R$.

Figure 6. Graphical representation for the first phase of the proposed scheme.

◇ Decoding at Receivers

At the end of phase 2, the interference at receiver k is removed by subtracting the terms $\sum_{j=1,j\neq k} \Pi_{kj}^{r} \mathbf{X}\ \mathbf{x}_j$ from the equalized signal $\tilde{\mathbf{y}}_k^r$, i.e., (ignoring the additive noise $\mathbf{n}_k^r$)

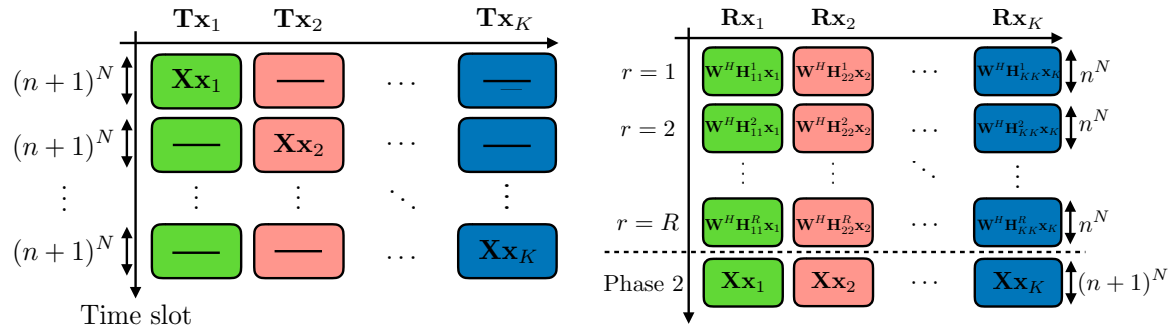
$$\mathbf{W}^H \mathbf{H}_{kk}^r \mathbf{x}_k = \tilde{\mathbf{y}}_k^r - \sum_{j=1,j\neq k} \Pi_{kj}^{r} \mathbf{X}\ \mathbf{x}_j. \tag{42}$$

Canceling the interference terms leaves each receiver $k, \forall k \in \{1,\ldots,K\}$ with Rn^N useful linear equations besides $(n+1)^N$ useful equations from transmitter k (from phase 2). At the end of phase 2, receiver k will collectively get the following signal,

$$\underbrace{\left[\mathbf{X}^H, (\mathbf{W}^H \mathbf{H}_{kk}^1)^H, \ldots, (\mathbf{W}^H \mathbf{H}_{kk}^R)^H\right]}_{\mathbf{B}_k}{}^H \mathbf{V}_k \begin{bmatrix} \mathbf{s}_k \\ \mathbf{u}_k \end{bmatrix}. \tag{43}$$

Therefore, at the end of phase 2, each receiver has enough linear equations of the desired symbols. In order to ensure decodability, we need to prove that $\mathbf{B}_k \mathbf{V}_k$ is full rank and hence each receiver will be able to decode its desired T_1 information symbols. First, we notice that $\mathbf{V}_k$ is full rank matrix and hence $\text{rank}(\mathbf{B}_k \mathbf{V}_k) = \text{rank}(\mathbf{B}_k)$ [25]. In Appendix B, we show that $\mathbf{B}_k$ is full rank. Figure 7 gives a detailed structure for the second phase of the transmission scheme.

Before we start the achievable secure rate analysis, we want to highlight first on the dimensions of the information symbols $\mathbf{s}_i \in \mathbb{C}^{T_1 \times 1}$ and the artificial noises $\mathbf{u}_i \in \mathbb{C}^{T_2 \times 1}, \forall i = 1, 2, \ldots, K$.



(a) Aggregate interference retransmission.

(b) Total received interference-free signals at receivers after cancelling the interference terms.

Figure 7. Graphical representation for the second phase of the proposed scheme.

4.2.3. Choice of T_1 and T_2 to Satisfy the Confidentiality Constraints

Without loss of generality, let us consider receiver 1. After decoding $\mathbf{s}_1$ and $\mathbf{u}_1$, receiver 1 will have RT equations of $\{\mathbf{s}_i\}_{i=2}^K$, $\{\mathbf{u}_i\}_{i=2}^K$ from phase one, and $(K-1)(n+1)^N$ equations of $\{\mathbf{s}_i\}_{i=2}^K$, $\{\mathbf{u}_i\}_{i=2}^K$ from phase two. Then, the total number of equations seen at receiver 1 is $RT+(K-1)(n+1)^N$. Hence, in order to keep the unintended information symbols of $(K-1)$ transmitters at this receiver secure, we require that the number of these equations must be at most equal to the total number of the artificial noise dimensions of the $(K-1)$ transmitters, i.e.,

$$RT+(K-1)(n+1)^N \leq (K-1)T_2. \tag{44}$$

Therefore, we choose T_2 as

$$T_2 = \left\lceil \frac{RT+(K-1)(n+1)^N}{K-1} \right\rceil. \tag{45}$$

Note that since $T = T_1 + T_2$, so we can get T_1 as follows:

$$T_1 = Rn^N + (n+1)^N - \left\lceil \frac{RT+(K-1)(n+1)^N}{K-1} \right\rceil. \tag{46}$$

We next compute the achievable secrecy rates and SDoF for the K-user interference channel with confidential messages and delayed CSIT.

4.3. Secrecy Rate and SDoF Calculation

Using stochastic encoding described in Appendix XII of [26], for a block length $B = RT + K(n+1)^N$, the following secure rate is achievable:

$$R_i = \frac{I(\mathbf{s}_i;\mathbf{y}_i|\Omega) - \max_{j\neq i} I(\mathbf{s}_i;\mathbf{y}_j|\mathbf{s}_{-i}^K,\Omega)}{RT+K(n+1)^N}, \ i=1,2,\ldots,K, \tag{47}$$

where $I(\mathbf{s}_i;\mathbf{y}_i|\Omega)$ is the mutual information between the information symbols vector $\mathbf{s}_i$ and $\mathbf{y}_i$, the received composite signal vector at the intended receiver i, given the knowledge of the channel coefficients. $I(\mathbf{s}_i;\mathbf{y}_j|\mathbf{s}_{-i}^K,\Omega)$ is the mutual information between $\mathbf{s}_i$ and $\mathbf{y}_j$, the received composite

signal vector at the unintended receiver j, i.e., the strongest adversary with respect to transmitter i, conditioned on the information symbols $\mathbf{s}_{-i}^{K}$. In terms of differential entropy, we can write,

$$\textbf{Term A:} \quad I(\mathbf{s}_i;\mathbf{y}_i|\Omega) = h(\mathbf{y}_i|\Omega) - h(\mathbf{y}_i|\mathbf{s}_i,\Omega),\ i=1,2,\ldots,K, \tag{48}$$

$$\textbf{Term B:} \quad I(\mathbf{s}_i;\mathbf{y}_j|\mathbf{s}_{-i}^{K},\Omega) = h(\mathbf{y}_j|\mathbf{s}_{-i}^{K},\Omega) - h(\mathbf{y}_j|\mathbf{s}_{1}^{K},\Omega),\ j=1,2,\ldots,K, j\neq i. \tag{49}$$

