Secure Degrees of Freedom for the MIMO Wire-tap Channel with a Multi-antenna Cooperative Jammer

Mohamed Nafea, *Member, IEEE,* and Aylin Yener, *Fellow, IEEE*

***Abstract*—In this paper, a multiple antenna wire-tap channel in the presence of a multi-antenna cooperative jammer is studied. In particular, the secure degrees of freedom (s.d.o.f.) of this channel is established, with N_t antennas at the transmitter, N_r antennas at the legitimate receiver, and N_e antennas at the eavesdropper, for all possible values of the number of antennas, N_c, at the cooperative jammer. In establishing the result, several different ranges of N_c need to be considered separately. The lower and upper bounds for these ranges of N_c are derived, and are shown to be tight. The achievability techniques developed rely on a variety of signaling, beamforming, and alignment techniques which vary according to the (relative) number of antennas at each terminal and whether the s.d.o.f. is integer valued. Specifically, it is shown that, whenever the s.d.o.f. is integer valued, Gaussian signaling for both transmission and cooperative jamming, linear precoding at the transmitter and the cooperative jammer, and linear processing at the legitimate receiver, are sufficient for achieving the s.d.o.f. of the channel. By contrast, when the s.d.o.f. is not an integer, the achievable schemes need to rely on structured signaling at the transmitter and the cooperative jammer, and joint signal space and signal scale alignment. The converse is established by combining an upper bound which allows for full cooperation between the transmitter and the cooperative jammer, with another upper bound which exploits the secrecy and reliability constraints.**

***Index Terms*—MIMO wiretap channel, cooperative jamming, secure degrees of freedom, beamforming, structured signaling, real interference alignment.**

I. Introduction

INFORMATION theoretically secure message transmission in noisy communication channels was first considered in the seminal work by Wyner [3]. Reference [4] subsequently identified the secrecy capacity of a general discrete memoryless wire-tap channel. Reference [5] studied the Gaussian wire-tap channel and its secrecy capacity. More recently, an extensive body of work was devoted to study a variety of network information theoretic models under secrecy constraint(s), see for example [6]–[23]. The secrecy capacity region for most of multi-terminal models remain open despite significant progress on bounds and associated insights. Recent work thus includes efforts that concentrate on characterizing the more tractable high signal-to-noise ratio (SNR) scaling behavior of secrecy capacity region for Gaussian multi-terminal models [21]–[26].

Among the multi-transmitter models studied, a recurrent theme in achievability is enlisting one or more terminals to transmit intentional interference with the specific goal of diminishing the reception capability of the eavesdropper, known as *cooperative jamming* [27]. For the Gaussian wire-tap channel, adding a cooperative jammer terminal transmitting Gaussian noise can improve the secrecy rate considerably [11], albeit not the scaling of the secrecy capacity with power at high SNR. Recently, reference [23] has shown that, for the Gaussian wire-tap channel, adding a cooperative jammer and utilizing structured codes for message transmission and cooperative jamming, provide an achievable secrecy rate scalable with power, i.e., a positive secure degrees of freedom (s.d.o.f.), an improvement from the zero degrees of freedom of the Gaussian wire-tap channel. More recently, reference [24] has proved that, for this channel, the s.d.o.f. $\frac{1}{2}$, achievable by codebooks constructed from integer lattices along with real interference alignment, is tight. References [25], [26] have subsequently identified the s.d.o.f. region for multi-terminal Gaussian wire-tap channel models.

While the above development is for single-antenna terminals, multiple antennas have also been utilized to improve secrecy rates and s.d.o.f. for several channel models, see for example [7]–[9], [21], [28]–[33]. The secrecy capacity of the multi-antenna (MIMO) wire-tap channel, identified in [28], scales with power only when the legitimate transmitter has an advantage over the eavesdropper in the number of antennas. It then follows naturally to utilize a cooperative jamming terminal to improve the secrecy rate and scaling for multi-antenna wire-tap channels as well which is the focus of this work.

In this paper, we study the multi-antenna wire-tap channel with a multi-antenna cooperative jammer. We characterize the high SNR scaling of the secrecy capacity, i.e., the s.d.o.f., of the channel with N_c antennas at the cooperative jammer, N_t antennas at the transmitter, N_r antennas at the receiver, N_e antennas at the eavesdropper, under the assumption of known channel state information at all terminals. The achievability and converse techniques both are methodologically developed for ranges of the parameters, i.e., the number of antennas at each terminal. The upper and lower bounds for all parameter

Manuscript received November 26, 2015; revised Febreuary 14, 2016; accepted July 17, 2017. Date of publication July 31, 2017. This work was supported by NSF Grants CCF 13-19338 and CNS 13-14719. This paper was presented in part at the 2014 IEEE Information Theory Workshop [1], and the 2015 IEEE International Conference on Communications [2].

M. Nafea is with the Department of Electrical Engineering, Pennsylvania State University, University Park, PA 16802 USA (e-mail: mnafea@psu.edu),

A. Yener is with the Department of Electrical Engineering, Pennsylvania State University, University Park, PA 16802 USA (e-mail: yener@ee.psu.edu).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2738011, IEEE Transactions on Information Theory

values are shown to match one another. Presenting this work in part, [1] and [2] are conference presentations reporting the resulting s.d.o.f. for $N_t = N_r$ only. Note that the s.d.o.f. results in this paper match the achievability results reported in [34], [35], which are special cases for $\{N_t = N_r = 1,\ N_c = N_e\}$, $\{N_t = N_r = N_e = N,\ N_c = 2N\}$, $\{N_t = N_r = N_e = N,\ N_c = 2N-1\}$, and with real channel gains.

We remark that secure degrees of freedom for single and multiple antenna wire-tap channels have recently been investigated under the assumption of unknown eavesdropper channel state information at the legitimate terminals. The secure degrees of freedom for the single-antenna wire-tap channel with multiple helpers, multiple-access wire-tap channel, and interference wire-tap channel, with unknown and static eavesdropper channel, have been derived in [36]. The strongly secure degrees of freedom of the multiple antenna wire-tap channel with unknown and *varying* eavesdropper channel is established in [32] by showing the existence of a universal scheme that can counter any eavesdropper state. [32] thus quantifies the reduction in degrees of freedom that results from universal immunity to eavesdropping. This paper, by contrast, addresses the improvement provided by adding a *multi-antenna helper* in the benchmark case that is the static and known channel state information for the MIMO wiretap channel.

The proposed achievable schemes for different ranges of the values for N_c, N_t, N_r, and N_e all involve linear precoding and linear receiver processing. The common goal to all these schemes is to perfectly align the cooperative jamming signals over the information signals observed at the eavesdropper while simultaneously enabling information and cooperative jamming signal separation at the legitimate receiver. We show that whenever the s.d.o.f. of the channel is integer valued, Gaussian signaling both at the transmitter and the cooperative jammer suffices to achieve the s.d.o.f. By contrast, non-integer s.d.o.f. requires structured signaling along with joint signal space and signal scale alignment in the complex plane [37], [38]. The necessity of structured signaling follows from the fact that fractional s.d.o.f. indicates sharing at least one spatial dimension between information and cooperative jamming signals at the receiver's signal space. In this case, sharing the same spatial dimension between Gaussian information and jamming signals, which have similar power scaling, does not provide positive degrees of freedom, and we need for structured signals that can be separated over this single dimension at high SNR. The tools that enable the signal scale alignment are available in the field of transcendental number theory [38]–[40], which we utilize.

The paper is organized as follows. Section II introduces the channel model, and Section III provides the main results. For clarity of exposition, we first present the converse and achievability for the MIMO wire-tap channel with $N_t = N_r = N$ in Sections IV and V. Section VI then extends the converse and achievability proofs for the case $N_t \neq N_r$. Section VII discusses the results of this work and Section VIII concludes the paper.

Overall, this study determines the value in jointly utilizing signal scale and spatial interference alignment techniques for secrecy and quantifies the impact of a multi-antenna helper for the MIMO wire-tap channel by settling the question of the secrecy prelog for the $(N_t \times N_r \times N_e)$ MIMO wire-tap channel in the presence of an N_c-antenna cooperative jammer, for all possible values of N_c. In contrast with the single antenna case, where integer lattice codes and real interference alignment suffice to achieve the s.d.o.f. of the channel, in the MIMO setting, one needs to utilize a variety of signaling, beamforming, and alignment techniques, in order to coordinate the transmitted and received signals for different values of N_t, N_r, N_e, and N_c.

II. Channel Model and Definitions

First, we remark the notation we use throughout the paper: Small letters denote scalars and capital letters denote random variables. Vectors are denoted by bold small letters, while matrices and random vectors are denoted by bold capital letters[1]. Sets are denoted using calligraphic fonts. All logarithms are taken to be base 2. Throughout the paper, we use $j = \sqrt{-1}$ to denote the imaginary unit in a complex number. $\mathbb{R}, \mathbb{C}$, $\mathbb{Q}$, and $\mathbb{Z}$ denote the sets of real, complex, rational, and real-valued integer numbers, respectively. $\mathbb{Z}_{\mathbb{C}}$ denotes the set of complex integers, i.e., $\mathbb{Z}_{\mathbb{C}} \triangleq \{n + jm : n, m \in \mathbb{Z}\}$. The set of integers $\{-Q, \cdots, Q\}$ is denoted by $(-Q, Q)_{\mathbb{Z}}$. $\mathbf{0}_{m\times n}$ denotes an $m \times n$ matrix of zeros, and $\mathbf{I}_n$ denotes an $n \times n$ identity matrix. For matrix $\mathbf{A}$, $\mathcal{N}(\mathbf{A})$ denotes its null space, $\det(\mathbf{A})$ denotes its determinant, and $||\mathbf{A}||$ denotes its *induced* norm. For vector $\mathbf{V}$, $||\mathbf{V}||$ denotes its Euclidean norm, and $\mathbf{V}_i^k$ denotes the ith to kth components in $\mathbf{V}$. We use $\mathbf{V}^n$ to denote the n-letter extension of the random vector $\mathbf{V}$, i.e., $\mathbf{V}^n = [\mathbf{V}(1) \cdots \mathbf{V}(n)]$. The operators ^{T}, ^{H}, and † denote the transpose, Hermitian, and pseudo inverse operations. A circularly symmetric Gaussian random vector with zero mean and covariance matrix $\mathbf{K}$ is denoted by $\mathcal{CN}(\mathbf{0}, \mathbf{K})$.

As the channel model, we consider the MIMO wire-tap channel with an N_t-antenna transmitter, N_r-antenna receiver, N_e-antenna eavesdropper, and an N_c-antenna cooperative jammer as depicted in Fig. 1. The received signals at the receiver and eavesdropper, at the nth channel use, are given by

$$\mathbf{Y}_r(n) = \mathbf{H}_t\mathbf{X}_t(n) + \mathbf{H}_c\mathbf{X}_c(n) + \mathbf{Z}_r(n) \quad (1)$$

$$\mathbf{Y}_e(n) = \mathbf{G}_t\mathbf{X}_t(n) + \mathbf{G}_c\mathbf{X}_c(n) + \mathbf{Z}_e(n), \quad (2)$$

where $\mathbf{X}_t(n)$ and $\mathbf{X}_c(n)$ are the transmitted signals from the transmitter and the cooperative jammer at the nth channel use. $\mathbf{H}_t \in \mathbb{C}^{N_r\times N_t}$, $\mathbf{H}_c \in \mathbb{C}^{N_r\times N_c}$ are the channel gain matrices from the transmitter and the cooperative jammer to the receiver, while $\mathbf{G}_t \in \mathbb{C}^{N_e\times N_t}$, $\mathbf{G}_c \in \mathbb{C}^{N_e\times N_c}$ are the channel gain matrices from the transmitter and the cooperative jammer to the eavesdropper. It is assumed that the channel gains are drawn independently from a *complex-valued* continuous distribution. All channel gains are assumed to be known at all terminals. $\mathbf{Z}_r(n)$ and $\mathbf{Z}_e(n)$ are the complex Gaussian noise at the receiver and eavesdropper at the nth channel use, where $\mathbf{Z}_r(n) \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_r})$ and $\mathbf{Z}_e(n) \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_e})$ for all n. $\mathbf{Z}_r(n)$ is independent from $\mathbf{Z}_e(n)$ and both are independent

[1]The distinction between matrices and random vectors is clear from the context.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2738011, IEEE Transactions on Information Theory

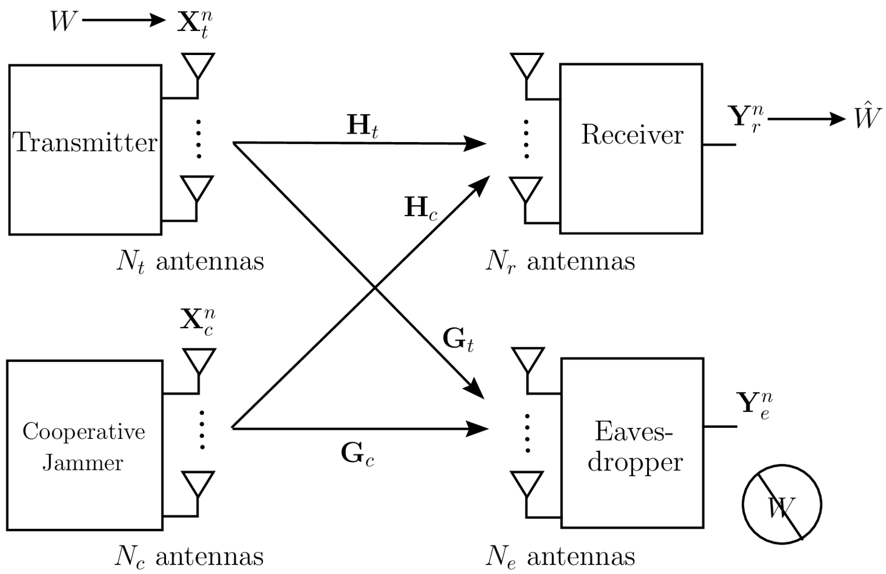


Fig. 1: $(N_t \times N_r \times N_e)$ multiple antenna wire-tap channel with an N_c-antenna cooperative jammer.

and identically distributed (i.i.d.) across the time index[2] n. The power constraints on the transmitted signals at the transmitter and the cooperative jammer are $\mathrm{E}\left(\mathbf{X}_t^H\mathbf{X}_t\right), \mathrm{E}\left(\mathbf{X}_c^H\mathbf{X}_c\right) \leq P$.

The transmitter aims to send a message W to the receiver, and keep it secret from the external eavesdropper. A stochastic encoder, which maps the message W to the transmitted signal $\mathbf{X}_t^n \in \mathcal{X}_t^n$, is used at the transmitter. The receiver uses its observation, $\mathbf{Y}_r^n \in \mathcal{Y}_r^n$, to obtain an estimate $\hat{W}$ of the transmitted message. Secrecy rate R_s is achievable if for any $\epsilon > 0$, there is a channel code $(2^{nR_s}, n)$ satisfying[3]

$$P_e \triangleq \Pr\left(\hat{W} \neq W\right) \leq \epsilon, \tag{3}$$

$$\frac{1}{n}H(W|\mathbf{Y}_e^n) \geq \frac{1}{n}H(W) - \epsilon. \tag{4}$$

The secrecy capacity of a channel, $C_s(P)$, is defined as the closure of all its achievable secrecy rates. For a channel with complex-valued coefficients, the maximum secure degrees of freedom (s.d.o.f.) is defined as

$$D_s \triangleq \lim_{P\to\infty} \frac{C_s(P)}{\log P}. \tag{5}$$

The cooperative jammer transmits the signal $\mathbf{X}_c^n \in \mathcal{X}_c^n$ in order to reduce the reception capability of the eavesdropper. However, this transmission affects the receiver as well, as interference. The jamming signal, $\mathbf{X}_c^n$, does not carry any information. Additionally, there is no shared secret between the transmitter and the cooperative jammer.

III. Main Result

We first state the s.d.o.f. results for $N_t = N_r = N$.