We collectively write the received signal $\mathbf{y}_i$ at receiver i over $RT + K(n+1)^N$ time slots as follows:

$$\mathbf{y}_i = \mathbf{A}_i\mathbf{V}\mathbf{q} + \mathbf{n}_i, \forall i = 1,2,\ldots,K, \tag{50}$$

where,

$$\mathbf{A}_i = \begin{bmatrix}\mathbf{C}_i\\ \hline \mathbf{D}_i\end{bmatrix},\ \mathbf{C}_i = \begin{bmatrix}\mathbf{H}_{i1}^{1} & \mathbf{H}_{i2}^{1} & \cdots & \mathbf{H}_{iK}^{1}\\ \mathbf{H}_{i1}^{2} & \mathbf{H}_{i2}^{2} & \cdots & \mathbf{H}_{iK}^{2}\\ \vdots & \vdots & \cdots & \vdots\\ \mathbf{H}_{i1}^{R} & \mathbf{H}_{i2}^{R} & \cdots & \mathbf{H}_{iK}^{R}\end{bmatrix},$$

$$\mathbf{D}_i = \text{blkdiag}(\tilde{\mathbf{H}}_{i1}\mathbf{X},\ldots,\tilde{\mathbf{H}}_{iK}\mathbf{X}), \tag{51}$$

where $\mathbf{A}_i$ has dimensions of $(RT + K(n+1)^N) \times KT$. Note that $\mathbf{A}_i$ is partitioned into two sub matrices $\mathbf{C}_i$ and $\mathbf{D}_i$. $\mathbf{C}_i$ consists of block matrices, where each block matrix has dimensions of $T \times T$ whose elements are i.i.d. drawn from a continuous distribution and hence, it is full rank, almost surely (i.e., $\text{rank}(\mathbf{C}_i) = RT$). $\mathbf{D}_i$ has a block diagonal structure (each block matrix has dimensions of $(n+1)^N \times T$) since the transmission in phase two of the scheme is done in TDMA fashion. Note that each block is a full rank matrix (i.e., $\text{rank}(\tilde{\mathbf{H}}_{ij}\mathbf{X}) = \text{rank}(\mathbf{X}) = (n+1)^N, \forall j = 1,\ldots,K$). The matrix $\mathbf{X}$ is a full rank matrix as proved in [26]. The matrix $\mathbf{V}$ can be written as follows:

$$\mathbf{V} = \text{blkdiag}(\mathbf{V}_1,\mathbf{V}_2,\ldots,\mathbf{V}_K), \tag{52}$$

where $\mathbf{V}$ is the block diagonal matrix with dimensions of $KT \times KT$. Furthermore, we write

$$\mathbf{q} = \begin{bmatrix}\mathbf{s}_1^T & \mathbf{u}_1^T & \mathbf{s}_2^T & \mathbf{u}_2^T & \cdots & \mathbf{s}_K^T & \mathbf{u}_K^T\end{bmatrix}^T \tag{53}$$

as a column vector of length KT, which contains the information symbols and the artificial noises of transmitters $1,\ldots,K$.

Before we proceed, we present two Lemmas; we provide the proof of Lemma 2 in Appendix C. For Lemma 3, we follow similar steps as in [17], the proof of this Lemma is provided in [26].

Lemma 2. *Let* $\mathbf{A}$ *be a matrix with dimension* $M \times N$ *and* $\mathbf{X} = (x_1,\ldots,x_N)^T$ *be a zero-mean jointly complex Gaussian random vector with covariance matrix* $P\mathbf{I}$*. Also, let* $\mathbf{N} = (n_1,\ldots,n_M)^T$ *be a zero-mean jointly complex Gaussian random vector with covariance matrix* $\sigma^2\mathbf{I}$*, independent of* $\mathbf{X}$*, then*

$$h(\mathbf{A}\mathbf{X}+\mathbf{N}) = \log(\pi e)^M + \sum_{i=1}^{\text{rank}(\mathbf{A})}\log(\lambda_i P + \sigma^2), \tag{54}$$

where $\{\lambda_i\}_{i=1}^{rank(\mathbf{A})}$ *are the singular values of* $\mathbf{A}$*.*

Lemma 3. *Consider two matrices* $\mathbf{A}_{M\times N}$ *and* $\mathbf{B}_{N\times M}$ *where* $M \leq N$*. The elements of matrix* $\mathbf{B}$ *are chosen independently from the entries of* $\mathbf{A}$ *at random from a continuous distribution. Then,*

$$\text{rank}(\mathbf{AB}) = \text{rank}(\mathbf{A}), \ \textit{almost surely}. \tag{55}$$

Without loss of generality, let us consider the first transmitter. The received signal at the first receiver after removing the $(K-1)$ interference terms is written as follows:

$$\tilde{\mathbf{y}}_1 = \underbrace{\begin{bmatrix} \mathbf{W}^H & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{W}^H & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{W}^H & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{I} \end{bmatrix}}_{\mathbf{\Psi}} \left(\underbrace{\begin{bmatrix} \mathbf{H}_{11}^1 \\ \mathbf{H}_{11}^2 \\ \vdots \\ \mathbf{H}_{11}^R \\ \tilde{\mathbf{H}}_{11}\mathbf{X} \end{bmatrix}}_{\mathbf{F}_1} \mathbf{V}_1 \underbrace{\begin{bmatrix} \mathbf{s}_1 \\ \mathbf{u}_1 \end{bmatrix}}_{\mathbf{x}_1} + \underbrace{\begin{bmatrix} \mathbf{n}_1^1 \\ \mathbf{n}_1^2 \\ \vdots \\ \mathbf{n}_1^R \\ \tilde{\mathbf{n}}_1 \end{bmatrix}}_{\mathbf{n}_1} \right). \tag{56}$$

◇ Lower Bounding **Term A**

We note that $\mathbf{s}_1 \to \mathbf{x}_1 \to \mathbf{y}_1 \to \tilde{\mathbf{y}}_1$ forms a Markov chain, thus

$$I(\mathbf{x}_1; \mathbf{y}_1|\Omega) \geq I(\mathbf{s}_1; \mathbf{y}_1|\Omega) \geq I(\mathbf{s}_1; \tilde{\mathbf{y}}_1|\Omega), \tag{57}$$

$$= h(\tilde{\mathbf{y}}_1|\Omega) - h(\tilde{\mathbf{y}}_1|\mathbf{s}_1, \Omega). \tag{58}$$

Using Equation (56), we can write $h(\tilde{\mathbf{y}}_1|\Omega)$ as follows:

$$h(\tilde{\mathbf{y}}_1|\Omega) = h(\mathbf{\Psi}(\mathbf{F}_1\mathbf{x}_1 + \mathbf{n}_1)), \tag{59}$$

$$= h(\mathbf{F}_1\mathbf{x}_1 + \mathbf{n}_1) + \log(\det(\Psi)), \tag{60}$$

$$= \log(\pi e)^{RT+(n+1)^N} + \sum_{i=1}^{r(\mathbf{F}_1)} \log(\lambda_i P + \sigma^2) + \log(\det(\Psi)), \tag{61}$$

$$\overset{(a)}{=} \log(\pi e)^{RT+(n+1)^N} + \sum_{i=1}^{T} \log(\lambda_i P + \sigma^2) + \log(\det(\Psi)), \tag{62}$$

where $\{\lambda_i\}_{i=1}^{r(\mathbf{F}_1)}$ are the singular values of $\mathbf{F}_1$. In (a), we note that $\mathbf{B}_1 = \mathbf{\Psi}\mathbf{F}_1$ is full rank. Using full rank decomposition Theorem [25], we conclude that $\mathbf{F}_1$ is also full rank, i.e., $\text{rank}(\mathbf{F}_1) = T$.

Now, we write $h(\tilde{\mathbf{y}}_1|\mathbf{s}_1, \Omega)$ as follows:

$$h(\tilde{\mathbf{y}}_1|\mathbf{s}_1, \Omega) = h(\mathbf{\Psi}(\tilde{\mathbf{F}}_1\mathbf{u}_1 + \mathbf{n}_1)), \tag{63}$$

$$= h(\tilde{\mathbf{F}}_1\mathbf{u}_1 + \mathbf{n}_1) + \log(\det(\Psi)), \tag{64}$$

$$= \log(\pi e)^{RT+(n+1)^N} + \sum_{i=1}^{r(\tilde{\mathbf{F}}_1)} \log(\lambda_i^{'} P + \sigma^2) + \log(\det(\Psi)), \tag{65}$$

$$\leq \log(\pi e)^{RT+(n+1)^N} + \sum_{i=1}^{T_2} \log(\lambda_i^{'} P + \sigma^2) + \log(\det(\Psi)), \tag{66}$$

where,

$$\tilde{\mathbf{F}}_1 = \begin{bmatrix} \mathbf{H}_{11}^1 \\ \mathbf{H}_{11}^2 \\ \vdots \\ \mathbf{H}_{11}^R \\ \tilde{\mathbf{H}}_{11}\mathbf{X} \end{bmatrix} \mathbf{V}_{1,\mathbf{u}_1}, \tag{67}$$

where $\tilde{\mathbf{F}}_1$ has dimensions of $RT + (n+1)^N \times T_2$, and $\{\lambda_i^{'}\}_{i=1}^{r(\tilde{\mathbf{F}}_1)}$ are the singular values of $\tilde{\mathbf{F}}_1$. $\mathbf{V}_{1,\mathbf{u}_1}$ has dimensions of $T \times T_2$. Note that, we can view the *mixing* matrix $\mathbf{V}_i$ being composed of two parts i.e., $\mathbf{V}_i = (\mathbf{V}_{i,\mathbf{s}_i} : \mathbf{V}_{i,\mathbf{u}_i}), \forall i \in 1, 2, \ldots, K$ where $\mathbf{V}_{i,\mathbf{s}_i}$ corresponding to the information symbol $\mathbf{s}_i$ and $\mathbf{V}_{i,\mathbf{u}_i}$ corresponding to the artificial noise $\mathbf{u}_i$.

From the substitution of Equations (62) and (66) into Equation (58), we obtain

$$I(\mathbf{x}_1; \mathbf{y}_1|\Omega) \geq I(\mathbf{s}_1; \mathbf{y}_1|\Omega) \geq I(\mathbf{s}_1; \tilde{\mathbf{y}}_1|\Omega), \tag{68}$$

$$\geq \sum_{i=1}^{T} \log(\lambda_i P + \sigma^2) - \sum_{i=1}^{T_2} \log(\lambda_i^{'} P + \sigma^2). \tag{69}$$

Before calculating the second term, i.e., **Term B**. We collectively write the received signal $\mathbf{y}_j$ at receiver j over $RT + K(n+1)^N$ time slots as follows:

$$\mathbf{y}_j = \mathbf{A}_j \mathbf{V} \mathbf{q} + \mathbf{n}_j, \tag{70}$$

where $\mathbf{A}_j$ is written as follows:

$$\mathbf{A}_j = \begin{bmatrix} \mathbf{C}_j \\ \hline \mathbf{D}_j \end{bmatrix}, \ \mathbf{C}_j = \begin{bmatrix} \mathbf{H}_{j1}^1 & \mathbf{H}_{j2}^1 & \cdots & \mathbf{H}_{jK}^1 \\ \mathbf{H}_{j1}^2 & \mathbf{H}_{j2}^2 & \cdots & \mathbf{H}_{jK}^2 \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{H}_{j1}^R & \mathbf{H}_{j2}^R & \cdots & \mathbf{H}_{jK}^R \end{bmatrix},$$

$$\mathbf{D}_j = \text{blkdiag}(\tilde{\mathbf{H}}_{j1}\mathbf{X}, \ldots, \tilde{\mathbf{H}}_{jK}\mathbf{X}). \tag{71}$$

where $\mathbf{A}_j$ has dimensions of $(RT + K(n+1)^N) \times KT$. The matrix V is written as follows:

$$\mathbf{V} = \text{blkdiag}(\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_K), \tag{72}$$

where $\mathbf{V}$ is the block diagonal matrix with dimensions of $KT \times KT$. Furthermore, we write

$$\mathbf{q} = \begin{bmatrix} \mathbf{s}_1^T & \mathbf{u}_1^T & \mathbf{s}_2^T & \mathbf{u}_2^T & \cdots & \mathbf{s}_K^T & \mathbf{u}_K^T \end{bmatrix}^T \tag{73}$$

as a column vector of length KT, which contains the information symbols and the artificial noises of transmitters $1, \ldots, K$.

◇ Upper Bounding **Term B**

Now, we can compute **Term B**, i.e., $I(\mathbf{s}_1; \mathbf{y}_j|\mathbf{s}_{-1}^K, \Omega)$ as follows:

$$I(\mathbf{s}_1; \mathbf{y}_j|\mathbf{s}_{-1}^K, \Omega) = h(\mathbf{y}_j|\mathbf{s}_{-1}^K, \Omega) - h(\mathbf{y}_j|\mathbf{s}_1^K, \Omega), \tag{74}$$

$$= h(\mathbf{A}_j \bar{\mathbf{V}} \bar{\mathbf{q}} + \mathbf{n}_j) - h(\mathbf{A}_j \tilde{\mathbf{V}} \tilde{\mathbf{q}} + \mathbf{n}_j), \tag{75}$$

where $\bar{\mathbf{V}}$ is a truncated version of $\mathbf{V}$ with dimensions $KT \times (T_1 + KT_2)$. Furthermore, we write

$$\bar{\mathbf{q}} = \begin{bmatrix} \mathbf{s}_1^T & \mathbf{u}_1^T & \cdots & \mathbf{u}_K^T \end{bmatrix}^T \tag{76}$$

as a column vector of length $T_1 + KT_2$, which contains T_1 information symbols and KT_2 artificial noises of transmitters. Also, $\tilde{\mathbf{V}}$ is a truncated version of $\mathbf{V}$ with dimensions $KT \times KT_2$. Furthermore, we write

$$\tilde{\mathbf{q}} = \begin{bmatrix} \mathbf{u}_1^T & \mathbf{u}_2^T & \cdots & \mathbf{u}_K^T \end{bmatrix}^T \tag{77}$$

as a column vector of length KT_2, which contains only KT_2 artificial noises of transmitters.

Using Equation (70), we can write $h(\mathbf{y}_j|\Omega)$ as follows:

$$h(\mathbf{y}_j|\mathbf{s}_{-1}^K,\Omega) \overset{(a)}{=} h(\mathbf{A}_j\bar{\mathbf{V}}\bar{\mathbf{q}}+\mathbf{n}_j) \tag{78}$$

$$= \log(\pi e)^{RT+K(n+1)^N} + \sum_{i=1}^{r(\mathbf{A}_j\bar{\mathbf{V}})} \log(\Lambda_i P+\sigma^2), \tag{79}$$

$$\leq \log(\pi e)^{RT+K(n+1)^N} + \sum_{i=1}^{r(\mathbf{A}_j\mathbf{V})} \log(\Lambda_i P+\sigma^2), \tag{80}$$

$$\overset{(b)}{=} \log(\pi e)^{RT+K(n+1)^N} + \sum_{i=1}^{r(\mathbf{A}_j)} \log(\Lambda_i P+\sigma^2). \tag{81}$$

In (a), we used Lemma 2. $\{\Lambda_i\}_{i=1}^{r(\mathbf{A}_j\mathbf{V})}$ are the singular values of $\mathbf{A}_j\mathbf{V}$. Note that in (b), $\mathbf{V}$ is an invertible matrix with rank KT, therefore $\text{rank}(\mathbf{A}_j\mathbf{V}) = \text{rank}(\mathbf{A}_j)$.