Theorem 1: The s.d.o.f. of the MIMO wire-tap channel with an N_c-antenna cooperative jammer, N antennas at each of the transmitter and receiver, and N_e antennas at the eavesdropper, is almost surely[4] *(a.s.)*

$$D_s = \begin{cases} [N+N_c-N_e]^+, & \text{if } 0 \leq N_c \leq N_e - \frac{\min\{N,N_e\}}{2} \\ N - \frac{\min\{N,N_e\}}{2}, & \text{if } N_e - \frac{\min\{N,N_e\}}{2} < N_c \leq \max\{N,N_e\} \\ \frac{N+N_c-N_e}{2}, & \text{if } \max\{N,N_e\} < N_c \leq N+N_e, \\ N, & \text{if } N_c > N+N_e \end{cases} \tag{6}$$

Proof: The proof for Theorem 1 is provided in Sections IV and V. ■

Next, in Theorem 2 below, we generalize the result in Theorem 1 to $N_t \neq N_r$.

Theorem 2: The s.d.o.f. of the MIMO wire-tap channel with an N_c-antenna cooperative jammer, N_t-antenna transmitter, N_r-antenna receiver, and N_e-antenna eavesdropper, is a.s.

$$D_s = \begin{cases} \min\left\{N_r, [N_c+N_t-N_e]^+\right\}, & \text{if } 0 \leq N_c \leq N_1 \\ \min\left\{N_t, N_r, \frac{N_r+[N_t-N_e]^+}{2}\right\}, & \text{if } N_1 < N_c \leq N_2 \\ \min\left\{N_t, N_r, \frac{N_c+N_t-N_e}{2}\right\}, & \text{if } N_2 < N_c \leq N_3, \\ \min\left\{N_t, N_r\right\}, & \text{if } N_c > N_3, \end{cases} \tag{7}$$

where,

$$N_1 = \min\left\{N_e, \left[\frac{N_r}{2} + \frac{N_e - N_t}{2 - 1_{N_e > N_t}}\right]^+\right\},$$

$$1_{N_e>N_t} = \begin{cases} 1, & \text{if } N_e > N_t \\ 0, & \text{if } N_e \leq N_t \end{cases}$$

$$N_2 = N_r + [N_e - N_t]^+,$$

$$N_3 = \max\left\{N_2, 2\min\left\{N_t, N_r\right\} + N_e - N_t\right\}.$$

[2] Throughout the paper, we omit the index n whenever possible.

[3] We consider weak secrecy throughout this paper.

[4] The subset of the channel gains for which the result does not hold has a Lebesgue measure zero.

Proof: The proof for Theorem 2 is provided in Section VI. ■

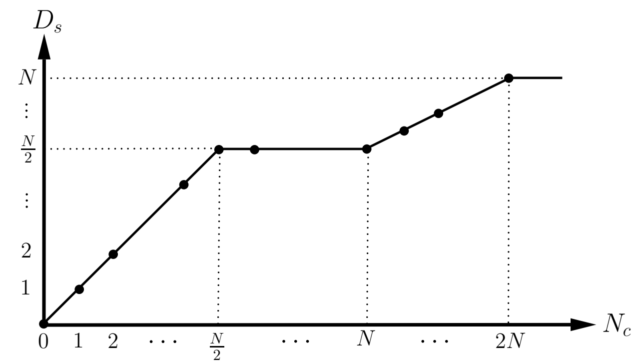
Remark 1: Theorem 2 provides a complete characterization for the s.d.o.f. of the channel. The s.d.o.f. at $N_c = N_3$ is equal to $\min\{N_t, N_r\}$, which is equal to the d.o.f of the $(N_t \times N_r)$ point-to-point MIMO Gaussian channel. Thus, increasing the number of antennas at the cooperative jammer, N_c, over N_3 cannot increase the s.d.o.f. over $\min\{N_t, N_r\}$.

Remark 2: For $N_t \geq N_r + N_e$, the s.d.o.f. of the channel is equal to N_r at $N_c = 0$, i.e., the maximum s.d.o.f. of the channel is achieved without the help of the cooperative jammer.

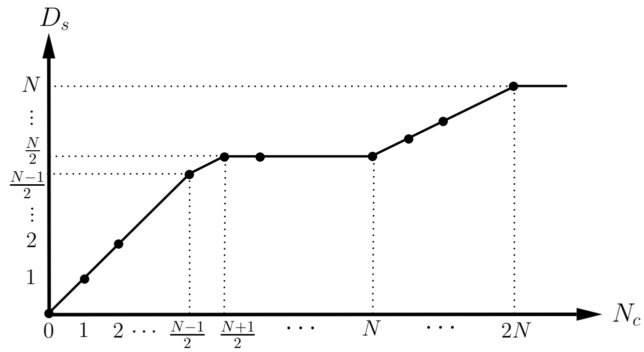
Remark 3: The converse proof for Theorem 2 involves combining two upper bounds for the s.d.o.f. derived for two different ranges of N_c. These two bounds are a straightforward generalization of those derived for the symmetric case in Theorem 1. However, combining them is more tedious since more cases of the number of antennas at the different terminals should be handled carefully. Achievability for Theorem 2 utilizes similar techniques to those used for Theorem 1 as well, where handling more cases is required. For clarity of exposition, we derive the s.d.o.f. for the symmetric case first in order to present the main ideas, and then utilize these ideas and generalize the result to the asymmetric case of Theorem 2.

For illustration purposes, the s.d.o.f. for $N_t = N_r = N_e = N$, and N_c varying from 0 to $2N$, is depicted in Fig. 2. The s.d.o.f. curves with N even and odd are shown in Fig. 2a and Fig. 2b, respectively.

We provide the discussion of the results of this work in Section VII.



(a) N even.



(b) N odd.

Fig. 2: Secure degrees of freedom for a MIMO wire-tap channel, with N antennas at each of its nodes, and a cooperative jammer with N_c antennas, where N_c varies from 0 to $2N$.

IV. Converse for $N_t = N_r = N$

In Section IV-A, we derive the upper bound for the s.d.o.f. for $0 \leq N_c \leq N_e$. In Section IV-B, we derive the upper bound for $\max\{N, N_e\} \leq N_c \leq N + N_e$. The two bounds are combined in Section IV-C to provide the desired upper bound in (6).

A. $0 \leq N_c \leq N_e$

Allow for full cooperation between the transmitter and the cooperative jammer. This cooperation cannot decrease the s.d.o.f. of the channel, and yields a MIMO wire-tap channel with $N+N_c$-antenna transmitter, N-antenna receiver, and N_e-antenna eavesdropper. It has been shown in [28] that, at high SNR, i.e., $P \to \infty$, the secrecy capacity of this channel, $C_s(P)$, takes the asymptotic form

$$C_s(P) = \log\det\left(\mathbf{I}_N + \frac{P}{p}\mathbf{H}\mathbf{G}^{\sharp}\mathbf{H}^H\right) + o(\log P), \qquad (8)$$

where $\lim_{P\to\infty} \frac{o(\log P)}{\log P} = 0$, $\mathbf{H} \in \mathbb{C}^{N\times(N+N_c)}$ and $\mathbf{G} \in \mathbb{C}^{N_e\times(N+N_c)}$ are the channel gains from the combined transmitter to the receiver and eavesdropper, and $\mathbf{G}^{\sharp}$ is the projection matrix onto the null space of $\mathbf{G}$, $\mathcal{N}(\mathbf{G})$. $p \triangleq \dim\left\{\mathcal{N}(\mathbf{H})^{\perp} \cap \mathcal{N}(\mathbf{G})\right\}$, where $\mathcal{N}(\mathbf{H})^{\perp}$ is the space orthogonal to the null space of $\mathbf{H}$. Due to the randomly generated channel gains, if a vector $\mathbf{x} \in \mathcal{N}(\mathbf{G})$, then $\mathbf{x} \in \mathcal{N}(\mathbf{H})^{\perp}$ a.s., for all $0 \leq N_c \leq N_e$. Thus, $p = \dim(\mathcal{N}(\mathbf{G})) = [N + N_c - N_e]^+$.

$\mathbf{H}\mathbf{G}^{\sharp}\mathbf{H}^H$ can be decomposed as

$$\mathbf{H}\mathbf{G}^{\sharp}\mathbf{H}^H = \mathbf{\Psi}\begin{bmatrix}\mathbf{0}_{(N-p)\times(N-p)} & \mathbf{0}_{(N-p)\times p} \\ \mathbf{0}_{p\times(N-p)} & \mathbf{\Omega}\end{bmatrix}\mathbf{\Psi}^H, \qquad (9)$$

where $\mathbf{\Psi} \in \mathbb{C}^{N\times N}$ is a unitary matrix and $\mathbf{\Omega} \in \mathbb{C}^{p\times p}$ is a non-singular matrix [28]. Let $\mathbf{\Psi} = [\mathbf{\Psi}_1\ \mathbf{\Psi}_2]$, where $\mathbf{\Psi}_1 \in \mathbb{C}^{N\times(N-p)}$ and $\mathbf{\Psi}_2 \in \mathbb{C}^{N\times p}$. Substituting (9) in (8) yields

$$\begin{aligned} C_s(P) &= \log\det\left(\mathbf{I}_N + \frac{P}{p}\mathbf{\Psi}_2\mathbf{\Omega}\mathbf{\Psi}_2^H\right) + o(\log P) && (10)\\ &= \log\det\left(\mathbf{I}_p + \frac{P}{p}\mathbf{\Omega}\mathbf{\Psi}_2^H\mathbf{\Psi}_2\right) + o(\log P) && (11)\\ &= \log P^p \det\left(\frac{1}{P}\mathbf{I}_p + \frac{1}{p}\mathbf{\Omega}\right) + o(\log P) && (12)\\ &= p\log P + o(\log P), && (13)\end{aligned}$$

where (11) follows from Sylvester's determinant identity and (12) follows from $\mathbf{\Psi}$ being unitary.

Thus, the s.d.o.f. of the original channel, for $0 \leq N_c \leq N_e$, is upper bounded as

$$\begin{aligned} D_s &\leq \lim_{P\to\infty}\frac{C_s(P)}{\log P} = \lim_{P\to\infty}\frac{p\log P + o(\log P)}{\log P} && (14)\\ &= [N + N_c - N_e]^+. && (15)\end{aligned}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2738011, IEEE Transactions on Information Theory

B. $\max\{N, N_e\} < N_c \leq N + N_e$

The upper bound we derive here is inspired by the converse of the single antenna Gaussian wire-tap channel with a single antenna cooperative jammer derived in [24], though as we will see shortly, the vector channel extension resulting from multiple antennas does require care. Let ϕ_i, for $i = 1, 2, \cdots, 10$, denote constants which do not depend on the power P.

The secrecy rate R_s can be upper bounded as follows

$$
\begin{aligned}
nR_s &= H(W) && (16)\\
&= H(W) - H(W|\mathbf{Y}_e^n) + H(W|\mathbf{Y}_e^n)\\
&\qquad - H(W|\mathbf{Y}_r^n) + H(W|\mathbf{Y}_r^n) && (17)\\
&\leq n\epsilon + H(W|\mathbf{Y}_e^n) - H(W|\mathbf{Y}_r^n, \mathbf{Y}_e^n) + n\delta && (18)\\
&= I(W; \mathbf{Y}_r^n|\mathbf{Y}_e^n) + n\phi_1 && (19)\\
&= h(\mathbf{Y}_r^n|\mathbf{Y}_e^n) - h(\mathbf{Y}_r^n|W, \mathbf{Y}_e^n) + n\phi_1 && (20)\\
&\leq h(\mathbf{Y}_r^n|\mathbf{Y}_e^n) - h(\mathbf{Y}_r^n|W, \mathbf{Y}_e^n, \mathbf{X}_t^n, \mathbf{X}_c^n) + n\phi_1 && (21)\\
&= h(\mathbf{Y}_r^n, \mathbf{Y}_e^n) - h(\mathbf{Y}_e^n) - h(\mathbf{Z}_r^n) + n\phi_1, && (22)
\end{aligned}
$$

where (18) follows since $H(W) - H(W|\mathbf{Y}_e^n) \leq n\epsilon$ by the secrecy constraint in (4), $H(W|\mathbf{Y}_r^n) \leq n\delta$ by Fano's inequality, and $H(W|\mathbf{Y}_r^n) \geq H(W|\mathbf{Y}_r^n, \mathbf{Y}_e^n)$ by the fact that conditioning does not increase entropy, (22) follows since $\mathbf{Z}_r^n$ is independent from $\{W, \mathbf{Y}_e^n, \mathbf{X}_t^n, \mathbf{X}_c^n\}$, and $\phi_1 = \epsilon + \delta$.

Let $\tilde{\mathbf{X}}_t = \mathbf{X}_t + \tilde{\mathbf{Z}}_t$ and $\tilde{\mathbf{X}}_c = \mathbf{X}_c + \tilde{\mathbf{Z}}_c$, where $\tilde{\mathbf{Z}}_t \sim \mathcal{CN}(\mathbf{0}, \mathbf{K}_t)$ and $\tilde{\mathbf{Z}}_c \sim \mathcal{CN}(\mathbf{0}, \mathbf{K}_c)$. The covariance matrices, $\mathbf{K}_t$ and $\mathbf{K}_c$, are chosen as $\mathbf{K}_t = \rho^2\mathbf{I}_N$ and $\mathbf{K}_c = \rho^2\mathbf{I}_{N_c}$, where $0 < \rho \leq 1/\max\left\{||\mathbf{H}_c^H||, \sqrt{||\mathbf{G}_t^H||^2 + ||\mathbf{G}_c^H||^2}\right\}$. Note that $\tilde{\mathbf{X}}_t$ and $\tilde{\mathbf{X}}_c$ are noisy versions of the transmitted signals $\mathbf{X}_t$ and $\mathbf{X}_c$, respectively. $\tilde{\mathbf{Z}}_t$ is independent from $\tilde{\mathbf{Z}}_c$ and both are independent from $\{\mathbf{X}_t, \mathbf{X}_c, \mathbf{Z}_r, \mathbf{Z}_e\}$. $\tilde{\mathbf{Z}}_t^n$ and $\tilde{\mathbf{Z}}_c^n$ are i.i.d. sequences of the random vectors $\tilde{\mathbf{Z}}_t$ and $\tilde{\mathbf{Z}}_c$. In addition, let $\tilde{\mathbf{Z}}_1 = -\mathbf{H}_t\tilde{\mathbf{Z}}_t - \mathbf{H}_c\tilde{\mathbf{Z}}_c + \mathbf{Z}_r$ and $\tilde{\mathbf{Z}}_2 = -\mathbf{G}_t\tilde{\mathbf{Z}}_t - \mathbf{G}_c\tilde{\mathbf{Z}}_c + \mathbf{Z}_e$. Note that $\tilde{\mathbf{Z}}_1 \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{\Sigma}_{\tilde{\mathbf{Z}}_1}\right)$ and $\tilde{\mathbf{Z}}_2 \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{\Sigma}_{\tilde{\mathbf{Z}}_2}\right)$, where $\mathbf{\Sigma}_{\tilde{\mathbf{Z}}_1} = \mathbf{H}_t\mathbf{K}_t\mathbf{H}_t^H + \mathbf{H}_c\mathbf{K}_c\mathbf{H}_c^H + \mathbf{I}_N$ and $\mathbf{\Sigma}_{\tilde{\mathbf{Z}}_2} = \mathbf{G}_t\mathbf{K}_t\mathbf{G}_t^H + \mathbf{G}_c\mathbf{K}_c\mathbf{G}_c^H + \mathbf{I}_{N_e}$. $\tilde{\mathbf{Z}}_1^n$ and $\tilde{\mathbf{Z}}_2^n$ are i.i.d. sequences of $\tilde{\mathbf{Z}}_1$ and $\tilde{\mathbf{Z}}_2$, since each of $\mathbf{Z}_r^n, \mathbf{Z}_e^n, \tilde{\mathbf{Z}}_t^n, \tilde{\mathbf{Z}}_c^n$ is i.i.d. across time. The choice of $\mathbf{K}_t$ and $\mathbf{K}_c$ above guarantees the finiteness of $h(\tilde{\mathbf{Z}}_t), h(\tilde{\mathbf{Z}}_c), h(\tilde{\mathbf{Z}}_1)$, and $h(\tilde{\mathbf{Z}}_2)$ as shown in Appendix A. Starting from (22), we have