Now, we write $h(\mathbf{y}_j|\mathbf{s}_1^K,\Omega)$ as follows:

$$h(\mathbf{y}_j|\mathbf{s}_1^K,\Omega) = \log(\pi e)^{RT+K(n+1)^N} + \sum_{i=1}^{r(\mathbf{A}_j\tilde{\mathbf{V}})} \log(\Lambda_i' P+\sigma^2), \tag{82}$$

$$\overset{(a)}{\geq} \log(\pi e)^{RT+K(n+1)^N} + \sum_{i=1}^{r(\mathbf{A}_j\tilde{\tilde{\mathbf{V}}})} \log(\Lambda_i' P+\sigma^2), \tag{83}$$

$$\overset{(b)}{=} \log(\pi e)^{RT+K(n+1)^N} + \sum_{i=1}^{r(\mathbf{A}_j)} \log(\Lambda_i' P+\sigma^2), \tag{84}$$

where $\{\Lambda_i'\}_{i=1}^{r(\mathbf{A}_j\tilde{\mathbf{V}})}$ are the singular values of $\mathbf{A}_j\tilde{\mathbf{V}}$. In (a), $\tilde{\tilde{\mathbf{V}}}$ is a truncated version of $\tilde{\mathbf{V}}$ with dimensions of $KT \times (RT+K(n+1)^N)$, therefore, $r(\mathbf{A}_j\tilde{\mathbf{V}}) \geq r(\mathbf{A}_j\tilde{\tilde{\mathbf{V}}})$. In (b), we used Lemma 3, i.e., $\text{rank}(\mathbf{A}_j\tilde{\tilde{\mathbf{V}}}) = \text{rank}(\mathbf{A}_j)$. The multiplication of $\mathbf{A}_j$ and $\tilde{\tilde{\mathbf{V}}}$ can be viewed as $RT+K(n+1)^N$ linear combinations of the KT rows of matrix $\tilde{\tilde{\mathbf{V}}}$, whose elements are generated independently of $\mathbf{A}_j$ from a continuous distribution. In Appendix XI of [26], we show that multiplying $\mathbf{A}_j$ with a non-square random matrix $\tilde{\tilde{\mathbf{V}}}$ does not reduce the rank of matrix $\mathbf{A}_j$, almost surely. Hence, from the above argument, in order to ensure that $\text{rank}(\mathbf{A}_j\tilde{\tilde{\mathbf{V}}}) = \text{rank}(\mathbf{A}_j)$, we must pick $RT+(K-1)(n+1)^N \leq (K-1)T_2$, which gives the reasoning behind the choice of the parameter T_2.

From the substitution of Equations (81) and (84) into Equation (75), we obtain

$$I(\mathbf{s}_1;\mathbf{y}_j|\Omega) \leq \sum_{i=1}^{r(\mathbf{A}_j)} \log(\Lambda_i P+\sigma^2) - \sum_{i=1}^{r(\mathbf{A}_j)} \log(\Lambda_i' P+\sigma^2). \tag{85}$$

Combining Equations (69) and (85), we have

$$R_1 \geq \frac{\sum_{i=1}^{T}\log(\lambda_i P+\sigma^2) - \sum_{i=1}^{T_2}\log(\lambda_i' P+\sigma^2) - \left(\sum_{i=1}^{r(\mathbf{A}_j)}\log(\Lambda_i P+\sigma^2) - \sum_{i=1}^{r(\mathbf{A}_j)}\log(\Lambda_i' P+\sigma^2)\right)}{RT+K(n+1)^N}, \tag{86}$$

$$\geq \underbrace{\frac{T\log(\lambda_{\min}P+\sigma^2) - T_2\log(\lambda_{\max}' P+\sigma^2)}{RT+K(n+1)^N}}_{\text{Term 1}} - \underbrace{\frac{\left(r(\mathbf{A}_j)\log(\Lambda_{\max}P+\sigma^2) - r(\mathbf{A}_j)\log(\Lambda_{\min}' P+\sigma^2)\right)}{RT+K(n+1)^N}}_{\text{Term 2}}, \tag{87}$$

where $\lambda_{\min} = \min_i\{\lambda_i\}_{i=1}^{r(\mathbf{F}_1)}$, $\lambda_{\max}' = \max_i\{\lambda_i'\}_{i=1}^{r(\tilde{\mathbf{F}}_1)}$, $\Lambda_{\min}' = \min_i\{\Lambda_i'\}_{i=1}^{r(\mathbf{A_j}\tilde{\mathbf{V}})}$ and $\Lambda_{\max} = \max_i\{\Lambda_i\}_{i=1}^{r(\mathbf{A_j}\mathbf{V})}$.

We now simplify the two terms as follows:

$$\text{Term 1} = \frac{T\log(\lambda_{\min}P+\sigma^2) - T_2\log(\lambda'_{\max}P+\sigma^2)}{RT+K(n+1)^N}, \tag{88}$$

$$= \frac{\left(Rn^N+(n+1)^N\right)\log(\lambda_{\min}P+\sigma^2) - \left\lceil\frac{R(Rn^N+(n+1)^N)+(K-1)(n+1)^N}{K-1}\right\rceil\log(\lambda'_{\max}P+\sigma^2)}{R(Rn^N+(n+1)^N)+K(n+1)^N}, \tag{89}$$

$$\overset{(a)}{\geq} \frac{\left(Rn^N+(n+1)^N\right)\log(\lambda_{\min}P+\sigma^2) - \left(\frac{R(Rn^N+(n+1)^N)+(K-1)(n+1)^N}{K-1}+1\right)\log(\lambda'_{\max}P+\sigma^2)}{R(Rn^N+(n+1)^N)+K(n+1)^N}, \tag{90}$$

where in (a), we used the property that $\lceil x\rceil < x+1$. Also,

$$\text{Term 2} = \frac{r(\mathbf{A}_j)\left(\log(\Lambda_{\max}P+\sigma^2) - \log(\Lambda'_{\min}P+\sigma^2)\right)}{RT+K(n+1)^N}, \tag{91}$$

$$= \frac{r(\mathbf{A}_j)\left(\log(\Lambda_{\max}P+\sigma^2) - \log(\Lambda'_{\min}P+\sigma^2)\right)}{R(Rn^N+(n+1)^N)+K(n+1)^N}, \tag{92}$$

$$\leq \frac{\left(R(Rn^N+(n+1)^N)+K(n+1)^N\right)\left(\log(\Lambda_{\max}P+\sigma^2) - \log(\Lambda'_{\min}P+\sigma^2)\right)}{R(Rn^N+(n+1)^N)+K(n+1)^N}, \tag{93}$$

$$= \log(\Lambda_{\max}P+\sigma^2) - \log(\Lambda'_{\min}P+\sigma^2). \tag{94}$$

Combining Equations (90) and (94) in Equation (87) and taking the limit $n \to \infty$, we get the following:

$$\lim_{n\to\infty} R_1 = \frac{(R+1)\log(\lambda_{\min}P+\sigma^2) - \left(\frac{R(R+1)}{K-1}+1\right)\log(\lambda'_{\max}P+\sigma^2)}{R(R+1)+K} - \log(\Lambda_{\max}P+\sigma^2) + \log(\Lambda'_{\min}P+\sigma^2). \tag{95}$$

Dividing R_1 by $\log(P)$ and letting $P \to \infty$, we get

$$d_1 = \lim_{P\to\infty}\frac{R_1}{\log(P)} = \frac{(R+1) - \left(\frac{R^2+R}{K-1}+1\right)}{R(R+1)+K}, = \frac{R(K-R-2)}{(K-1)\times[R(R+1)+K]}. \tag{96}$$

Therefore, the achievable secure sum degrees of freedom ($\text{SDoF}^{\text{ach.}}_{\text{IC-CM}}$) is obtained as

$$\text{SDoF}^{\text{ach.}}_{\text{IC-CM}} = \frac{KR(K-R-2)}{(K-1)\times[R(R+1)+K]}. \tag{97}$$

Hence, this completes the proof of Theorem 1.