$$
\begin{aligned}
nR_s &\leq h(\mathbf{Y}_r^n, \mathbf{Y}_e^n) - h(\mathbf{Y}_e^n) + n\phi_2 && (23)\\
&= h(\mathbf{Y}_r^n, \mathbf{Y}_e^n, \tilde{\mathbf{X}}_t^n, \tilde{\mathbf{X}}_c^n) - h(\tilde{\mathbf{X}}_t^n, \tilde{\mathbf{X}}_c^n|\mathbf{Y}_r^n, \mathbf{Y}_e^n)\\
&\qquad - h(\mathbf{Y}_e^n) + n\phi_2 && (24)\\
&\leq h(\tilde{\mathbf{X}}_t^n, \tilde{\mathbf{X}}_c^n) + h(\mathbf{Y}_r^n, \mathbf{Y}_e^n|\tilde{\mathbf{X}}_t^n, \tilde{\mathbf{X}}_c^n)\\
&\qquad - h(\tilde{\mathbf{X}}_t^n, \tilde{\mathbf{X}}_c^n|\mathbf{Y}_r^n, \mathbf{Y}_e^n, \mathbf{X}_t^n, \mathbf{X}_c^n) - h(\mathbf{Y}_e^n) + n\phi_2 && (25)\\
&\leq h(\tilde{\mathbf{X}}_t^n) + h(\tilde{\mathbf{X}}_c^n) + h(\mathbf{Y}_r^n|\tilde{\mathbf{X}}_t^n, \tilde{\mathbf{X}}_c^n)\\
&\qquad + h(\mathbf{Y}_e^n|\tilde{\mathbf{X}}_t^n, \tilde{\mathbf{X}}_c^n) - h(\tilde{\mathbf{Z}}_t^n, \tilde{\mathbf{Z}}_c^n) - h(\mathbf{Y}_e^n) + n\phi_2 && (26)\\
&= h(\tilde{\mathbf{X}}_t^n) + h(\tilde{\mathbf{X}}_c^n) + h(\tilde{\mathbf{Z}}_1^n|\tilde{\mathbf{X}}_t^n, \tilde{\mathbf{X}}_c^n) + h(\tilde{\mathbf{Z}}_2^n|\tilde{\mathbf{X}}_t^n, \tilde{\mathbf{X}}_c^n)\\
&\qquad - h(\mathbf{Y}_e^n) + n\phi_3 && (27)\\
&\leq h(\tilde{\mathbf{X}}_t^n) + h(\tilde{\mathbf{X}}_c^n) + h(\tilde{\mathbf{Z}}_1^n) + h(\tilde{\mathbf{Z}}_2^n) - h(\mathbf{Y}_e^n) + n\phi_3 && (28)\\
&= h(\tilde{\mathbf{X}}_t^n) + h(\tilde{\mathbf{X}}_c^n) - h(\mathbf{Y}_e^n) + n\phi_4, && (29)
\end{aligned}
$$

where (26) follows since $\tilde{\mathbf{Z}}_t^n$ and $\tilde{\mathbf{Z}}_c^n$ are independent from $\{\mathbf{X}_t^n, \mathbf{X}_c^n, \mathbf{Y}_r^n, \mathbf{Y}_e^n\}$, $\phi_2 = \phi_1 - h(\mathbf{Z}_r)$, $\phi_3 = \phi_2 - h(\tilde{\mathbf{Z}}_t) - h(\tilde{\mathbf{Z}}_c)$, and $\phi_4 = \phi_3 + h(\tilde{\mathbf{Z}}_1) + h(\tilde{\mathbf{Z}}_2)$.

We have utilized the noisy versions $\tilde{\mathbf{X}}_t = \mathbf{X}_t + \tilde{\mathbf{Z}}_t$ and $\tilde{\mathbf{X}}_c = \mathbf{X}_c + \tilde{\mathbf{Z}}_c$ instead of $\mathbf{X}_t$, $\mathbf{X}_c$ so that (24)-(29) hold whether $\mathbf{X}_t, \mathbf{X}_c$ are continuous or discrete random vectors. This requires continuing the analysis with stochastically equivalent versions of $\mathbf{Y}_r, \mathbf{Y}_e$ in which they are expressed as functions of $\tilde{\mathbf{X}}_t$ and/or $\tilde{\mathbf{X}}_c$. To do so, we divide the Gaussian noise $\mathbf{Z}_r, \mathbf{Z}_e$ into sums of other independent Gaussian noise variables. The infinite divisibility of the Gaussian distribution ensures such division of $\mathbf{Z}_r, \mathbf{Z}_e$. We now consider the following two cases.

Case 1: $N_e \leq N$

We first lower bound $h(\mathbf{Y}_e^n)$ in (29) as follows. Using the infinite divisibility of Gaussian distribution, we can express a stochastically equivalent form of $\mathbf{Z}_e$, denoted by $\mathbf{Z}_e'$, as

$$\mathbf{Z}_e' = \mathbf{G}_t\tilde{\mathbf{Z}}_t + \tilde{\mathbf{Z}}_e. \qquad (30)$$

where[5] $\tilde{\mathbf{Z}}_e \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_e} - \mathbf{G}_t\mathbf{K}_t\mathbf{G}_t^H)$ is independent from $\{\tilde{\mathbf{Z}}_t, \tilde{\mathbf{Z}}_c, \mathbf{X}_t, \mathbf{X}_c, \mathbf{Z}_r\}$. $\tilde{\mathbf{Z}}_e^n$ is an i.i.d. sequence of the random vectors $\tilde{\mathbf{Z}}_e$. Using (30), a stochastically equivalent form of $\mathbf{Y}_e^n$ is

$${\mathbf{Y}'}_e^n = \mathbf{G}_t\tilde{\mathbf{X}}_t^n + \mathbf{G}_c\mathbf{X}_c^n + \tilde{\mathbf{Z}}_e^n. \qquad (31)$$

Let $\mathbf{X}_t = [X_{t,1} \cdots X_{t,N}]^T$, $\tilde{\mathbf{Z}}_t = [\tilde{Z}_{t,1} \cdots \tilde{Z}_{t,N}]^T$, and $\tilde{\mathbf{X}}_t = [\tilde{\mathbf{X}}_{t_1}^T \ \ \tilde{\mathbf{X}}_{t_2}^T]^T$, where $\tilde{\mathbf{X}}_{t_1} = [\tilde{X}_{t,1} \cdots \tilde{X}_{t,N_e}]^T$, $\tilde{\mathbf{X}}_{t_2} = [\tilde{X}_{t,N_e+1} \cdots \tilde{X}_{t,N}]^T$, and $\tilde{X}_{t,k} = X_{t,k} + \tilde{Z}_{t,k}$, $k = 1, 2, \cdots, N$. In addition, let $\mathbf{G}_t = [\mathbf{G}_{t_1} \ \mathbf{G}_{t_2}]$, where $\mathbf{G}_{t_1} \in \mathbb{C}^{N_e \times N_e}$ and[6] $\mathbf{G}_{t_2} \in \mathbb{C}^{N_e \times (N - N_e)}$. Using (31), we have

$$
\begin{aligned}
h(\mathbf{Y}_e^n) &= h({\mathbf{Y}'}_e^n) = h(\mathbf{G}_t\tilde{\mathbf{X}}_t^n + \mathbf{G}_c\mathbf{X}_c^n + \tilde{\mathbf{Z}}_e^n) && (32)\\
&\geq h(\mathbf{G}_t\tilde{\mathbf{X}}_t^n) = h(\mathbf{G}_{t_1}\tilde{\mathbf{X}}_{t_1}^n + \mathbf{G}_{t_2}\tilde{\mathbf{X}}_{t_2}^n) && (33)\\
&\geq h(\mathbf{G}_{t_1}\tilde{\mathbf{X}}_{t_1}^n + \mathbf{G}_{t_2}\tilde{\mathbf{X}}_{t_2}^n|\tilde{\mathbf{X}}_{t_2}^n) = h(\mathbf{G}_{t_1}\tilde{\mathbf{X}}_{t_1}^n|\tilde{\mathbf{X}}_{t_2}^n) && (34)\\
&= h(\tilde{\mathbf{X}}_{t_1}^n|\tilde{\mathbf{X}}_{t_2}^n) + n\log|\det(\mathbf{G}_{t_1})|. && (35)
\end{aligned}
$$

where the inequality in (33) follows since $\{\mathbf{G}_t\tilde{\mathbf{X}}_t^n\}$ and $\{\mathbf{G}_c\mathbf{X}_c^n + \tilde{\mathbf{Z}}_e^n\}$ are independent, as for two independent random vectors $\mathbf{X}$ and $\mathbf{Y}$, we have $h(\mathbf{X} + \mathbf{Y}) \geq h(\mathbf{X})$.

Substituting (35) in (29) results in

$$
\begin{aligned}
nR_s &\leq h(\tilde{\mathbf{X}}_{t_1}^n, \tilde{\mathbf{X}}_{t_2}^n) + h(\tilde{\mathbf{X}}_c^n) - h(\tilde{\mathbf{X}}_{t_1}^n|\tilde{\mathbf{X}}_{t_2}^n)\\
&\qquad - n\log|\det(\mathbf{G}_{t_1})| + n\phi_4 && (36)\\
&= h(\tilde{\mathbf{X}}_{t_2}^n) + h(\tilde{\mathbf{X}}_c^n) + n\phi_5, && (37)
\end{aligned}
$$

where $\phi_5 = \phi_4 - \log|\det(\mathbf{G}_{t_1})|$.

We now exploit the reliability constraint in (3) to derive another upper bound for R_s, which we combine with the bound in (37) in order to obtain the desired bound for the s.d.o.f. when $N_e \leq N$ and $N < N_c \leq N + N_e$. The reliability constraint in (3) can be achieved only if [41]

$$
\begin{aligned}
nR_s &\leq I(\mathbf{X}_t^n; \mathbf{Y}_r^n) = h(\mathbf{Y}_r^n) - h(\mathbf{Y}_r^n|\mathbf{X}_t^n) && (38)\\
&= h(\mathbf{Y}_r^n) - h(\mathbf{H}_c\mathbf{X}_c^n + \mathbf{Z}_r^n). && (39)
\end{aligned}
$$

[5] The choice of $\mathbf{K}_t$ guarantees that $\mathbf{I}_{N_e} - \mathbf{G}_t\mathbf{K}_t\mathbf{G}_t^H$ is a valid covariance matrix.

[6] Note that when $N_e = N$, the vector $\tilde{\mathbf{X}}_{t_2}$ and the matrix $\mathbf{G}_{t_2}$ vanish and the analysis below holds in the same manner, by discarding $\tilde{\mathbf{X}}_{t_2}^n$ and $\mathbf{G}_{t_2}$.

Similar to (30), a stochastically equivalent form of $\mathbf{Z}_r$ is given by

$$\mathbf{Z}'_r = \mathbf{H}_c \tilde{\mathbf{Z}}_c + \tilde{\mathbf{Z}}_r, \tag{40}$$

where[7] $\tilde{\mathbf{Z}}_r \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N - \mathbf{H}_c\mathbf{K}_c\mathbf{H}_c^H)$ is independent from $\{\tilde{\mathbf{Z}}_t, \tilde{\mathbf{Z}}_c, \mathbf{X}_t, \mathbf{X}_c, \mathbf{Z}_e\}$. $\tilde{\mathbf{Z}}_r^n$ is an i.i.d. sequence of the random vectors $\tilde{\mathbf{Z}}_r$.

Let $\mathbf{X}_c = [X_{c,1} \cdots X_{c,N_c}]^T$, $\tilde{\mathbf{Z}}_c = [\tilde{Z}_{c,1} \cdots \tilde{Z}_{c,N_c}]^T$, and $\tilde{\mathbf{X}}_c = [\tilde{\mathbf{X}}_{c_1}^T \ \tilde{\mathbf{X}}_{c_2}^T]^T$, where $\tilde{\mathbf{X}}_{c_1} = [\tilde{X}_{c,1} \ \ldots \ \tilde{X}_{c,N}]^T$, $\tilde{\mathbf{X}}_{c_2} = [\tilde{X}_{c,N+1} \ \cdots \ \tilde{X}_{c,N_c}]^T$, and $\tilde{X}_{c,k} = X_{c,k} + \tilde{Z}_{c,k}$, $k = 1, 2, \cdots, N_c$. In addition, let $\mathbf{H}_c = [\mathbf{H}_{c_1} \ \mathbf{H}_{c_2}]$, where $\mathbf{H}_{c_1} \in \mathbb{C}^{N\times N}$ and $\mathbf{H}_{c_2} \in \mathbb{C}^{N\times(N_c-N)}$. Using (40), we have

$$\begin{aligned}
h(\mathbf{H}_c\mathbf{X}_c^n + \mathbf{Z}_r^n) &= h(\mathbf{H}_c\mathbf{X}_c^n + {\mathbf{Z}'_r}^n) = h(\mathbf{H}_c\tilde{\mathbf{X}}_c^n + \tilde{\mathbf{Z}}_r^n) && (41)\\
&\geq h(\mathbf{H}_c\tilde{\mathbf{X}}_c^n) = h(\mathbf{H}_{c_1}\tilde{\mathbf{X}}_{c_1}^n + \mathbf{H}_{c_2}\tilde{\mathbf{X}}_{c_2}^n) && (42)\\
&\geq h(\mathbf{H}_{c_1}\tilde{\mathbf{X}}_{c_1}^n | \tilde{\mathbf{X}}_{c_2}^n) && (43)\\
&= h(\tilde{\mathbf{X}}_{c_1}^n | \tilde{\mathbf{X}}_{c_2}^n) + n\log|\det(\mathbf{H}_{c_1})|. && (44)
\end{aligned}$$

Substituting (44) in (39) yields

$$nR_s \leq h(\mathbf{Y}_r^n) - h(\tilde{\mathbf{X}}_{c_1}^n | \tilde{\mathbf{X}}_{c_2}^n) - n\log|\det(\mathbf{H}_{c_1})|. \tag{45}$$

Let $\mathbf{Y}_r = [Y_{r,1} \ \cdots \ Y_{r,N}]^T$. Summing (37) and (45) results in

$$\begin{aligned}
nR_s &\leq \frac{1}{2}\Big\{h(\mathbf{Y}_r^n) + h(\tilde{\mathbf{X}}_{t_2}^n) + h(\tilde{\mathbf{X}}_{c_2}^n)\Big\} + n\phi_6 && (46)\\
&\leq \frac{1}{2}\sum_{i=1}^{n}\Bigg\{\sum_{k=1}^{N} h(Y_{r,k}(i)) + \sum_{k=N_e+1}^{N} h(\tilde{X}_{t,k}(i))\\
&\qquad + \sum_{k=N+1}^{N_c} h(\tilde{X}_{c,k}(i))\Bigg\} + n\phi_6, && (47)
\end{aligned}$$

where $\phi_6 = \frac{1}{2}(\phi_5 - \log|\det(\mathbf{H}_{c_1})|)$.

In Appendix B, we show, for $i = 1, \cdots, n$, $k = 1, \cdots, N$, and $m = 1, \cdots, N_c$, that

$$h(Y_{r,k}(i)) \leq \log 2\pi e + \log(1 + h^2 P) \tag{48}$$

$$h(\tilde{X}_{t,k}(i)), h(\tilde{X}_{c,m}(i)) \leq \log 2\pi e + \log(\rho^2 + P), \tag{49}$$

where $h^2 = \max_k \left(||\mathbf{h}_{t,k}^r||^2 + ||\mathbf{h}_{c,k}^r||^2\right)$; $\mathbf{h}_{t,k}^r$ and $\mathbf{h}_{c,k}^r$ denote the transpose of the kth row vectors of $\mathbf{H}_t$ and $\mathbf{H}_c$, respectively. Using (47), (48), and (49), we have

$$R_s \leq \frac{N}{2}\log(1 + h^2 P) + \frac{N_c - N_e}{2}\log(\rho^2 + P) + \phi_7, \tag{50}$$

where $\phi_7 = \phi_6 + \frac{N+N_c-N_e}{2}\log 2\pi e$. Using (5), we get

$$\begin{aligned}
D_s &\leq \lim_{P\to\infty} \frac{\frac{N}{2}\log(1+h^2P) + \frac{N_c-N_e}{2}\log(\rho^2+P) + \phi_7}{\log P} && (51)\\
&= \frac{N + N_c - N_e}{2}. && (52)
\end{aligned}$$

Thus, the s.d.o.f. for $N_e \leq N$ and $N < N_c \leq N + N_e$, is upper bounded by $\frac{N+N_c-N_e}{2}$.