5. Proof of Theorem 2

We follow a similar achievability scheme presented in Section 4, however, the main differences are the number of information symbols, the artificial noises used for transmission and the number of rounds in the first phase of the scheme. The goal of each transmitter is to securely send $T_1 = Rn^N + (n+1)^N - \lceil\frac{RT+K(n+1)^N}{K}\rceil$ information symbols to its corresponding receiver and keeping all messages secure against the external eavesdropper.

The total number of equations seen at the eavesdropper is $RT + K(n+1)^N$ of $\{\mathbf{s}_i\}_{i=1}^K$, $\{\mathbf{u}_i\}_{i=1}^K$. Hence, in order to keep the unintended information symbols of K transmitters at this receiver secure, we require that the number of these equations must be at most equal to the total number of the artificial noise dimensions of the K transmitters, i.e.,

$$RT + K(n+1)^N \leq KT_2. \tag{98}$$

Therefore, we choose T_2 as

$$T_2 = \left\lceil \frac{RT + K(n+1)^N}{K} \right\rceil. \tag{99}$$

Since $T = T_1 + T_2$, so we can get T_1 as follows:

$$T_1 = Rn^N + (n+1)^N - \left\lceil \frac{RT + K(n+1)^N}{K} \right\rceil, \quad T = Rn^N + (n+1)^N. \tag{100}$$

To this end, this scheme leads to the following achievable SDoF:

$$\text{SDoF}_{\text{IC-EE}}^{\text{ach.}} \overset{(b)}{=} \frac{R(K-R-1)}{R(R+1)+K}. \tag{101}$$

Since the achieved SDoF in Equation (101) is a concave function of R. Hence, getting the optimal R^* is obtained by equating the first derivative of the function with zero. Therefore, the optimal R^* is

$$R^* = \lfloor \sqrt{K} \rfloor - 1, \tag{102}$$

$$> \sqrt{K} - 2. \tag{103}$$

Now we approximate the obtained SDoF as follows:

$$\text{SDoF}_{\text{IC-EE}}^{\text{ach.}} = \frac{(\sqrt{K}-2)(K-\sqrt{K}+1)}{(\sqrt{K}-2)(\sqrt{K}-1)+K}, \tag{104}$$

$$\overset{(a)}{=} \frac{K\sqrt{K} - 3K + 3\sqrt{K} - 2}{2K - 3\sqrt{K} + 2}, \tag{105}$$

$$> \frac{K\sqrt{K} - 3K + 3\sqrt{K} - 2}{2K}, \tag{106}$$

$$\overset{(b)}{=} \frac{\sqrt{K}-3}{2} + \frac{3\sqrt{K}-2}{2K}, \tag{107}$$

$$> \frac{1}{2}(\sqrt{K}-3)^+, \tag{108}$$

where in (a), the term $-3\sqrt{K}+2$ in the denominator is negative, $\forall K \geq 1$, so neglecting this term gives us Equation (106). In step (b), since the term $3\sqrt{K}-2$ is positive, hence omitting this term gives Equation (108).

Secrecy Rate and SDoF Calculation

For a transmission of block length $B = RT + K(n+1)^N$, the achievable secure rate $R_i, i = 1, 2, \ldots, K$ is defined as

$$R_i = \frac{I(\mathbf{s}_i; \mathbf{y}_i|\Omega) - I(\mathbf{s}_i; \mathbf{z}|\mathbf{s}_{-i}^K, \Omega)}{RT + K(n+1)^N}, \; i = 1, 2, \ldots, K, \tag{109}$$

where $I(\mathbf{s}_i; \mathbf{y}_i|\Omega)$ is the mutual information between the information symbols vector $\mathbf{s}_i$ and $\mathbf{y}_i$, the received composite signal vector at the intended receiver i, given the knowledge of the channel coefficients. $I(\mathbf{s}_i; \mathbf{z}|\mathbf{s}_{-i}^K, \Omega)$ is the mutual information between $\mathbf{s}_i$ and $\mathbf{z}$, the received composite signal vector at the external eavesdropper, conditioned on the information symbols $\mathbf{s}_{-i}^K$. Note that $\mathbf{z}$ is collectively written as

$$\mathbf{z} = \mathbf{A}_z \mathbf{V} \mathbf{q}_z + \mathbf{n}_z, \tag{110}$$

where,

$$\mathbf{A}_z = \begin{bmatrix} \mathbf{C}_z \\ \hline \mathbf{D}_z \end{bmatrix}, \ \mathbf{C}_z = \begin{bmatrix} \mathbf{G}_1^1 & \mathbf{G}_2^1 & \cdots & \mathbf{G}_K^1 \\ \mathbf{G}_1^2 & \mathbf{G}_2^2 & \cdots & \mathbf{G}_K^2 \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{G}_1^R & \mathbf{G}_2^R & \cdots & \mathbf{G}_K^R \end{bmatrix},$$
$$\mathbf{D}_z = \text{blkdiag}(\tilde{\mathbf{G}}_1 \mathbf{X}, \ldots, \tilde{\mathbf{G}}_K \mathbf{X}). \tag{111}$$

where $\mathbf{A}_z$ has dimensions of $(RT + K(n+1)^N) \times KT = KT_2 \times KT$. Each $\mathbf{G}$ is a matrix represents the channel gains between each transmitter and the external eavesdropper.

The analysis of the achievable secure rate and SDoF follows similar steps as those in Section 4.3. This completes the proof of Theorem 2.

6. Proof of Theorem 3

We follow the same transmission scheme presented in Section 4. The goal of each transmitter is to securely send $T_1 = Rn^N + (n+1)^N - \lceil \frac{RT+(K-1)(n+1)^N}{K-1} \rceil$ information symbols to its corresponding receiver and keeping all messages secure against the external eavesdropper and the unintended receivers.

We have two secrecy constraints must be satisfied, i.e.,

$$RT + (K-1)(n+1)^N \leq (K-1)T_2, \text{ (confidential messages)}, \tag{112}$$
$$RT + K(n+1)^N \leq KT_2, \text{ (eavesdropper)}, \tag{113}$$

Equation (113) can be re-written as

$$RT + (K-1)(n+1)^N + (n+1)^N \leq (K-1)T_2 + T_2. \tag{114}$$

So if we pick T_2 as

$$T_2 = \left\lceil \frac{RT + (K-1)(n+1)^N}{K-1} \right\rceil \tag{115}$$

we need to check that $T_2 \geq (n+1)^N$. T_2 can be written as

$$T_2 = \left\lceil \frac{R(Rn^N + (n+1)^N) + (K-1)(n+1)^N}{K-1} \right\rceil, \tag{116}$$
$$> \frac{R(Rn^N + (n+1)^N) + (K-1)(n+1)^N}{K-1} - 1, \tag{117}$$
$$= \frac{R^2 n^N + R(n+1)^N + (K-1)(n+1)^N}{K-1} - 1, \tag{118}$$
$$\overset{(a)}{=} \frac{R^2}{(K-1)} n^N + \frac{R}{K-1}(n+1)^N + (n+1)^N - 1 > (n+1)^N, \tag{119}$$

where in (a), the first two terms are positive, hence, T_2 is strictly greater than $(n+1)^N$. To this end, we conclude that the two secrecy constraints are satisfied. Hence, we achieve the same SDoF of Theorem 1, i.e.,

$$\text{SDoF}^{\text{ach.}}_{\text{IC-CM-EE}} = \frac{KR(K-R-2)}{(K-1)\times[R(R+1)+K]} > \frac{1}{2}(\sqrt{K}-6)^{+}. \tag{120}$$

Secrecy Rate and SDoF Calculation

For a transmission of block length $B = RT + K(n+1)^N$, the achievable secure rate $R_i, i = 1,2,\ldots,K$ is defined as

$$R_i = \frac{I(\mathbf{s}_i;\mathbf{y}_i|\Omega) - \max\left[\max_{j\neq i} I(\mathbf{s}_i;\mathbf{y}_j|\mathbf{s}^K_{-i},\Omega), I(\mathbf{s}_i;\mathbf{z}|\mathbf{s}^K_{-i},\Omega)\right]}{RT+K(n+1)^N},\ i=1,2,\ldots,K. \tag{121}$$

The analysis of the achievable secure rate and SDoF follows similar steps as those in Section 4.3. This completes the proof of Theorem 3.