Case 2: $N_e > N$

Another stochastically equivalent form of $\mathbf{Z}_e$ is

$$\mathbf{Z}''_e = \mathbf{G}_t\tilde{\mathbf{Z}}_t + \mathbf{G}_c\tilde{\mathbf{Z}}_c + \tilde{\mathbf{Z}}'_e. \tag{53}$$

where[8] $\tilde{\mathbf{Z}}'_e \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_e} - \mathbf{G}_t\mathbf{K}_t\mathbf{G}_t^H - \mathbf{G}_c\mathbf{K}_c\mathbf{G}_c^H)$ is independent from $\{\tilde{\mathbf{Z}}_t, \tilde{\mathbf{Z}}_c, \mathbf{X}_t, \mathbf{X}_c, \mathbf{Z}_r\}$. $\tilde{\mathbf{Z}}_e'^n$ is an i.i.d. sequence of the random vectors $\tilde{\mathbf{Z}}'_e$. Using (53), another stochastically equivalent form of $\mathbf{Y}_e^n$ is given by

$$\mathbf{Y}_e''^n = \mathbf{G}_t\tilde{\mathbf{X}}_t + \mathbf{G}_c\tilde{\mathbf{X}}_c^n + \tilde{\mathbf{Z}}_e'^n. \tag{54}$$

Let us rewrite $\tilde{\mathbf{X}}_c$ and $\mathbf{H}_c$ as follows. $\tilde{\mathbf{X}}_c = [\tilde{\mathbf{X}}_{c_1}'^T \ \tilde{\mathbf{X}}_{c_2}'^T]^T$, where $\tilde{\mathbf{X}}'_{c_1} = [\tilde{X}_{c,1} \cdots \tilde{X}_{c,N_e-N}]^T$, $\tilde{\mathbf{X}}'_{c_2} = [\tilde{\mathbf{X}}_{c_{21}}'^T \ \tilde{\mathbf{X}}_{c_{22}}'^T]^T$, $\tilde{\mathbf{X}}'_{c_{21}} = [\tilde{X}_{c,N_e-N+1} \cdots \tilde{X}_{c,N_e}]^T$, and $\tilde{\mathbf{X}}'_{c_{22}} = [\tilde{X}_{c,N_e+1} \cdots \tilde{X}_{c,N_c}]^T$. $\mathbf{H}_c = [\mathbf{H}'_{c_1} \ \mathbf{H}'_{c_2}]$, where $\mathbf{H}'_{c_1} \in \mathbb{C}^{N\times(N_e-N)}$, $\mathbf{H}'_{c_2} = [\mathbf{H}'_{c_{21}} \ \mathbf{H}'_{c_{22}}]$, $\mathbf{H}'_{c_{21}} \in \mathbb{C}^{N\times N}$, and $\mathbf{H}'_{c_{22}} \in \mathbb{C}^{N\times(N_c-N_e)}$. Let $\mathbf{G}_c = [\mathbf{G}_{c_1} \ \mathbf{G}_{c_2}]$, where $\mathbf{G}_{c_1} \in \mathbb{C}^{N_e\times(N_e-N)}$ and $\mathbf{G}_{c_2} \in \mathbb{C}^{N_e\times(N+N_c-N_e)}$. Using (54), we have

$$\begin{aligned}
h(\mathbf{Y}_e^n) &= h({\mathbf{Y}''_e}^n) = h\Bigg([\mathbf{G}_t \ \mathbf{G}_{c_1}]\begin{bmatrix}\tilde{\mathbf{X}}_t^n\\ \tilde{\mathbf{X}}_{c_1}'^n\end{bmatrix} + \mathbf{G}_{c_2}\tilde{\mathbf{X}}_{c_2}'^n + \tilde{\mathbf{Z}}_e'^n\Bigg) && (55)\\
&\geq h(\tilde{\mathbf{X}}_t^n, \tilde{\mathbf{X}}_{c_1}'^n | \tilde{\mathbf{X}}_{c_2}'^n) + n\log|\det[\mathbf{G}_t \ \mathbf{G}_{c_1}]| && (56)\\
&= h(\tilde{\mathbf{X}}_t^n) + h(\tilde{\mathbf{X}}_{c_1}'^n | \tilde{\mathbf{X}}_{c_2}'^n) + n\log|\det[\mathbf{G}_t \ \mathbf{G}_{c_1}]|, && (57)
\end{aligned}$$

where (57) follows since $\tilde{\mathbf{X}}_t^n$ and $\tilde{\mathbf{X}}_c^n$ are independent. Substituting (57) in (29) gives

$$nR_s \leq h(\tilde{\mathbf{X}}_{c_2}'^n) + n\phi_8, \tag{58}$$

where $\phi_8 = \phi_4 - \log|\det[\mathbf{G}_t \ \mathbf{G}_{c_1}]|$.

In order to obtain another upper bound for R_s, which we combine with (58) to obtain the desired bound for $N_e > N$ and $N_e < N_c \leq N + N_e$, we proceed as follows. Consider a modified channel where the first $N_e - N$ antennas at the cooperative jammer are removed, i.e., the cooperative jammer uses only the last $N + N_c - N_e$ out of its N_c antennas. The transmitted signals in the modified channel are $\mathbf{X}_t^n$ and $\mathbf{X}_{c_2}'^n$, and hence, the legitimate receiver receives

$$\bar{\mathbf{Y}}_r^n = \mathbf{H}_t\mathbf{X}_t^n + \mathbf{H}'_{c_2}\mathbf{X}_{c_2}'^n + \mathbf{Z}_r^n. \tag{59}$$

Let R and $\bar{R}$ denote reliable communication rates, i.e., the achievable rates without the secrecy constraint, for the original and the modified channels, respectively. Since the cooperative jamming signal is additive interference for the legitimate receiver, the reliable communication rate of the modified channel, $\bar{R}$, is an upper bound for that of the original channel, R. Since R_s satisfies the reliability and secrecy constraints in (3) and (4), we have that

$$\begin{aligned}
nR_s &\leq nR \leq n\bar{R}\\
&\leq I(\mathbf{X}_t^n; \bar{\mathbf{Y}}_r^n) = h(\bar{\mathbf{Y}}_r^n) - h(\mathbf{H}'_{c_2}\mathbf{X}_{c_2}'^n + \mathbf{Z}_r^n). && (60)
\end{aligned}$$

Let $\tilde{\mathbf{Z}}_{c_2} = [\tilde{Z}_{c,N_e-N+1} \cdots \tilde{Z}_{c,N_c}]^T \sim \mathcal{CN}(\mathbf{0}, \mathbf{K}'_c)$, where $\mathbf{K}'_c = \rho^2\mathbf{I}_{N+N_c-N_e}$. Another stochastically equivalent form

[7]The choice of $\mathbf{K}_c$ guarantees that $\mathbf{I}_N - \mathbf{H}_c\mathbf{K}_c\mathbf{H}_c^H$ is a valid covariance matrix.

[8]The choice of $\mathbf{K}_t$ and $\mathbf{K}_c$ guarantees that $\mathbf{I}_{N_e} - \mathbf{G}_t\mathbf{K}_t\mathbf{G}_t^H - \mathbf{G}_c\mathbf{K}_c\mathbf{G}_c^H$ is a valid covariance matrix.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2738011, IEEE Transactions on Information Theory

of $\mathbf{Z}_r$ is $\mathbf{Z}''_r = \mathbf{H}'_{c_2}\tilde{\mathbf{Z}}_{c_2} + \tilde{\mathbf{Z}}'_r$, where[9] $\tilde{\mathbf{Z}}'_r \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N - \mathbf{H}'_{c_2}\mathbf{K}'_c\mathbf{H}'^H_{c_2})$ is independent from $\{\tilde{\mathbf{Z}}_t, \tilde{\mathbf{Z}}_c, \mathbf{X}_t, \mathbf{X}_c, \mathbf{Z}_e\}$, and $\tilde{\mathbf{Z}}'^n_r$ is an i.i.d. sequence of $\tilde{\mathbf{Z}}'_r$. Thus, using (60), we have

$$\begin{aligned} nR_s &\le h(\bar{\mathbf{Y}}^n_r) - h(\mathbf{H}'_{c_2}\tilde{\mathbf{X}}'^n_{c_2} + \tilde{\mathbf{Z}}'^n_r) && (61)\\ &\le h(\bar{\mathbf{Y}}^n_r) - h(\mathbf{H}'_{c_2}\tilde{\mathbf{X}}'^n_{c_2}) && (62)\\ &\le h(\bar{\mathbf{Y}}^n_r) - h(\tilde{\mathbf{X}}'^n_{c_{21}}|\tilde{\mathbf{X}}'^n_{c_{22}}) - n\log|\det(\mathbf{H}'_{c_{21}})|. && (63) \end{aligned}$$

Let $\bar{\mathbf{Y}}_r = [\bar{Y}_{r,1} \cdots \bar{Y}_{r,N}]^T$. Summing (58) and (63) yields

$$\begin{aligned} nR_s &\le \frac{1}{2}\left\{h(\bar{\mathbf{Y}}^n_r) + h(\tilde{\mathbf{X}}'^n_{c_{22}})\right\} + n\phi_9 && (64)\\ &\le \frac{1}{2}\sum_{i=1}^{n}\left\{\sum_{k=1}^{N} h(\bar{Y}_{r,k}(i)) + \sum_{k=N_e+1}^{N_c} h(\tilde{X}_{c,k}(i))\right\} + n\phi_9, && (65) \end{aligned}$$

where $\phi_9 = \frac{1}{2}\{\phi_8 - \log|\det(\mathbf{H}'_{c_{21}})|\}$. In Appendix B, we also show that

$$h(\bar{Y}_{r,k}(i)) \le \log 2\pi e + \log(1 + \bar{h}^2 P), \qquad (66)$$

where $\bar{h}^2 = \max_k \left(||\mathbf{h}^r_{t,k}||^2 + ||\mathbf{h}'^r_{c,k}||^2\right)$; $\mathbf{h}'^r_{c,k}$ denotes the transpose of the kth row vector of $\mathbf{H}'_{c_2}$.

Similar to case 1, using (65), (66), and (49), the secrecy rate is bounded as

$$R_s \le \frac{N}{2}\log(1 + \bar{h}^2 P) + \frac{N_c - N_e}{2}\log(\rho^2 + P) + n\phi_{10}, \qquad (67)$$

where $\phi_{10} = \phi_9 + \frac{N+N_c-N_e}{2}\log 2\pi e$. Thus, the s.d.o.f., for $N_e > N$ and $N_e < N_c \le N + N_e$, is upper bounded as

$$D_s \le \frac{N + N_c - N_e}{2}. \qquad (68)$$

C. Obtaining the Upper Bound

For $N_e \le N$, the upper bound for the s.d.o.f. derived in Section IV-A is equal to $N + N_c - N_e$ for all $0 \le N_c \le N_e$. In addition, the upper bound derived in Section IV-B, at $N_c = N$, is equal to $N - \frac{N_e}{2}$, c.f. equations (15) and (52). As the former upper bound is greater than the latter for all $\frac{N_e}{2} < N_c \le N$, the s.d.o.f. is upper bounded by $N - \frac{N_e}{2}$ for all $\frac{N_e}{2} < N_c \le N$. Combining these statements, we have the following upper bound for the s.d.o.f. for $N_e \le N$:

$$D_s \le \begin{cases} N + N_c - N_e, & \text{if } 0 \le N_c \le \frac{N_e}{2}\\ N - \frac{N_e}{2}, & \text{if } \frac{N_e}{2} < N_c \le N\\ \frac{N+N_c-N_e}{2}, & \text{if } N < N_c \le N + N_e,\\ N, & \text{if } N_c > N + N_e. \end{cases} \qquad (69)$$

Similarly, when $N_e > N$ and for all $N_e - \frac{N}{2} < N_c \le N_e$, the upper bound derived for $0 \le N_c \le N_e$ in Section IV-A is greater than the upper bound derived in Section IV-B at $N_c = N_e$. Thus, the s.d.o.f. for $N_e - \frac{N}{2} < N_c \le N_e$ is upper bounded by $\frac{N}{2}$. In addition, the upper bound in (15) is equal to zero for all $0 \le N_c \le N_e - N$. Thus, the s.d.o.f. for $N_e > N$ is upper bounded as:

$$D_s \le \begin{cases} 0, & \text{if } 0 \le N_c \le N_e - N\\ N + N_c - N_e, & \text{if } N_e - N < N_c \le N_e - \frac{N}{2}\\ \frac{N}{2}, & \text{if } N_e - \frac{N}{2} < N_c \le N_e\\ \frac{N+N_c-N_e}{2}, & \text{if } N_e < N_c \le N + N_e,\\ N, & \text{if } N_c > N + N_e. \end{cases} \qquad (70)$$

By combining the bounds for $N_e \le N$ in (69) and for $N_e > N$ in (70), we obtain the upper bound for the s.d.o.f. in (6). In the next section, we will show the achievability of (6).

V. Achievablility for $N_t = N_r = N$

In this section, we provide the achievability proof for Theorem 1 by showing the achievability of (69) when $N_e \le N$, and the achievability of (70) when $N_e > N$. For both $N_e \le N$ and $N_e > N$, we divide the range of the number of antennas at the cooperative jammer, N_c, into five ranges and propose an achievable scheme for each range. For all the achievable schemes in this section, we have the n-letter signals, $\mathbf{X}^n_t$ and $\mathbf{X}^n_c$, as i.i.d. sequences. Since $\mathbf{X}^n_c$ is independent from $\mathbf{X}^n_t$, and each of them is i.i.d. across time, we have in effect a memoryless wire-tap channel and the secrecy rate

$$R_s = [I(\mathbf{X}_t; \mathbf{Y}_r) - I(\mathbf{X}_t; \mathbf{Y}_e)]^+, \qquad (71)$$

is achievable by *stochastic encoding* at the transmitter [4].

The transmitted signals at the transmitter and the cooperative jammer, for each of the following schemes, are

$$\mathbf{X}_t = \mathbf{P}_t\mathbf{U}_t, \quad \mathbf{X}_c = \mathbf{P}_c\mathbf{V}_c, \qquad (72)$$

where $\mathbf{U}_t = [U_1 \cdots U_d]^T$ and $\mathbf{V}_c = [V_1 \cdots V_l]^T$ are the information and cooperative jamming streams, respectively. $\mathbf{P}_t = [\mathbf{p}_{t,1} \cdots \mathbf{p}_{t,d}] \in \mathbb{C}^{N\times d}$ and $\mathbf{P}_c = [\mathbf{p}_{c,1} \cdots \mathbf{p}_{c,l}] \in \mathbb{C}^{N_c\times l}$ are the precoding matrices at the transmitter and the cooperative jammer.

Signaling, precoding, and decoding techniques utilized in this proof vary according to the relative number of antennas at the different terminals and whether the s.d.o.f. of the channel is integer valued or not an integer. In particular, we show that Gaussian signaling both for transmission and cooperative jamming is sufficient to achieve the integer valued s.d.o.f., while achieving non-integer s.d.o.f. requires structured signaling and cooperative jamming along with a combination of linear receiver processing, and the complex field equivalent of real interference alignment [37], [38]. Additionally, the linear precoding at the transmitter and the cooperative jammer depends on whether N_e is equal to, smaller than, or larger than N, and whether the number of antennas at the cooperative jammer, N_c, results in a s.d.o.f. for the channel that is before, after, or at the flat s.d.o.f. range in the s.d.o.f. plot versus N_c. This leads to an achievability proof that involves 10 distinct achievable schemes, which differ from each other in the type of signals used (Gaussian or structured), and/or precoding at the transmitter and cooperative jammer, and/or decoding at the legitimate receiver.

[9]The choice of $\mathbf{K}_c$ guarantees that $\mathbf{I}_N - \mathbf{H}'_{c_2}\mathbf{K}'_c\mathbf{H}'^H_{c_2}$ is a valid covariance matrix.

Remark 4: Note that integer-valued s.d.o.f. can also be achieved using structured signals. However, Gaussian signaling often outperforms structured signaling for finite SNR; see for example [42, Fig. 2]. Although our focus in this paper is on characterizing the s.d.o.f., i.e., secrecy rate scaling at high SNR, for the channel, we use Gaussian signaling whenever possible for the achievability for this reason.

In order to extend real interference alignment to complex channels, we need to utilize different results than those used for real channels. For real channels, to analyze the decoder performance, reference [43] proposed utilizing the convergence part of Khintchine-Groshev theorem in the field of Diophantine approximation [44], which deals with the approximation of real numbers with rational numbers. For complex channels, transforming the channel into a real channel with twice the dimensions, as is usually the convention, is not sufficient here, since real interference alignment relies on the linear independence over rational numbers of the channel gains, which does not continue to hold after such channel transformation. Luckily, we can utilize a result in the field of classification of transcendental complex numbers, which provides a bound on the absolute value of a complex algebraic number with rational coefficients in terms of its height, i.e., the maximum coefficient [38]–[40]. For complex channel coefficients, this result ends up playing the same role of the Khintchine-Groshev theorem for real coefficients.

Before continuing with the achievability proof for the different cases, we state the following lemma, which is utilized to show the linear independence between the directions of the received streams at the legitimate receiver.

Lemma 1: Consider two matrices $\mathbf{E}_1 \in \mathbb{C}^{N\times K}$ and $\mathbf{E}_2 \in \mathbb{C}^{K\times M}$, where $N, M < K$. If the matrix $\mathbf{E}_2$ is full column rank and the matrix $\mathbf{E}_1$ has all of its entries independently and randomly drawn according to a continuous distribution, then $\operatorname{rank}(\mathbf{E}_1\mathbf{E}_2) = \min(N, M)$ *a.s.*

Proof: The proof of Lemma 1 is given in Appendix C. ■

A. Case 1: $N_e \le N$ and $0 \le N_c \le \frac{N_e}{2}$

The s.d.o.f. for this case is equal to $N + N_c - N_e$, i.e., integer valued, for which we utilize Gaussian signaling and cooperative jamming. Since $N_e \le N$, the transmitter exploits this advantage by sending a part of its signal invisible to the eavesdropper. There is no need for linear precoding at the cooperative jammer for this case. Increasing the number of the cooperative jammer antennas, N_c, increases the s.d.o.f. of the channel.

The transmitted signals, $\mathbf{X}_t$ and $\mathbf{X}_c$, are given by (72) with $d = N + N_c - N_e$, $l = N_c$, $\mathbf{U}_t \sim \mathcal{CN}(\mathbf{0}, \bar{P}\mathbf{I}_d)$, $\mathbf{V}_c \sim \mathcal{CN}(\mathbf{0}, \bar{P}\mathbf{I}_l)$, and $\bar{P} = \frac{1}{\alpha}P$, in accordance with the power constraints on the transmitted signals at the transmitter and the cooperative jammer, where $\alpha = \max\left\{l, \sum_{i=1}^{d} ||\mathbf{p}_{t,i}||^2\right\}$ is a constant which does not depend on the power P. The precoders $\mathbf{P}_c$ and $\mathbf{P}_t$ are given by $\mathbf{P}_c = \mathbf{I}_l$, and

$$\mathbf{P}_t = [\mathbf{P}_{t,a} \;\; \mathbf{P}_{t,n}] \in \mathbb{C}^{N\times d}, \tag{73}$$

where $\mathbf{P}_{t,a} = \mathbf{G}_t^{\dagger}\mathbf{G}_c$ in order to align the information streams over the cooperative jamming streams at the eavesdropper, and the $N - N_e$ columns of $\mathbf{P}_{t,n}$ are chosen to span $\mathcal{N}(\mathbf{G}_t)$. The achievability scheme for this case, when $N = 4$, $N_e = 2$, and $N_c = 1$, is depicted in Fig. 3.

Since $N_c \le \frac{N_e}{2}$, the total number of superposed received streams at the receiver, $2N_c + N - N_e$, is less than or equal to the number of its available spatial dimensions, N. Thus, the receiver can decode all the information and cooperative jamming streams at high SNR. Using (1), (2), and (72), the received signals at the receiver and the eavesdropper are

$$\mathbf{Y}_r = \begin{bmatrix}\mathbf{H}_t\mathbf{P}_t & \mathbf{H}_c\end{bmatrix}\begin{bmatrix}\mathbf{U}_t\\ \mathbf{V}_c\end{bmatrix} + \mathbf{Z}_r, \tag{74}$$

$$\mathbf{Y}_e = \begin{bmatrix}\mathbf{G}_t\mathbf{G}_t^{\dagger}\mathbf{G}_c & \mathbf{0}_{N_e\times(N-N_e)}\end{bmatrix}\begin{bmatrix}\mathbf{U}_{t1}^{\;l}\\ \mathbf{U}_{t\,l+1}^{\;d}\end{bmatrix} + \mathbf{G}_c\mathbf{V}_c + \mathbf{Z}_e \tag{75}$$

$$= \mathbf{G}_c(\mathbf{U}_{t1}^{\;l} + \mathbf{V}_c) + \mathbf{Z}_e. \tag{76}$$

We lower bound the secrecy rate in (71) as follows. First, in order to compute $I(\mathbf{X}_t;\mathbf{Y}_r)$, we show that the matrix $[\mathbf{H}_t\mathbf{P}_t \;\; \mathbf{H}_c] \in \mathbb{C}^{N\times(d+l)}$ in (74) is full column-rank a.s.

The columns of $\mathbf{P}_{t,a} = \mathbf{G}_t^{\dagger}\mathbf{G}_c$ are linearly independent a.s. due to the randomly generated channel gains, and the $N - N_e$ columns of $\mathbf{P}_{t,n}$ are linearly independent as well, since they span an $N - N_e$-dimensional subspace. In addition, each of the columns of $\mathbf{P}_{t,a}$ is linearly independent from the columns of $\mathbf{P}_{t,n}$ a.s. since $\mathbf{G}_t\mathbf{P}_{t,a} = \mathbf{G}_c$, and hence $\mathbf{G}_t\mathbf{p}_{t_i} \neq \mathbf{0}$ for all $i = 1, 2, \cdots, l$. Thus $\mathbf{P}_t = [\mathbf{P}_{t,a} \; \mathbf{P}_{t,n}]$ is full column rank a.s. The matrix $[\mathbf{H}_t\mathbf{P}_t \;\; \mathbf{H}_c]$ can be written as

$$\begin{bmatrix}\mathbf{H}_t\mathbf{P}_t & \mathbf{H}_c\end{bmatrix} = \begin{bmatrix}\mathbf{H}_t & \mathbf{H}_c\end{bmatrix}\begin{bmatrix}\mathbf{P}_t & \mathbf{0}_{N\times l}\\ \mathbf{0}_{l\times d} & \mathbf{I}_l\end{bmatrix}. \tag{77}$$

The matrix $[\mathbf{H}_t \;\; \mathbf{H}_c]$ has all of its entries independently and randomly drawn according to a continuous distribution, while the second matrix on the right hand side (RHS) of (77) is full column rank a.s. By applying Lemma 1 to (77), we have that the matrix $[\mathbf{H}_t\mathbf{P}_t \;\; \mathbf{H}_c]$ is full column rank a.s. Thus, using (74), we obtain the lower bound

$$I(\mathbf{X}_t;\mathbf{Y}_r) \ge d\log P + o(\log P). \tag{78}$$

Next, using (76), we upper bound $I(\mathbf{X}_t;\mathbf{Y}_e)$ as follows:

$$I(\mathbf{X}_t;\mathbf{Y}_e) = h(\mathbf{Y}_e) - h(\mathbf{Y}_e|\mathbf{X}_t) \tag{79}$$

$$= h(\mathbf{G}_c(\mathbf{U}_{t1}^{\;l} + \mathbf{V}_c) + \mathbf{Z}_e) - h(\mathbf{G}_c\mathbf{V}_c + \mathbf{Z}_e) \tag{80}$$

$$= \log\frac{\det(\mathbf{I}_{N_e} + 2\bar{P}\mathbf{G}_c\mathbf{G}_c^H)}{\det(\mathbf{I}_{N_e} + \bar{P}\mathbf{G}_c\mathbf{G}_c^H)} \tag{81}$$

$$= \log\frac{\det(\mathbf{I}_l + 2\bar{P}\mathbf{G}_c^H\mathbf{G}_c)}{\det(\mathbf{I}_l + \bar{P}\mathbf{G}_c^H\mathbf{G}_c)} \tag{82}$$

$$= \log\frac{2^l\det(\frac{1}{2}\mathbf{I}_l + \bar{P}\mathbf{G}_c^H\mathbf{G}_c)}{\det(\mathbf{I}_l + \bar{P}\mathbf{G}_c^H\mathbf{G}_c)} \tag{83}$$

$$\le l. \tag{84}$$

Substituting (78) and (84) in (71), we have

$$R_s \ge d\log P + o(\log P) - l \tag{85}$$

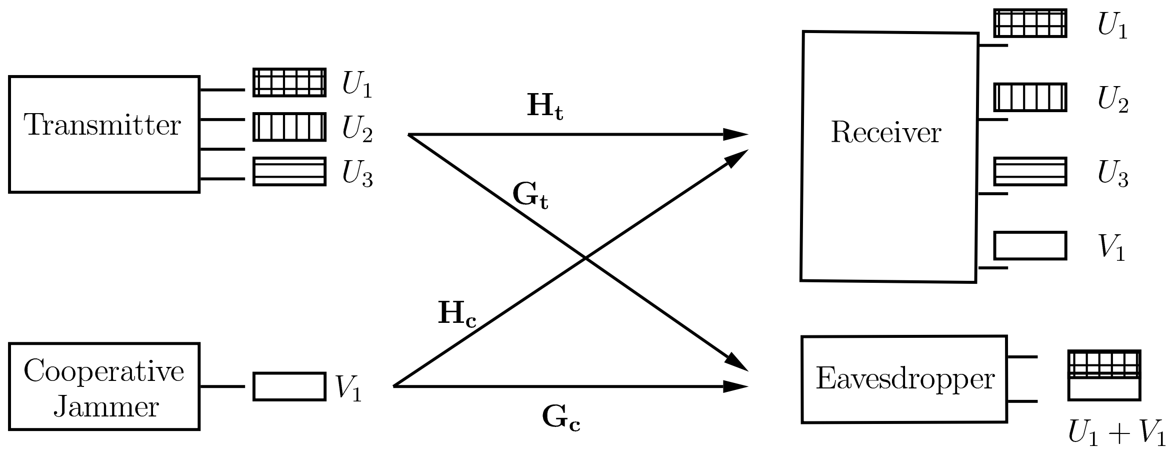


Fig. 3: An example for the achievability scheme for Case 1, when $N = 4$, $N_e = 2$, and $N_c = 1$.

$$= (N + N_c - N_e) \log P + o(\log P) - N_c, \tag{86}$$

and hence, using (5), we conclude that the achievable s.d.o.f. is $D_s \geq N + N_c - N_e$.

B. Case 2: $N_e \leq N$, $\frac{N_e}{2} < N_c \leq N$, and N_e is even

Unlike case 1, the s.d.o.f. for this case does not increase by increasing N_c. For all N_c in this case, the transmitter sends the same number of information streams, while the cooperative jammer utilizes a linear precoder which allows for discarding any unnecessary antennas. The s.d.o.f. here is integer valued, and we use Gaussian signaling for transmission and cooperative jamming.

In particular, for N_e is even, $N_c = \frac{N_e}{2}$, and $N_e \leq N$, the achievable s.d.o.f., using the scheme in Section V-A, is equal to $N - \frac{N_e}{2}$. However, from (69), we observe that the s.d.o.f. is upper bounded by $N - \frac{N_e}{2}$ for all $\frac{N_e}{2} < N_c \leq N$. Thus, when $N_e \leq N$ and N_e is even, the scheme for $N_c = \frac{N_e}{2}$ in Section V-A can be used to achieve the s.d.o.f. for all $\frac{N_e}{2} < N_c \leq N$ by discarding the remaining $N_c - \frac{N_e}{2}$ antennas. That is, the cooperative jammer uses the precoder

$$\mathbf{P}_c = \begin{bmatrix} \mathbf{I}_l \\ \mathbf{0}_{(N_c - l)\times l} \end{bmatrix}, \tag{87}$$

with $l = \frac{N_e}{2}$, to utilize only $\frac{N_e}{2}$ out of its N_c antennas, and the transmitter utilizes

$$\mathbf{P}_t = [\mathbf{P}_{t,a}\ \mathbf{P}_{t,n}], \tag{88}$$

$\mathbf{P}_{t,a} = \mathbf{G}_t^\dagger \mathbf{G}_c \mathbf{P}_c \in \mathbb{C}^{N\times l}$, $\mathbf{P}_{t,n} \in \mathbb{C}^{N\times(N-N_e)}$ is defined as in (73), in order to send $d = N - \frac{N_e}{2}$ Gaussian information streams. Following the same analysis as in the previous case, the achievable s.d.o.f. is $N - \frac{N_e}{2}$ for all $\frac{N_e}{2} < N_c \leq N$, where N_e is even and $N_e \leq N$.

C. Case 3: $N_e \leq N$, $\frac{N_e}{2} < N_c \leq N$, and N_e is odd

The s.d.o.f. for this case is equal to $N - \frac{N_e}{2}$, which is not an integer. As Gaussian signaling cannot achieve fractional s.d.o.f. for the channel, we utilize structured signaling both for transmission and cooperative jamming for this case. In particular, we propose utilizing *joint* signal space alignment and the complex field equivalent of real interference alignment [37], [38].

The decoding scheme at the receiver is as follows. The receiver projects its received signal over a direction that is orthogonal to all but one information and one cooperative jamming streams. Then, the receiver decodes these two streams from the projection using complex field analogy of real interference alignment. Finally, the receiver removes the decoded information and cooperative jamming streams from its received signal, leaving $N - 1$ spatial dimensions for the other $N - \frac{N_e+1}{2}$ information and $\frac{N_e-1}{2}$ cooperative jamming streams.

Before continuing with the details for the achievability scheme for this case, we provide the following example, which illustrates the ideas utilized for this case.

Example 1: Consider a multi-antenna wire-tap channel with 4-antenna transmitter, 4-antenna receiver, 3-antenna eavesdropper, and 2-antenna cooperative jammer as shown in Fig. 4.
The transmitter sends 3 structured information streams, U_1, U_2, U_3, and the cooperative jammer sends 2 structured jamming streams, V_1, V_2. The streams U_1, V_1 are integer valued, while the streams U_2, U_3, V_2, are complex integers. That is, $U_2 = U_{2,\mathrm{Re}} + jU_{2,\mathrm{Im}}, U_3 = U_{3,\mathrm{Re}} + jU_{3,\mathrm{Im}}$, and $V_2 = V_{2,\mathrm{Re}} + jV_{2,\mathrm{Im}}$, where $\{U_1, U_{2,\mathrm{Re}}, U_{2,\mathrm{Im}}, U_{3,\mathrm{Re}}, U_{3,\mathrm{Im}}, V_1, V_{2,\mathrm{Re}}, V_{2,\mathrm{Im}}\}$ are i.i.d. random variables uniform over a set of integer that scales with the transmit power as it will be explained later in (89). The transmitter chooses its precoder as in (88) so that U_3 is sent over $\mathcal{N}(\mathbf{G}_t)$, and hence U_3 is invisible to the eavesdropper, and that U_1, V_1 and U_2, V_2 are perfectly aligned at the eavesdropper. The cooperative jammer chooses its precoder as in (87) so that it utilizes only 2 out of its 3 antennas to send V_1, V_2. The legitimate receiver projects its received signal over a single dimension that is orthogonal to $\{U_2, U_3, V_2\}$, and hence, only U_1 and V_1 remain in this dimension. The received signal after projection is of the form $f_1 U_1 + f_2 V_1 + Z$, where f_1, f_2 are the coefficients resulting from multiplying the channel gains with the projection matrix, and Z is the projection of the Gaussian noise over

the single dimension. The receiver utilizes a hard decision decoder which maps $f_1U_1 + f_2V_1 + Z$ to the nearest point in the constellation of $f_1U_1 + f_2V_1$. It has been shown in [43] that U_1, V_1 can be uniquely decoded from $f_1U_1 + f_2V_1$. Thus, the receiver decodes U_1, V_1, subtracts them from its original received signal, and then utilizes the remaining 3 dimensions in its signal space to decode U_2, U_3, V_2. Thus, the receiver utilizes 2.5 dimensions to decode the information streams, i.e., 2.5 useful dimensions, where each of U_2 and U_3 is decoded from a separate dimension while both U_1 and V_1 are decoded from a single dimension (each occupies half of that dimension), leading to 2.5 achievable s.d.o.f. for the channel.