7. Conclusions

In this paper, we studied the K-user interference channel with three secrecy constrained channel models and delayed CSIT: We showed that for the K-user interference channel with confidential messages, the sum secure degrees of freedom (SDoF) is at least $\frac{1}{2}(\sqrt{K}-6)$, and scales with square root of the number of users. Also, we showed that for the K-user interference channel with an external eavesdropper, $\frac{1}{2}(\sqrt{K}-3)$ SDoF is achievable. For the K-user interference channel with confidential messages and an external eavesdropper, we showed that $\frac{1}{2}(\sqrt{K}-6)$ is achievable. To achieve these results, we have proposed novel secure retrospective interference alignment schemes which satisfy both secrecy and decodability at receivers. To the best of our knowledge, this is the first result showing scaling of SDoF for the interference channel with secrecy constraints and delayed CSIT. An interesting open problem is to investigate the optimality of these schemes, and finding upper bounds on SDoF with delayed CSIT for these channel models.

Author Contributions: Conceptualization, M.S., R.T. and M.L.; Formal analysis, M.S., R.T. and M.L.; Investigation, M.S., R.T. and M.L.; Methodology, M.S., R.T. and M.L.; Validation, M.S., R.T. and M.L.; Visualization, M.S., R.T. and M.L.; Writing—original draft, M.S., R.T. and M.L.; Writing—review and editing, M.S., R.T. and M.L.

Funding: This work of M.S., R.T., and M.L. was supported in part by U.S. NSF through grants CCF-1559758 and CNS-1715947, CNS-1564477, and ONR YIP grant N00014-16-1-2650, and was presented in part at ISIT conference, Colorado, USA, June 2018.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. (Proof of Lemma 1 and Corollary 4)

By taking the first derivative of Equation (20) with respect to the number of rounds R, we get

$$\frac{\partial}{\partial R}\ \text{SDoF}^{\text{ach.}}_{\text{IC-CM}}(K,R)) = \frac{K(K^2-K(R^2+2R+2)+R^2)}{(K-1)(K+R^2+R)^2}. \tag{A1}$$

For $R < R^*$, the function $\text{SDoF}^{\text{ach.}}_{\text{IC-CM}}(K, R)$ strictly increases and for $R > R^*$ the function strictly decreases, where R^* is given by

$$R^* = \frac{-K + K \times \sqrt{1 + \frac{(K-1)(K-2)}{K}}}{K-1}. \tag{A2}$$

Alternatively,

$$\frac{\partial}{\partial R} \text{SDoF}^{\text{ach.}}_{\text{IC-CM}}(K, R) > 0, \forall R < R^*, \tag{A3}$$

$$\frac{\partial}{\partial R} \text{SDoF}^{\text{ach.}}_{\text{IC-CM}}(K, R) < 0, \forall R > R^*. \tag{A4}$$

Figure A1 shows the behavior of the achievable sum SDoF as a function of the number of rounds R. The optimal value of R can be obtained by equating the first derivative of SDoF_{sum} to zero as follows:

$$\frac{\partial}{\partial R} \text{SDoF}^{\text{ach.}}_{\text{IC-CM}}(K, R) K, R) = \frac{\partial}{\partial R} \frac{KR(K - R - 2)}{(K-1) \times [R(R+1) + K]} = 0, \tag{A5}$$

$$\propto \frac{\partial}{\partial R} \frac{R(K - R - 2)}{R(R+1) + K} = 0. \tag{A6}$$

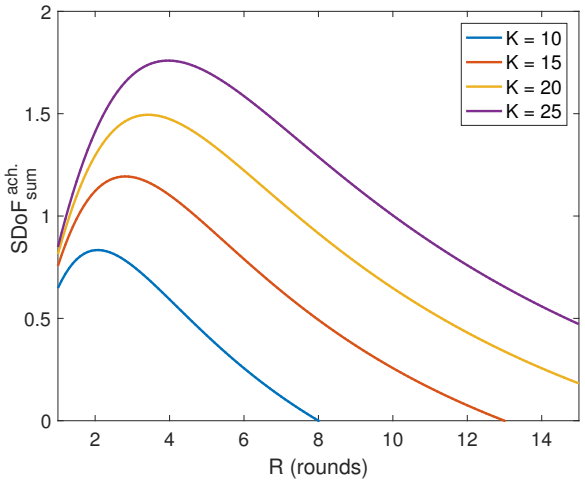


Figure A1. Plot for the achievable sum SDoF as a function of the number of rounds R for different number of user K.

After differentiating Equation (A6), we will get the following:

$$R^2(K-1) + 2KR - K(K-2) = 0. \tag{A7}$$

The solution of the previous equation is

$$R^* \overset{(a)}{=} \left\lfloor \frac{-K + K \times \sqrt{1 + \frac{(K-1)(K-2)}{K}}}{K-1} \right\rfloor, \tag{A8}$$

$$> \left\lfloor \frac{-K + K \times \sqrt{1 + \frac{(K-1)(K-2)}{K}}}{K} \right\rfloor, \tag{A9}$$

$$= \left\lfloor -1 + \sqrt{1 + \frac{(K-1)(K-2)}{K}} \right\rfloor, \tag{A10}$$

$$\overset{(b)}{>} \sqrt{\frac{(K-1)(K-2)}{K}} - 2, \tag{A11}$$

$$= \sqrt{\frac{K^2 - 3K + 2}{K}} - 2, \tag{A12}$$

$$> \sqrt{\frac{K^2 - 3K}{K}} - 2 \tag{A13}$$

$$= \sqrt{K-3} - 2, \tag{A14}$$

$$\overset{(c)}{>} \sqrt{K} - 3 - 2 \tag{A15}$$

$$= \sqrt{K} - 5 = R_{\text{lb}}, \tag{A16}$$

where in (a), since $R^* \in \mathbb{N}$, we apply the floor rounding operator on the obtained value of R. In (b), we used the property of the floor operator, i.e., $\lfloor x \rfloor > x - 1$. In (c), the term $\sqrt{K-3}$ is greater than $\sqrt{K} - 3, \forall K \geq 1$.