Now, we continue with the detailed explanation for the achievability scheme for this case. The transmitted signals are given by (72), with $d = N - \frac{N_e-1}{2}$, $l = \frac{N_e+1}{2}$, $\mathbf{P}_c, \mathbf{P}_t$ are defined as in (87) and (88), and $U_i = U_{i,\mathrm{Re}} + jU_{i,\mathrm{Im}}$, $V_k = V_{k,\mathrm{Re}} + jV_{k,\mathrm{Im}}$, $i = 2, 3, \cdots, d$ and $k = 2, 3, \cdots, l$. The random variables U_1, V_1, $\{U_{i,\mathrm{Re}}\}_{i=2}^d$, $\{U_{i,\mathrm{Im}}\}_{i=2}^d$, $\{V_{i,\mathrm{Re}}\}_{i=2}^l$, and $\{V_{i,\mathrm{Im}}\}_{i=2}^l$ are i.i.d. uniform over the set $\{a(-Q,Q)_{\mathbb{Z}}\}$. The values for a and the integer Q are chosen as

$$Q = \left\lfloor P^{\frac{1-\epsilon}{2+\epsilon}} \right\rfloor = P^{\frac{1-\epsilon}{2+\epsilon}} - \nu \tag{89}$$

$$a = \gamma P^{\frac{3\epsilon}{2(2+\epsilon)}}, \tag{90}$$

in order to satisfy the power constraints, where ϵ is an arbitrarily small positive number, and ν, γ are constants that do not depend on the power P. Justification for the choice of a and Q is provided in Appendix D.

The received signal at the eavesdropper is

$$\mathbf{Y}_e = \tilde{\mathbf{G}}_c(\mathbf{U}_{t1}^{\,l} + \mathbf{V}_c) + \mathbf{Z}_e, \tag{91}$$

where $\tilde{\mathbf{G}}_c = \mathbf{G}_c\mathbf{P}_c$. We upper bound the second term in (71), $I(\mathbf{X}_t; \mathbf{Y}_e)$, as follows:

$$\begin{aligned}
I(\mathbf{X}_t;\mathbf{Y}_e) &\le I(\mathbf{X}_t;\mathbf{Y}_e,\mathbf{Z}_e) && (92)\\
&= I(\mathbf{X}_t;\mathbf{Y}_e|\mathbf{Z}_e) && (93)\\
&= H(\mathbf{Y}_e|\mathbf{Z}_e) - H(\mathbf{Y}_e|\mathbf{Z}_e,\mathbf{X}_t) && (94)\\
&= H\left(\tilde{\mathbf{G}}_c(\mathbf{U}_{t1}^{\,l}+\mathbf{V}_c)\right) - H\left(\tilde{\mathbf{G}}_c\mathbf{V}_c\right) && (95)\\
&= H(\mathbf{U}_{t1}^{\,l}+\mathbf{V}_c) - H(\mathbf{V}_c) && (96)\\
&= H\left(\begin{array}{c} U_1+V_1,\ U_{2,\mathrm{Re}}+V_{2,\mathrm{Re}},\ U_{2,\mathrm{Im}}+V_{2,\mathrm{Im}},\\ \cdots,\ U_{l,\mathrm{Re}}+V_{l,\mathrm{Re}},\ U_{l,\mathrm{Im}}+V_{l,\mathrm{Im}}\end{array}\right)\\
&\quad - H\left(V_1,\ V_{2,\mathrm{Re}},\ V_{2,\mathrm{Im}},\ \cdots,\ V_{l,\mathrm{Re}},\ V_{l,\mathrm{Im}}\right) && (97)\\
&\le \log(4Q+1)^{2l-1} - \log(2Q+1)^{2l-1} && (98)\\
&= (2l-1)\log\frac{4Q+1}{2Q+1} && (99)\\
&\le 2l-1, && (100)
\end{aligned}$$

where (93) follows since $\mathbf{X}_t$ and $\mathbf{Z}_e$ are independent, and (98) follows since the entropy of a uniform random variable over the set $\{a(-2Q,2Q)_{\mathbb{Z}}\}$ upper bounds the entropy of each of $U_1+V_1, U_{2,\mathrm{Re}}+V_{2,\mathrm{Re}}, U_{2,\mathrm{Im}}+V_{2,\mathrm{Im}}, \cdots, U_{l,\mathrm{Im}}+V_{l,\mathrm{Im}}$. Equation (96) follows since the mappings $\mathbf{U}_{t1}^{\,l} + \mathbf{V}_c \mapsto \tilde{\mathbf{G}}_c(\mathbf{U}_{t1}^{\,l} + \mathbf{V}_c)$ and $\mathbf{V}_c \mapsto \tilde{\mathbf{G}}_c\mathbf{V}_c$ are bijective. The reason for this is that the entries of $\tilde{\mathbf{G}}_c$ are *rationally independent*, as illustrated in Definition 1 below, and that $(\mathbf{U}_{t1}^{\,l} + \mathbf{V}_c)$ and $\mathbf{V}_c$ belong to $\mathbb{Z}_{\mathbb{C}}^l$.

Definition 1: A set of complex numbers $\{c_1, c_2, \cdots, c_L\}$ are rationally independent, i.e., linearly independent over $\mathbb{Q}$, if there is no set of rational numbers $\{r_i\}$, $r_i \neq 0$ for all $i = 1, 2, \cdots, L$, such that $\sum_{i=1}^L r_ic_i = 0$.

Next, we derive a lower bound for $I(\mathbf{X}_t; \mathbf{Y}_r)$. The received signal at the legitimate receiver is given by

$$\mathbf{Y}_r = \mathbf{A}\mathbf{U}_t + \mathbf{H}_c'\mathbf{V}_c + \mathbf{Z}_r, \tag{101}$$

where $\mathbf{A} = \mathbf{H}_t\mathbf{P}_t = [\mathbf{a}_1\ \mathbf{a}_2\ \cdots\ \mathbf{a}_d]$ and $\mathbf{H}_c' = \mathbf{H}_c\mathbf{P}_c = [\mathbf{h}_{c,1}\ \mathbf{h}_{c,2}\ \cdots\ \mathbf{h}_{c,l}]$. The receiver chooses $\mathbf{b} \in \mathbb{C}^N$ such that $\mathbf{b} \perp \mathrm{span}\{\mathbf{a}_2, \cdots, \mathbf{a}_d, \mathbf{h}_{c,2}, \cdots, \mathbf{h}_{c,l}\}$ and obtains

$$\widetilde{\mathbf{Y}}_r = \mathbf{D}\mathbf{Y}_r \tag{102}$$

where

$$\mathbf{D} \triangleq \begin{bmatrix} & \mathbf{b}^H & \\ \mathbf{0}_{(N-1)\times 1} & & \mathbf{I}_{N-1}\end{bmatrix}. \tag{103}$$

Note that $(d-1)+(l-1) = N - \frac{N_e-1}{2} + \frac{N_e+1}{2} - 2 = N-1$, and hence the dimension of $\mathrm{span}\{\mathbf{a}_2, \cdots, \mathbf{a}_d, \mathbf{h}_{c,2}, \cdots, \mathbf{h}_{c,l}\}$ is at most $N-1$. This shows the existence of a vector $\mathbf{b} \in \mathbb{C}^N$ such that $\mathbf{b} \perp \mathrm{span}\{\mathbf{a}_2, \cdots, \mathbf{a}_d, \mathbf{h}_{c,2}, \cdots, \mathbf{h}_{c,l}\}$.

Due to the fact that channel gains are continuous and randomly generated, $\mathbf{a}_1$ and $\mathbf{h}_{c,1}$ are linearly independent from $\mathrm{span}\{\mathbf{a}_2, \cdots, \mathbf{a}_d, \mathbf{h}_{c,2}, \cdots, \mathbf{h}_{c,l}\}$, and hence, $\mathbf{b}$ is not orthogonal to $\mathbf{a}_1$ and $\mathbf{h}_{c,1}$ a.s. Thus, we have

$$\begin{aligned}\widetilde{\mathbf{Y}}_r = \begin{bmatrix}\widetilde{Y}_{r_1}\\ \widetilde{\mathbf{Y}}_{r_2}^N\end{bmatrix} &= \begin{bmatrix}\mathbf{b}^H\mathbf{a}_1 & \mathbf{0}_{1\times(d-1)}\\ & \widetilde{\mathbf{A}}\end{bmatrix}\begin{bmatrix}U_1\\ \mathbf{U}_{t2}^{\,d}\end{bmatrix}\\
&+ \begin{bmatrix}\mathbf{b}^H\mathbf{h}_{c,1} & \mathbf{0}_{1\times(l-1)}\\ & \widetilde{\mathbf{H}}_c\end{bmatrix}\begin{bmatrix}V_1\\ \mathbf{V}_{c2}^{\,l}\end{bmatrix} + \begin{bmatrix}\mathbf{b}^H\mathbf{Z}_r\\ \mathbf{Z}_{r2}^{\,N}\end{bmatrix},\end{aligned} \tag{104}$$

where $\widetilde{\mathbf{A}} = [\tilde{\mathbf{a}}_1\ \tilde{\mathbf{a}}_2\ \cdots\ \tilde{\mathbf{a}}_d] \in \mathbb{C}^{(N-1)\times d}$, $\tilde{\mathbf{a}}_i = \mathbf{a}_{i2}^{\,N}$ for all $i = 1, 2, \cdots, d$. Similarly, $\widetilde{\mathbf{H}}_c = [\tilde{\mathbf{h}}_{c,1}\ \tilde{\mathbf{h}}_{c,2}\ \cdots\ \tilde{\mathbf{h}}_{c,l}] \in \mathbb{C}^{(N-1)\times l}$, where $\tilde{\mathbf{h}}_{c,i} = \mathbf{h}_{c,i2}^{\,N}$ for all $i = 1, 2, \cdots, l$.

Next, the receiver uses $\widetilde{Y}_{r_1}$ to decode the information stream U_1 and the cooperative jamming stream V_1 as follows. Let $Z' = \mathbf{b}^H\mathbf{Z}_r \sim \mathcal{CN}(0, ||\mathbf{b}||^2)$, $f_1 = \mathbf{b}^H\mathbf{a}_1$, and $f_2 = \mathbf{b}^H\mathbf{h}_{c,1}$. Thus, $\widetilde{Y}_{r_1}$ is given by

$$\widetilde{Y}_{r_1} = f_1U_1 + f_2V_1 + Z'. \tag{105}$$

Once again, with randomly generated channel gains, $f_1 = \mathbf{b}^H\mathbf{a}_1$ and $f_2 = \mathbf{b}^H\mathbf{h}_{c,1}$ are rationally independent a.s. Thus, the mapping $(U_1, V_1) \mapsto f_1U_1 + f_2V_1$ is invertible [43]. The receiver employs a hard decision decoder which maps $\widetilde{Y}_{r_1} \in \mathcal{Y}_{r_1}$ to the nearest point in the constellation $\mathcal{R}_1 = f_1\mathcal{U}_1 + f_2\mathcal{V}_1$, where $\mathcal{U}_1, \mathcal{V}_1 = \{a(-Q,Q)_{\mathbb{Z}}\}$. Then, the receiver passes the output of the hard decision decoder through the bijective mapping $f_1U_1 + f_2V_1 \mapsto (U_1, V_1)$ in order to decode both U_1 and V_1.

The receiver can now use

$$\begin{aligned}\bar{\bar{\mathbf{Y}}}_r &= \widetilde{\mathbf{Y}}_{r_2}^N - \tilde{\mathbf{a}}_1U_1 - \tilde{\mathbf{h}}_{c,1}V_1 && (106)\\
&= \begin{bmatrix}\tilde{\mathbf{a}}_2 & \cdots & \tilde{\mathbf{a}}_d\end{bmatrix}\mathbf{U}_{t2}^{\,d} + \begin{bmatrix}\tilde{\mathbf{h}}_{c,2} & \cdots & \tilde{\mathbf{h}}_{c,l}\end{bmatrix}\mathbf{V}_{c2}^{\,l} + \mathbf{Z}_{r2}^{\,N} && (107)\\
&= \mathbf{B}\begin{bmatrix}\mathbf{U}_{t2}^{\,d}\\ \mathbf{V}_{c2}^{\,l}\end{bmatrix} + \mathbf{Z}_{r2}^{\,N}, && (108)\end{aligned}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2738011, IEEE Transactions on Information Theory

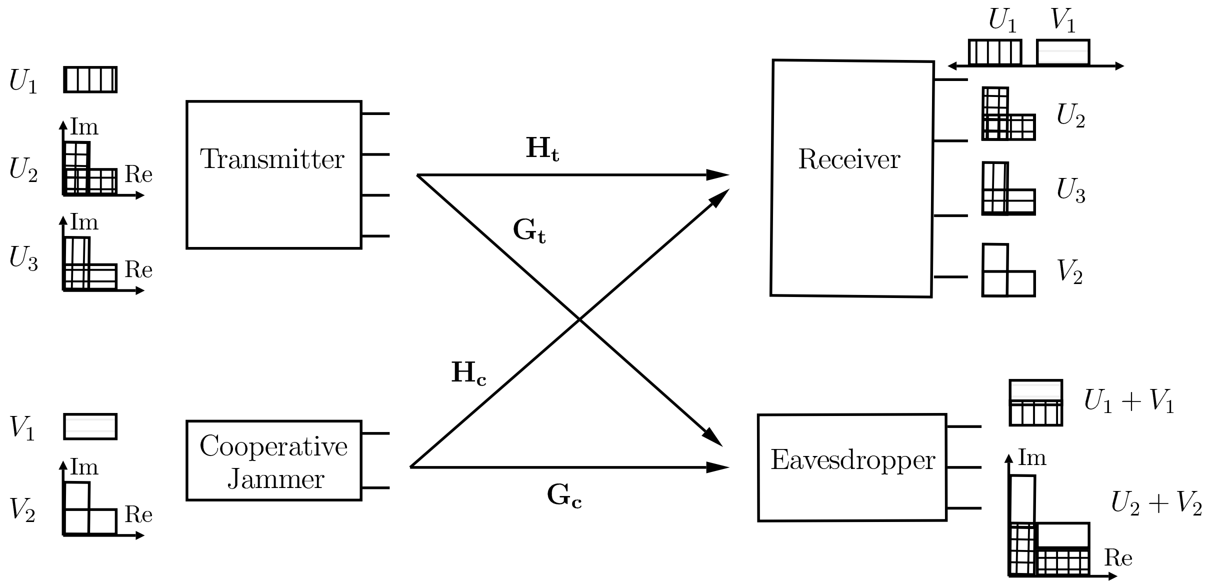


Fig. 4: An example for the achievability scheme for Case 3, when $N = 4$, $N_e = 3$, and $N_c = 2$.

to decode $U_2, \cdots, U_d$, where,

$$\mathbf{B} \triangleq \begin{bmatrix}\tilde{\mathbf{a}}_2 & \cdots & \tilde{\mathbf{a}}_d & \tilde{\mathbf{h}}_{c,2} & \cdots & \tilde{\mathbf{h}}_{c,l}\end{bmatrix} \in \mathbb{C}^{(N-1)\times(N-1)}, \quad (109)$$

is full rank a.s. To show that $\mathbf{B}$ is full rank a.s., let $\bar{\mathbf{H}}_t$ and $\bar{\mathbf{H}}_c$ be generated by removing the first row from $\mathbf{H}_t$ and $\mathbf{H}_c$, and let $\bar{\mathbf{P}}_t$ and $\bar{\mathbf{P}}_c$ be generated by removing the first column from $\mathbf{P}_t$ and $\mathbf{P}_c$, respectively. $\mathbf{B}$ can be rewritten as

$$\mathbf{B} = \begin{bmatrix}\bar{\mathbf{H}}_t & \bar{\mathbf{H}}_c\end{bmatrix} \begin{bmatrix}\bar{\mathbf{P}}_t & \mathbf{0}_{N\times(l-1)} \\ \mathbf{0}_{N_c\times(d-1)} & \bar{\mathbf{P}}_c\end{bmatrix}. \quad (110)$$

Note that $\begin{bmatrix}\bar{\mathbf{H}}_t \; \bar{\mathbf{H}}_c\end{bmatrix}$ has all of its entries independently and randomly drawn from a continuous distribution, and the second matrix in the RHS of (110) is full column rank. Using Lemma 1, the matrix $\mathbf{B}$ is full rank a.s.

Hence, by zero forcing, the receiver obtains

$$\widehat{\mathbf{Y}}_r = \mathbf{B}^{-1}\bar{\bar{\mathbf{Y}}}_r = \begin{bmatrix}\mathbf{U}_{t2}^{d} \\ \mathbf{V}_{c2}^{l}\end{bmatrix} + \bar{\mathbf{Z}}_r, \quad (111)$$

where $\bar{\mathbf{Z}}_r = \mathbf{B}^{-1}\mathbf{Z}_{r2}^{N} \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{B}^{-1}\mathbf{B}^{-H}\right)$. Thus, at high SNR, the receiver can decode the other information streams, $U_2, \cdots, U_d$, from $\widehat{\mathbf{Y}}_r$.

The mutual information between the transmitter and receiver is lower bounded as follows:

$$\begin{aligned}
I(\mathbf{X}_t; \mathbf{Y}_r) &\geq I(\mathbf{U}_t; \widetilde{\mathbf{Y}}_r) && (112)\\
&= I(U_1, \mathbf{U}_{t2}^{d}; \widetilde{Y}_{r_1}, \widetilde{\mathbf{Y}}_{r_2}^{N}) && (113)\\
&= I(U_1, \mathbf{U}_{t2}^{d}; \widetilde{Y}_{r_1}) + I(U_1, \mathbf{U}_{t2}^{d}; \widetilde{\mathbf{Y}}_{r_2}^{N} | \widetilde{Y}_{r_1}) && (114)\\
&= I(U_1; \widetilde{Y}_{r_1}) + I(\mathbf{U}_{t2}^{d}; \widetilde{Y}_{r_1} | U_1) + I(U_1; \widetilde{\mathbf{Y}}_{r_2}^{N} | \widetilde{Y}_{r_1}) \\
&\quad + I(\mathbf{U}_{t2}^{d}; \widetilde{\mathbf{Y}}_{r_2}^{N} | U_1, \widetilde{Y}_{r_1}) && (115)\\
&\geq I(U_1; \widetilde{Y}_{r_1}) + I(\mathbf{U}_{t2}^{d}; \widetilde{\mathbf{Y}}_{r_2}^{N} | U_1, \widetilde{Y}_{r_1}), && (116)
\end{aligned}$$

where (112) follows since $\mathbf{U}_t - \mathbf{X}_t - \mathbf{Y}_r - \widetilde{\mathbf{Y}}_r$ forms a Markov chain. We next lower bound each term in the RHS of (116).

We lower bound the first term, $I(U_1; \widetilde{Y}_{r_1})$ as follows, see also [37], [43]. Let P_{e_1} denote the probability of error in decoding U_1 at the receiver, i.e., $P_{e_1} \triangleq \Pr\left(\hat{U}_1 \neq U_1\right)$, where $\hat{U}_i$, $i = 1, 2, \cdots, d$, is the estimate of U_i at the legitimate receiver. Thus, using Fano's inequality, we have

$$\begin{aligned}
I(U_1; \widetilde{Y}_{r_1}) &= H(U_1) - H(U_1 | \widetilde{Y}_{r_1}) && (117)\\
&\geq H(U_1) - 1 - P_{e_1} \log |\mathcal{U}_1| && (118)\\
&= (1 - P_{e_1}) \log(2Q + 1) - 1. && (119)
\end{aligned}$$

From (105), since the mapping $(U_1, V_1) \mapsto f_1 U_1 + f_2 V_1$ is invertible, the only source of error in decoding U_1 from $\widetilde{Y}_{r_1}$ is the additive Gaussian noise Z'. Note that, since $Z' \sim \mathcal{CN}(0, ||\mathbf{b}||^2)$, $\mathrm{Re}\{Z'\}$ and $\mathrm{Im}\{Z'\}$ are i.i.d. with $\mathcal{N}\left(0, \frac{||\mathbf{b}||^2}{2}\right)$ distribution, and $|Z'| \sim \text{Rayleigh}\left(\frac{||\mathbf{b}||}{\sqrt{2}}\right)$. Thus, we have

$$\begin{aligned}
P_{e_1} &\triangleq \Pr\left(\hat{U}_1 \neq U_1\right) && (120)\\
&\leq \Pr\left((\hat{U}_1, \hat{V}_1) \neq (U_1, V_1)\right) && (121)\\
&\leq \Pr\left(|Z'| \geq \frac{d_{\min}}{2}\right) && (122)\\
&= \exp\left(\frac{-d_{\min}^2}{4||\mathbf{b}||^2}\right), && (123)
\end{aligned}$$

where $d_{\min}$ is the minimum distance between the points in the constellation $\mathcal{R}_1 = f_1 \mathcal{U}_1 + f_2 \mathcal{V}_1$.

In order to upper bound P_{e_1}, we lower bound $d_{\min}$. To do so, similar to [37], we extend real interference alignment [43] to complex channels. In particular, we utilize the following results from number theory:

Definition 2: [38] *The Diophantine exponent $\omega(\mathbf{z})$ of $\mathbf{z} \in \mathbb{C}^n$ is defined as*

$$\omega(\mathbf{z}) \triangleq \sup \left\{ \begin{array}{c} v : |p + \mathbf{z}.\mathbf{q}| \leq (||\mathbf{q}||_\infty)^{-v} \text{ for} \\ \text{infinetly many } \mathbf{q} \in \mathbb{Z}^n, p \in \mathbb{Z} \end{array} \right\}, \quad (124)$$

where $\mathbf{q} = [q_1 \; q_2 \; \cdots \; q_n]^T$ and $||\mathbf{q}||_\infty = \max_i |q_i|$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2738011, IEEE Transactions on Information Theory

Lemma 2: [38] *For almost all* $\mathbf{z} \in \mathbb{C}^n$*, the Diophantine exponent* $\omega(\mathbf{z})$ *is equal to* $\frac{n-1}{2}$.

Lemma 2 implies the following:

Corollary 1: For almost all $\mathbf{z} \in \mathbb{C}^n$ *and for all* $\epsilon > 0$*,*

$$|p + \mathbf{z}.\mathbf{q}| > (\max_i |q_i|)^{-\frac{(n-1+\epsilon)}{2}}, \tag{125}$$

holds for all $\mathbf{q} \in \mathbb{Z}^n$ *and* $p \in \mathbb{Z}$ *except for finitely many of them.*

Since the number of integers that violate the inequality in (125) is finite, there exists a constant κ such that, for almost all $\mathbf{z} \in \mathbb{C}^n$ and all $\epsilon > 0$, the inequality

$$|p + \mathbf{z}.\mathbf{q}| > \kappa(\max_i |q_i|)^{-\frac{(n-1+\epsilon)}{2}}, \tag{126}$$

holds for all $\mathbf{q} \in \mathbb{Z}^n$ and $p \in \mathbb{Z}$.

Thus, for almost all channel gains, the minimum distance $d_{\min}$ is lower bounded as follows:

$$\begin{aligned}
d_{\min} &= \inf_{Y'_{r_1}, Y''_{r_1} \in \mathcal{R}_1} |Y'_{r_1} - Y''_{r_1}| && (127)\\
&= \inf_{U_1, V_1 \in \{a(-2Q, 2Q)_{\mathbb{Z}}\}} |f_1 U_1 + f_2 V_1| && (128)\\
&= \inf_{U_1, V_1 \in (-2Q, 2Q)_{\mathbb{Z}}} a|f_1| \left| U_1 + \frac{f_2}{f_1} V_1 \right| && (129)\\
&\geq \kappa \frac{a|f_1|}{(2Q)^{\frac{\epsilon}{2}}} && (130)\\
&\geq \kappa\gamma |f_1| 2^{-\frac{\epsilon}{2}} P^{\frac{\epsilon}{2}}, && (131)
\end{aligned}$$

where (130) follows from (126), and (131) follows by substituting (89) and (90) in (130). Substituting (131) in (123) gives the following bound on P_{e_1},

$$P_{e_1} \leq \exp(-\mu P^{\epsilon}), \tag{132}$$

where $\mu = \frac{\kappa^2 \gamma^2 |f_1|^2 2^{-\epsilon}}{4||\mathbf{b}||^2}$ is a constant which does not depend on the power P. Thus, using (119) and (132), we have

$$I(U_1; \widetilde{Y}_{r_1}) \geq (1 - \exp(-\mu P^{\epsilon})) \log(2Q+1) - 1. \tag{133}$$

Next, we lower bound the second term in the RHS of (116), $I(\mathbf{U}_{t2}^{d}; \widetilde{\mathbf{Y}}_{r_2}^{N} | U_1, \widetilde{Y}_{r_1})$. Let $\widetilde{\mathbf{B}} = \begin{bmatrix} \mathbf{0}_{(N-1)\times 1} & \mathbf{I}_{N-1} \end{bmatrix} - \frac{1}{f_2} \tilde{\mathbf{h}}_{c,1} \mathbf{b}^H$, and

$$\bar{\bar{\mathbf{Y}}}'_r = \mathbf{B} \begin{bmatrix} \mathbf{U}_{t2}^{d} \\ \mathbf{V}_{c2}^{l} \end{bmatrix} + \widetilde{\mathbf{B}} \mathbf{Z}_r \tag{134}$$

$$\widehat{\mathbf{Y}}'_r = \mathbf{B}^{-1} \bar{\bar{\mathbf{Y}}}'_r = \begin{bmatrix} \mathbf{U}_{t2}^{d} \\ \mathbf{V}_{c2}^{l} \end{bmatrix} + \mathbf{B}^{-1} \widetilde{\mathbf{B}} \mathbf{Z}_r, \tag{135}$$

where $\mathbf{B}$ is defined as in (109). Thus, we have

$$\begin{aligned}
&I\left(\mathbf{U}_{t2}^{d}; \widetilde{\mathbf{Y}}_{r_2}^{N} | U_1, \widetilde{Y}_{r_1}\right)\\
&= I\left(\mathbf{U}_{t2}^{d}; \widetilde{\mathbf{A}} \mathbf{U}_t + \widetilde{\mathbf{H}}_c \mathbf{V}_c + \mathbf{Z}_{r2}^{N} \middle| U_1, f_2 V_1 + Z'\right) && (136)\\
&= I\left(\mathbf{U}_{t2}^{d}; \mathbf{B} \begin{bmatrix} \mathbf{U}_{t2}^{d} \\ \mathbf{V}_{c2}^{l} \end{bmatrix} + \mathbf{Z}_{r2}^{N} - \frac{1}{f_2} \tilde{\mathbf{h}}_{c,1} \mathbf{b}^H \mathbf{Z}_r \middle| f_2 V_1 + Z'\right) && (137)\\
&= I(\mathbf{U}_{t2}^{d}; \bar{\bar{\mathbf{Y}}}'_r | f_2 V_1 + Z') && (138)\\
&\geq I(\mathbf{U}_{t2}^{d}; \bar{\bar{\mathbf{Y}}}'_r) && (139)\\
&\geq I(\mathbf{U}_{t2}^{d}; \widehat{\mathbf{Y}}'_r) && (140)\\
&= H(\mathbf{U}_{t2}^{d}) - H(\mathbf{U}_{t2}^{d} | \widehat{\mathbf{Y}}'_r) && (141)\\
&\geq H(\mathbf{U}_{t2}^{d}) - P_{e_2}^{d} \log(2Q+1)^{2(d-1)} - 1 && (142)\\
&= 2(d-1)\left(1 - P_{e_2}^{d}\right) \log(2Q+1) - 1, && (143)
\end{aligned}$$

where $P_{e_2}^{d} \triangleq \Pr\left((\hat{U}_2, \hat{U}_3, \cdots, \hat{U}_d) \neq (U_2, U_3, \cdots, U_d)\right)$, (136) follows from (104), (139) follows since $\mathbf{U}_{t2}^{d}$ and $f_2 V_1 + Z'$ are independent, (140) follows since $\mathbf{U}_{t2}^{d} - \bar{\bar{\mathbf{Y}}}'_r - \widehat{\mathbf{Y}}'_r$ forms a Markov chain, and (142) follows from Fano's inequality.

Let $\widehat{\mathbf{Z}}_r \triangleq \mathbf{\Theta} \mathbf{Z}_r = [\hat{Z}_{r_2} \ \cdots \ \hat{Z}_{r_N}]^T$, where $\mathbf{\Theta} = \mathbf{B}^{-1} \widetilde{\mathbf{B}}$. Thus, $\widehat{\mathbf{Z}}_r \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Theta}\mathbf{\Theta}^H)$ and $|\hat{Z}_{r_i}| \sim \mathrm{Rayleigh}(\sigma_i)$, where $\sigma_i^2 = \mathbf{\Theta}\mathbf{\Theta}^H(i,i)$, $i = 2, 3, \cdots, N$. Using the union bound, we have

$$\begin{aligned}
P_{e_2}^{d} &= \Pr\left((\hat{U}_2, \hat{U}_3, \cdots, \hat{U}_d) \neq (U_2, U_3, \cdots, U_d)\right) && (144)\\
&\leq \sum_{i=2}^{d} \Pr\left(\hat{U}_i \neq U_i\right) && (145)\\
&\leq \sum_{i=2}^{d} \Pr\left(|\hat{Z}_{r_i}| \geq \frac{a}{2}\right) && (146)\\
&= \sum_{i=2}^{d} \exp\left(-\frac{a^2}{8\sigma_i^2}\right) && (147)\\
&\leq (d-1) \exp\left(-\frac{\gamma^2}{8\sigma_{\max}^2} P^{\frac{3\epsilon}{2+\epsilon}}\right) && (148)\\
&= (d-1) \exp(-\mu' P^{\epsilon'}), && (149)
\end{aligned}$$

where $\sigma_{\max} = \max_i \sigma_i$, $\mu' = \frac{\gamma^2}{8\sigma_{\max}^2}$, $\epsilon' = \frac{3\epsilon}{2+\epsilon}$, and (148) follows by substituting (90) in (147).

Substituting (149) in (143) yields

$$\begin{aligned}
&I\left(\mathbf{U}_{t2}^{d}; \widetilde{\mathbf{Y}}_{r_2}^{N} | U_1, \widetilde{Y}_{r_1}\right)\\
&\geq \left(2d - 2 - 2(d-1)^2 \exp(-\mu' P^{\epsilon'})\right) \log(2Q+1) - 1. && (150)
\end{aligned}$$

Using (89), (116), (133), and (150), we have

$$\begin{aligned}
I(\mathbf{X}_t; \mathbf{Y}_r) &\geq \Big[2d - 1 - \exp(-\mu P^{\epsilon})\\
&\quad - 2(d-1)^2 \exp(-\mu' P^{\epsilon'})\Big] \log\left(2P^{\frac{1-\epsilon}{2+\epsilon}} - 2\nu + 1\right) - 2 && (151)\\
&= \frac{1-\epsilon}{2+\epsilon} \left[2d - 1 - \exp(-\mu P^{\epsilon}) - 2(d-1)^2 \exp(-\mu' P^{\epsilon'})\right]\\
&\qquad \times \log P + o(\log P). && (152)
\end{aligned}$$

Using the upper bound in (100) and the lower bound in (152), we get

$$\begin{aligned}
R_s &\geq \frac{1-\epsilon}{2+\epsilon} \Big[2d - 1 - \exp(-\mu P^{\epsilon})\\
&\quad - 2(d-1)^2 \exp(-\mu' P^{\epsilon'})\Big] \log P + o(\log P) - (2l - 1) && (153)\\
&= \frac{1-\epsilon}{2+\epsilon} \Big[2N - N_e - \exp(-\mu P^{\epsilon})
\end{aligned}$$

$$-\frac{1}{2}(2N-N_e-1)^2\exp(-\mu' P^{\epsilon'})\Big]\log P+o(\log P)-N_e. \tag{154}$$

Thus, it follows that the s.d.o.f. is lower bounded as

$$D_s \geq \frac{(1-\epsilon)(2N-N_e)}{2+\epsilon}. \tag{155}$$

Since $\epsilon > 0$ can be chosen arbitrarily small, we can achieve s.d.o.f. of $N-\frac{N_e}{2}$.

D. Case 4: $N_e \leq N$, $N < N_c \leq N+N_e$, and $N+N_c-N_e$ is even

Since $N_c > N$ for this case, the cooperative jammer, unlike the previous three cases, chooses its precoder such that N_c-N of its jamming streams are sent invisible to the receiver, in order to allow for more space for the information streams at the receiver. The s.d.o.f. for this case is integer valued, which we can achieve using Gaussian information and cooperative jamming streams.