Appendix B. (Proof of Linear Independence in Equation (43))

In this Appendix, we show that by the end of the transmission scheme, each receiver gets $T = Rn^N + (n+1)^N$ linear independent equations of the desired signals (i.e., the information symbols and the artificial noises). Then, we need to show that the following matrix

$$\mathbf{B}_k = \left[\mathbf{X}^H, (\mathbf{W}^H \mathbf{H}_{kk}^1)^H, \ldots, (\mathbf{W}^H \mathbf{H}_{kk}^R)^H \right]^H \tag{A17}$$

is full rank. Since $\mathbf{B}_k$ is a square matrix, then it is sufficient to show that $\det(\mathbf{B}_k) \neq 0, \forall k = 1, 2, \ldots, K$. Without loss of generality, we consider receiver 1 which has the following matrix

$$\mathbf{B}_1 = \left[\mathbf{X}^H, (\mathbf{W}^H \mathbf{H}_{11}^1)^H, \ldots, (\mathbf{W}^H \mathbf{H}_{11}^R)^H \right]^H. \tag{A18}$$

Since $\det(\mathbf{B}_1) = \det(\mathbf{B}_1^H)$, we will instead show that $\det(\mathbf{B}_1^H) \neq 0$, which is given as follows:

$$\mathbf{B}_1^H = \left[\mathbf{X}^H, (\mathbf{H}_{11}^1)^H \mathbf{W}, \ldots, (\mathbf{H}_{11}^R)^H \mathbf{W} \right]. \tag{A19}$$

Note that $\mathbf{W}$ and $\mathbf{X}$ are function of the diagonal entries of the channels $\{(\mathbf{H}_{kj}^r)^H\}_{k \neq j}$, $\forall k, j = 1, \ldots, K$ and $r = 1, 2, \ldots, R$. More specifically, the entries of $(\mathbf{H}_{kj}^r)^H$ are $h_{kj}^r(t), \forall t = 1, 2, \ldots, T$. $\mathbf{B}_1$ depends on $\{(\mathbf{H}_{kj}^r)^H\}_{k \neq j}$ plus $(\mathbf{H}_{11}^r)^H$ whose elements are $h_{11}^r(t), \forall t = 1, \ldots, T$ and, $\forall r = 1, \ldots, R$. For notation convenience, let us denote these channel coefficients as $\mu_{mt}, \forall t = 1, 2, \ldots, T$ and $\forall m = 1, 2, \ldots, N + R$. The elements of $\mathbf{B}_1^H$ are written as a monomial function of the random variables $\mu_{it}, \forall i = 1, \ldots, N + R$ and $\forall t = 1, 2, \ldots, T$ as follows:

$$B_1^H(t, p) = \prod_{i=1}^{N+R} (\mu_{it})^{n_i(p)}, \tag{A20}$$

where $n_i(p) \in \mathbb{Z}_+$ is the exponent of the random variable μ_{it}. Note that for two different columns p_1 and p_2, $(n_1(p_1), n_2(p_1), \ldots, n_{N+R}(p_1)) \neq (n_1(p_2), n_2(p_2), \ldots, n_{N+R}(p_2))$. More specifically, the structure of $\mathbf{X}^H$ is as follows:

$$\mathbf{X}^H = \begin{bmatrix} 1 & \mu_{11} & \mu_{21} & \cdots & \mu_{11}^n \mu_{21}^n \ldots \mu_{N1}^n \\ 1 & \mu_{12} & \mu_{22} & \cdots & \mu_{12}^n \mu_{22}^n \ldots \mu_{N2}^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \mu_{1T}^n & \mu_{2T}^n & \cdots & \mu_{1T}^n \mu_{2T}^n \ldots \mu_{NT}^n \end{bmatrix}, \tag{A21}$$

and for $(\mathbf{H}_{11}^r)^H \mathbf{W}$ as

$$(\mathbf{H}_{11}^r)^H \mathbf{W} = \begin{bmatrix} \kappa_1^r & \kappa_1^r \mu_{11} & \kappa_1^r \mu_{21} & \cdots & \kappa_1^r \chi_1^{n-1} \\ \kappa_2^r & \kappa_2^r \mu_{12} & \kappa_2^r \mu_{22} & \cdots & \kappa_2^r \chi_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \kappa_T^r & \kappa_T^r \mu_{1T}^n & \kappa_T^r \mu_{2T}^n & \cdots & \kappa_T^r \chi_T^{n-1} \end{bmatrix}, \tag{A22}$$

where $\kappa_t^r = \mu_{N+r,t}, \forall t = 1, 2, \ldots, T, \forall r = 1, 2, \ldots, R$ and $\chi_t^{n-1} = \mu_{1t}^n \mu_{2t}^n \ldots \mu_{Nt}^n$. The full matrix $\mathbf{B}_1^H$ is written as follows:

$$\mathbf{B}_1^H = \begin{bmatrix} 1 & \cdots & \chi_1^n & \kappa_1^1 & \cdots & \kappa_1^1 \chi_1^{n-1} & \cdots & \kappa_1^R & \cdots & \kappa_1^R \chi_1^{n-1} \\ 1 & \cdots & \chi_2^n & \kappa_2^1 & \cdots & \kappa_2^1 \chi_2^{n-1} & \cdots & \kappa_2^R & \cdots & \kappa_2^R \chi_2^{n-1} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 1 & \cdots & \chi_T^n & \kappa_T^1 & \cdots & \kappa_T^1 \chi_T^{n-1} & \cdots & \kappa_T^R & \cdots & \kappa_T^R \chi_T^{n-1} \end{bmatrix}. \tag{A23}$$

The determinant of matrix $\mathbf{B}_1^H$ can be written as follows:

$$\det(\mathbf{B}_1^H) = a_{1,1} C_{1,1} + a_{1,2} C_{1,2} + \cdots + a_{1,T} C_{1,T}, \tag{A24}$$

where $\mathbf{C}_{1,j}$ is the cofactor matrix corresponding after removing the 1st row and the jth column with coefficient $a_{1,j}$. Now we will show that $\mathbf{B}_1^H$ is full rank by contradiction. The zero determinant assumption implies one of the following two events:

1. $\mu_{m1}, m \in \{1, 2, \ldots, N+R\}$ takes a value equal to one of the roots of the polynomial equation.
2. All the cofactors of the polynomials are zero.

For the first event, none of the cofactors depends on the random variables $\mu_{m1}, m \in \{1, 2, \ldots, N+R\}$. Note that $\mu_{m1}, m \in \{1, 2, \ldots, N+R\}$ are drawn from a continuous distribution, then the probability of these random variables that take finitely many values as a solution for the polynomial is zero almost surely. Therefore, the second event happens with probability greater than zero, which implies

$$C_{1,p} = 0, \forall p \in \{1, \ldots, T\}. \tag{A25}$$

Then $C_{1,T} = 0$ with probability higher than zero. $C_{1,T} = 0$ implies that the determinant of the matrix obtained by stripping off the first row and last column of $\mathbf{B}_1^H$ is equal to zero with non-zero probability. Repeating the process of stripping off each row and column, it will end up with 1×1 matrix with value one which contradicts the assumption that $C_{1,T} = 0$. It is worth noting that stripping off the rows and columns procedure preserves the structure of the matrix which means that the cofactors do not depend on the coefficients. To conclude, the determinant of $\mathbf{B}_1^H$ does not equal zero almost surely, which implies that the desired symbols are decoded successfully at the receiver side with probability one.

Appendix C. (Proof of Lemma 2)

Note that $\mathbf{AX}+\mathbf{N}$ is a jointly complex Gaussian vector with zero mean and covariance $P\mathbf{A}\mathbf{A}^H + \sigma^2\mathbf{I}$. From [27], $h(\mathbf{AX}+\mathbf{N})$ is written as

$$h(\mathbf{AX}+\mathbf{N}) = \log(\pi e)^M \det\left(P\mathbf{A}\mathbf{A}^H + \sigma^2\mathbf{I}_M\right). \tag{A26}$$

It is worth noting that $\mathbf{A}\mathbf{A}^H$ is positive semi-definite, with eigenvalue decomposition $\mathbf{Q}\mathbf{D}\mathbf{Q}^H$, where $\mathbf{D}$ is a diagonal matrix with r non-zero eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_r$ where $r = \text{rank}(\mathbf{A}\mathbf{A}^H) = \text{rank}(\mathbf{A})$. Then,

$$h(\mathbf{AX}+\mathbf{N}) = \log(\pi e)^M \det\left(P\mathbf{Q}\mathbf{D}\mathbf{Q}^H + \sigma^2\mathbf{I}_M\right). \tag{A27}$$

Next, we use the Sylvester's identity for determinants, i.e., $\det(\mathbf{I}+\mathbf{AB}) = \det(\mathbf{I}+\mathbf{BA})$. Using this identity, we have

$$h(\mathbf{AX}+\mathbf{N}) = \log(\pi e)^M \det\left(P\mathbf{D}\mathbf{Q}^H\mathbf{Q} + \sigma^2\mathbf{I}_M\right). \tag{A28}$$

Since $\mathbf{Q}$ is a unitary matrix, (i.e., $\mathbf{Q}^H\mathbf{Q} = \mathbf{Q}\mathbf{Q}^H = \mathbf{I}$), we have the following:

$$h(\mathbf{AX}+\mathbf{N}) = \log(\pi e)^M \prod_{i=1}^{r}\left(\lambda_i P + \sigma^2\right), \tag{A29}$$

$$= \log(\pi e)^M + \sum_{i=1}^{r}\log\left(\lambda_i P + \sigma^2\right). \tag{A30}$$

This completes the proof of Lemma 2.

References

1. Maddah-Ali, M.A.; Tse, D. Completely stale transmitter channel state information is still very useful. *IEEE Trans. Inf. Theory* **2012**, *58*, 4418–4431.
2. Cadambe, V.R.; Jafar, S.A. Interference alignment and degrees of freedom of the *K*-user interference channel. *IEEE Trans. Inf. Theory* **2008**, *54*, 3425–3441. [CrossRef]
3. Ghasemi, A.; Motahari, A.S.; Khandani, A.K. On the degrees of freedom of X channel with delayed CSIT. In Proceedings of the IEEE International Symposium on Information Theory (ISIT), St. Petersburg, Russia, 31 July–5 August 2011; pp. 767–770.
4. Maggi, L.; Cottatellucci, L. Retrospective interference alignment for interference channels with delayed feedback. In Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC), Shanghai, China, 1–4 April 2012; pp. 453–458.
5. Abdoli, M.J.; Ghasemi, A.; Khandani, A.K. On the degrees of freedom of *K*-user SISO interference and X channels with delayed CSIT. *IEEE Trans. Inf. Theory* **2013**, *59*, 6542–6561. [CrossRef]
6. Kang, M.G.; Choi, W. Ergodic interference alignment with delayed feedback. *IEEE Signal Process. Lett.* **2013**, *20*, 511–514. [CrossRef]
7. Maleki, H.; Jafar, S.A.; Shamai, S. Retrospective interference alignment over interference networks. *IEEE J. Sel. Top. Signal Process.* **2012**, *6*, 228–240. [CrossRef]
8. Castanheira, D.; Silva, A.; Gameiro, A. Retrospective Interference Alignment: Degrees of Freedom Scaling With Distributed Transmitters. *IEEE Trans. Inf. Theory* **2017**, *63*, 1721–1730. [CrossRef]
9. Yener, A.; Ulukus, S. Wireless physical-layer security: Lessons learned from information theory. *Proc. IEEE* **2015**, *103*, 1814–1825. [CrossRef]
10. Mukherjee, P.; Tandon, R.; Ulukus, S. *Physical Layer Security with Delayed, Hybrid and Alternating Channel State Knowledge, Information Theoretic Security and Privacy of Information Sources*; Cambridge University Press: Cambridge, UK, 2017.

11. Liang, Y.; Poor, H.V.; Shamai, S. Information theoretic security. *Found. Trends® Commun. Inf. Theory* **2009**, *5*, 355–580. [CrossRef]
12. He, X.; Yener, A. *K*-user interference channels: Achievable secrecy rate and degrees of freedom. In Proceedings of the IEEE Information Theory Workshop (ITW) on Networking and Information Theory, Volos, Greece, 10–12 June 2009; pp. 336–340.
13. Koyluoglu, O.O.; El Gamal, H.; Lai, L.; Poor, H.V. Interference alignment for secrecy. *IEEE Trans. Inf. Theory* **2011**, *57*, 3323–3332. [CrossRef]
14. Xie, J.; Ulukus, S. Real interference alignment for the *K*-user Gaussian interference compound wiretap channel. In Proceedings of the 48th Annual Allerton Conference on Communication, Control, and Computing, Allerton, IL, USA, 29 September–1 October 2010; pp. 1252–1257.
15. Bassily, R.; Ulukus, S. Ergodic secret alignment. *IEEE Trans. Inf. Theory* **2012**, *58*, 1594–1611. [CrossRef]
16. Gou, T.; Jafar, S.A. On the secure degrees of freedom of wireless X networks. In Proceedings of the 46th Annual Allerton Conference on Communication, Control, and Computing, Urbana-Champaign, IL, USA, 23–26 September 2008; pp. 826–833.
17. Nafea, M.; Yener, A. How many antennas does a cooperative jammer need for achieving the degrees of freedom of multiple antenna Gaussian channels in the presence of an eavesdropper? In Proceedings of the 51st Annual Allerton Conference on Communication, Control, and Computing, Monticello, IL, USA, 2–4 October 2013; pp. 774–780.
18. Yang, S.; Kobayashi, M. Secure communication in *K*-user multi-antenna broadcast channel with state feedback. In Proceedings of the IEEE International Symposium on Information Theory (ISIT), Hong Kong, China, 14–19 June 2015; pp. 1976–1980.
19. Xie, J.; Ulukus, S. Secure degrees of freedom of *K*-user Gaussian interference channels: A unified view. *IEEE Trans. Inf. Theory* **2015**, *61*, 2647–2661. [CrossRef]
20. Xie, J.; Ulukus, S. Secure degrees of freedom of one-hop wireless networks. *IEEE Trans. Inf. Theory* **2014**, *60*, 3359–3378. [CrossRef]
21. Mukherjee, P.; Ulukus, S. Secrecy in MIMO Networks with no eavesdropper CSIT. *IEEE Trans. Commun.* **2017**, *65*, 4382–4391. [CrossRef]
22. Mukherjee, P.; Tandon, R.; Ulukus, S. Secure degrees of freedom region of the two-user MISO broadcast channel with alternating CSIT. *IEEE Trans. Inf. Theory* **2017**, *63*, 3823–3853. [CrossRef]
23. Liang, Y.; Kramer, G.; Poor, H.V.; Shamai, S. Compound wiretap channels. *EURASIP J. Wirel. Commun. Netw.* **2009**, *2009*, 5. [CrossRef]
24. Liu, R.; Maric, I.; Spasojevic, P.; Yates, R.D. Discrete memoryless interference and broadcast channels with confidential messages: Secrecy rate regions. *IEEE Trans. Inf. Theory* **2008**, *54*, 2493–2507. [CrossRef]
25. Horn, R.A.; Johnson, C.R. *Matrix Analysis*; Cambridge University Press: Cambridge, UK, 2012.
26. Seif, M.; Tandon, R.; Li, M. Secure Retrospective Interference Alignment. *arXiv* **2018**, arXiv:1801.03494.
27. Cover, T.M.; Thomas, J.A. *Elements of Information Theory*; John Wiley & Sons: Hoboken, NJ, USA, 2012.