The transmitted signals are given by (72), with $d=\frac{N+N_c-N_e}{2}$, $l=\frac{N_c+N_e-N}{2}$, $\mathbf{U}_t \sim \mathcal{CN}\left(\mathbf{0},\bar{P}\mathbf{I}_d\right)$, $\mathbf{V}_c \sim \mathcal{CN}\left(\mathbf{0},\bar{P}\mathbf{I}_l\right)$,

$$\mathbf{P}_c = [\mathbf{P}_{c,\mathrm{I}}\ \mathbf{P}_{c,n}], \tag{156}$$

where $\mathbf{P}_{c,\mathrm{I}}$ is given by

$$\mathbf{P}_{c,\mathrm{I}} = \begin{bmatrix} \mathbf{I}_g \\ \mathbf{0}_{(N_c-g)\times g}\end{bmatrix}, \tag{157}$$

$g=\frac{N_e+N-N_c}{2}$, and $\mathbf{P}_{c,n} \in \mathbb{C}^{N_c\times(N_c-N)}$ is a matrix whose columns span $\mathcal{N}(\mathbf{H}_c)$, $\mathbf{P}_t$ is defined as in Section V-B, and $\bar{P}=\frac{1}{\alpha'}P$, where $\alpha' = \max\left\{\sum_{i=1}^{d}||\mathbf{p}_{t,i}||^2, g+\sum_{i=g+1}^{l}||\mathbf{p}_{c,i}||^2\right\}$. At high SNR, the receiver can decode the d information and the g cooperative jamming streams, where $d+g=N$.

The received signals at the legitimate receiver and the eavesdropper are given by

$$\mathbf{Y}_r = \mathbf{H}_t\mathbf{P}_t\mathbf{U}_t + \begin{bmatrix}\mathbf{H}_c\mathbf{P}_{c,\mathrm{I}} & \mathbf{0}_{N\times(N_c-N)}\end{bmatrix}\begin{bmatrix}\mathbf{V}_{c1}^{\ g} \\ \mathbf{V}_{c\,g+1}^{\ l}\end{bmatrix} + \mathbf{Z}_r \tag{158}$$

$$= \begin{bmatrix}\mathbf{H}_t\mathbf{P}_t & \mathbf{H}_c\mathbf{P}_{c,\mathrm{I}}\end{bmatrix}\begin{bmatrix}\mathbf{U}_t \\ \mathbf{V}_{c1}^{\ g}\end{bmatrix} + \mathbf{Z}_r \tag{159}$$

$$\mathbf{Y}_e = \tilde{\mathbf{G}}_c(\mathbf{U}_{t1}^{\ l} + \mathbf{V}_c) + \mathbf{Z}_e, \tag{160}$$

where $\tilde{\mathbf{G}}_c = \mathbf{G}_c\mathbf{P}_c$.

The matrix $[\mathbf{H}_t\mathbf{P}_t\ \mathbf{H}_c\mathbf{P}_{c,\mathrm{I}}] \in \mathbb{C}^{N\times N}$ in (159) can be rewritten as

$$\begin{bmatrix}\mathbf{H}_t\mathbf{P}_t & \mathbf{H}_c\mathbf{P}_{c,\mathrm{I}}\end{bmatrix} = \begin{bmatrix}\mathbf{H}_t & \mathbf{H}_c\end{bmatrix}\begin{bmatrix}\mathbf{P}_t & \mathbf{0}_{N\times g} \\ \mathbf{0}_{N_c\times d} & \mathbf{P}_{c,\mathrm{I}}\end{bmatrix}. \tag{161}$$

By applying Lemma 1 on (161), the matrix $[\mathbf{H}_t\mathbf{P}_t\ \mathbf{H}_c\mathbf{P}_{c,\mathrm{I}}]$ is full rank a.s. Thus,

$$I(\mathbf{X}_t;\mathbf{Y}_r) \geq d\log P + o(\log P). \tag{162}$$

Using similar steps as from (79) to (84), we can show that

$$I(\mathbf{X}_t;\mathbf{Y}_e) = \log\frac{\det(\mathbf{I}_l + 2\bar{P}\tilde{\mathbf{G}}_c^H\tilde{\mathbf{G}}_c)}{\det(\mathbf{I}_l + \bar{P}\tilde{\mathbf{G}}_c^H\tilde{\mathbf{G}}_c)} \leq l. \tag{163}$$

Thus, the achievable secrecy rate in (71) is lower bounded as

$$R_s \geq d\log P + o(\log P) - l \tag{164}$$

$$= \frac{N+N_c-N_e}{2}\log P + o(\log P) - \frac{N_c+N_e-N}{2}, \tag{165}$$

and, using (5), the s.d.o.f. is lower bounded as

$$D_s \geq \frac{N+N_c-N_e}{2}. \tag{166}$$

E. Case 5: $N_e \leq N$, $N < N_c \leq N+N_e$, and $N+N_c-N_e$ is odd

As in case 3, the s.d.o.f. for this case is not an integer, and as in case 4, we have $N_c > N$, which allows the cooperative jammer to send some signals invisible to the receiver. Consequently, the achievable scheme for this case combines the techniques used in Sections V-C and V-D.

The transmitted signals are given by (72) with $d=\frac{N+N_c-N_e+1}{2}$, $l=\frac{N_c+N_e-N+1}{2}$, $\mathbf{P}_t$ and $\mathbf{P}_c$ are defined as in Section V-D with $g=\frac{N_e+N-N_c+1}{2}$, and $\mathbf{U}_t$, $\mathbf{V}_c$ are defined as in Section V-C. Similar to the proof in Appendix D, the values of Q and a are chosen as in (89) and (90), with

$$\gamma = \frac{1}{\left(\max\left\{\begin{array}{l}|||\mathbf{p}_{t,1}||^2 + 2\sum_{i=2}^{d}||\mathbf{p}_{t,i}||^2, \\ 2g-1+2\sum_{i=g+1}^{l}||\mathbf{p}_{c,i}||^2\end{array}\right\}\right)^{\frac{1}{2}}}, \tag{167}$$

and ν are constants that do not depend on the power P.

The legitimate receiver uses the projection and cancellation technique described in Section V-C in order to decode the information streams. The received signal at the eavesdropper is the same as in (160), with $l=\frac{N_c+N_e-N+1}{2}$. Similar to the derivation from (92) to (100), we have

$$I(\mathbf{X}_t;\mathbf{Y}_e) \leq 2l-1. \tag{168}$$

Let $\mathbf{A} = \mathbf{H}_t\mathbf{P}_t = [\mathbf{a}_1\cdots\mathbf{a}_d]$, and $\mathbf{H}'_c = \mathbf{H}_c\mathbf{P}_{c,\mathrm{I}} = [\mathbf{h}_{c,1}\cdots\mathbf{h}_{c,g}]$. The received signal at the legitimate receiver is

$$\mathbf{Y}_r = \begin{bmatrix}\mathbf{A} & \mathbf{H}'_c\end{bmatrix}\begin{bmatrix}\mathbf{U}_t \\ \mathbf{V}_{c1}^{\ g}\end{bmatrix} + \mathbf{Z}_r. \tag{169}$$

As in case 3, we have that $d+g-2=N-1$, and hence the dimension of span$\left\{\mathbf{a}_2,\cdots,\mathbf{a}_d,\mathbf{h}_{c_2},\cdots,\mathbf{h}_{c_g}\right\}$ is at most $N-1$, and there exists $\mathbf{b}\in\mathbb{C}^N$ such that $\mathbf{b}$ is orthogonal to span$\left\{\mathbf{a}_2,\cdots,\mathbf{a}_d,\mathbf{h}_{c_2},\cdots,\mathbf{h}_{c_g}\right\}$. The receiver chooses such $\mathbf{b}$ and multiplies its received signal by the matrix $\mathbf{D}$ given in (103) to obtain $\widetilde{\mathbf{Y}}_r = \left[\widetilde{Y}_{r_1}\ (\widetilde{\mathbf{Y}}_{r_2}^N)^T\right]^T$, where

$$\widetilde{Y}_{r_1} = f_1U_1 + f_2V_1 + Z', \tag{170}$$

$$\widetilde{\mathbf{Y}}_{r_2}^N = \widetilde{\mathbf{A}}\mathbf{U}_t + \widetilde{\mathbf{H}}_c\mathbf{V}_{c1}^{\ g} + \mathbf{Z}_{r2}^{\ N}, \tag{171}$$

f_1, f_2, Z', $\widetilde{\mathbf{A}}$, and $\widetilde{\mathbf{H}}_c$, are defined as in Section V-C. In order to decode U_1 and V_1, the receiver passes $\widetilde{Y}_{r_1}$ through a hard decision decoder, $\widetilde{Y}_{r_1} \mapsto f_1U_1+f_2V_1$, and passes the output of the hard decision decoder through the bijective map $f_1U_1+f_2V_1 \mapsto (U_1,V_1)$, where f_1 and f_2 are rationally independent.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2738011, IEEE Transactions on Information Theory

Using similar steps to the derivation from (112) to (152) in Section V-C, we obtain

$$I(\mathbf{X}_t;\mathbf{Y}_r) \geq \frac{1-\epsilon}{2+\epsilon}\Big[2d-1-\exp\left(-\mu P^{\epsilon}\right) \\ -2(d-1)^2\exp(-\mu' P^{\epsilon'})\Big]\log P + o(\log P), \quad (172)$$

where $\epsilon > 0$ is arbitrarily small, $\epsilon' = \frac{3\epsilon}{2+\epsilon}$, and μ, μ' are constants which do not depend on P.

Thus, the achievable secrecy rate in (71) is lower bounded as

$$R_s \geq \frac{1-\epsilon}{2+\epsilon}\Big[2d-1-\exp(-\mu P^{\epsilon})-(d-1)^2\exp(-\mu' P^{\epsilon'})\Big] \\ \times \log P + o(\log P) - (2l-1) \quad (173)$$

$$= \frac{1-\epsilon}{2+\epsilon}\Big[N+N_c-N_e-\exp(-\mu P^{\epsilon}) \\ -\frac{1}{2}(N+N_c-N_e-1)^2\exp(-\mu' P^{\epsilon'})\Big]\log P \\ + o(\log P) - (N_c+N_e-N), \quad (174)$$

and hence the s.d.o.f is lower bounded as

$$D_s \geq \frac{(1-\epsilon)(N+N_c-N_e)}{2+\epsilon}. \quad (175)$$

Since $\epsilon > 0$ can be chosen arbitrarily small, $D_s = \frac{N+N_c-N_e}{2}$ is achievable for this case, which completes the achievability of (69). Next, we show the achievability of (70), where $N_e > N$, i.e., the eavesdropper has more antennas than the legitimate receiver.

F. Case 6: $N_e > N$ and $N_e - N < N_c \leq N_e - \frac{N}{2}$

Unlike the previous five cases, since $N_e > N$, no information streams can be sent invisible to the eavesdropper. In fact, the precoder at the transmitter is not adequate for achieving the alignment of the information and cooperative jamming streams at the eavesdropper. We need to design both precoders at the transmitter and the cooperative jammer to take part in achieving the alignment condition. The s.d.o.f. here is integer valued, and hence we can utilize Gaussian streams.

The transmitted signals are given by (72), with $d = l = N + N_c - N_e$, and $\mathbf{U}_t, \mathbf{V}_c \sim \mathcal{CN}\left(\mathbf{0}, \bar{P}\mathbf{I}_d\right)$. The matrices $\mathbf{P}_t$ and $\mathbf{P}_c$ are chosen as follows. Let $\mathbf{G} = [\mathbf{G}_t \ \ -\mathbf{G}_c] \in \mathbb{C}^{N_e\times(N+N_c)}$, and let $\mathbf{Q} \in \mathbb{C}^{(N+N_c)\times d}$ be a matrix whose columns are randomly[10] chosen to span $\mathcal{N}(\mathbf{G})$. Write the matrix $\mathbf{Q}$ as $\mathbf{Q} = \left[\mathbf{Q}_1^T \ \ \mathbf{Q}_2^T\right]^T$, where $\mathbf{Q}_1 \in \mathbb{C}^{N\times d}$ and $\mathbf{Q}_2 \in \mathbb{C}^{N_c\times d}$. Set $\mathbf{P}_t = \mathbf{Q}_1$ and $\mathbf{P}_c = \mathbf{Q}_2$. $\bar{P} = \frac{1}{\alpha''}P$, where $\alpha'' = \max\left\{\sum_{i=1}^d ||\mathbf{p}_{t,i}||^2, \sum_{i=1}^d ||\mathbf{p}_{c,i}||^2\right\}$.

The choice of $\mathbf{P}_t$ and $\mathbf{P}_c$ results in $\mathbf{G}_t\mathbf{P}_t = \mathbf{G}_c\mathbf{P}_c$. Thus, the eavesdropper receives

$$\mathbf{Y}_e = \mathbf{G}_c\mathbf{P}_c(\mathbf{U}_t + \mathbf{V}_c) + \mathbf{Z}_e. \quad (176)$$

Similar to going from (79) to (84), it follows that we have

$$I(\mathbf{X}_t;\mathbf{Y}_e) \leq N + N_c - N_e. \quad (177)$$

[10]Out of all possible sets of $d = N+N_c-N_e$ linearly independent vectors which span $\mathcal{N}(\mathbf{G})$, the columns of $\mathbf{Q}$ are the elements of one randomly chosen set.

The received signal at the receiver in turn is given by

$$\mathbf{Y}_r = \begin{bmatrix}\mathbf{H}_t\mathbf{P}_t & \mathbf{H}_c\mathbf{P}_c\end{bmatrix}\begin{bmatrix}\mathbf{U}_t\\ \mathbf{V}_c\end{bmatrix} + \mathbf{Z}_r. \quad (178)$$

Note that, without conditioning on $\mathbf{G}_t$ and $\mathbf{G}_c$, the matrix $\mathbf{Q}$ has all of its entries independently and randomly drawn according to a continuous distribution. Thus, each of $\mathbf{P}_t$ and $\mathbf{P}_c$ is full column rank a.s. Thus, by using Lemma 1, we can show that the matrix $[\mathbf{H}_t\mathbf{P}_t \ \ \mathbf{H}_c\mathbf{P}_c]$ is full column rank a.s. Using (178), we have

$$I(\mathbf{X}_t;\mathbf{Y}_r) \geq (N + N_c - N_e)\log P + o(\log P). \quad (179)$$

Hence, using (177), (179), (71), and (5), the s.d.o.f. is lower bounded as $D_s \geq N + N_c - N_e$.

G. Case 7: $N_e > N$, $N_e - \frac{N}{2} < N_c \leq N_e$, and N is even

The s.d.o.f. for this case does not increase by increasing N_c. The scheme in Section V-F for $N_c = N_e - \frac{N}{2}$, i.e., $d = \frac{N}{2}$, can be used to achieve the s.d.o.f. for all $N_e - \frac{N}{2} < N_c \leq N_e$, when $N_e > N$ and N is even. However, since $\dim(\mathcal{N}(\mathbf{G})) = N + N_c - N_e > \frac{N}{2}$, the $d = \frac{N}{2}$ columns of the matrix $\mathbf{Q}$ are randomly chosen as linearly independent vectors from $\mathcal{N}(\mathbf{G})$. Following the same analysis as in Section V-F, we can show that the s.d.o.f. is lower bounded as $D_s \geq \frac{N}{2}$.

H. Case 8: $N_e > N$, $N_e - \frac{N}{2} < N_c \leq N_e$, and N is odd

The difference here from Section V-G is that the s.d.o.f. is not an integer, and hence, structured signaling for transmission and cooperative jamming is needed, and the difference from V-C is that $N_e > N$, and hence both the precoders at the transmitter and cooperative jammer have to participate in achieving the alignment condition at the eavesdropper.

The transmitted signals are given by (72), with $d = l = \frac{N+1}{2}$, $\mathbf{U}_t$ and $\mathbf{V}_c$ are defined as in Section V-C, and the values for Q and a are chosen as in (89) and (90), with

$$\gamma = \frac{1}{\left(\max\left\{\begin{array}{c} |||\mathbf{p}_{t,1}||^2 + 2\sum_{i=2}^d ||\mathbf{p}_{t,i}||^2, \\ ||\mathbf{p}_{c,1}||^2 + 2\sum_{i=2}^d ||\mathbf{p}_{c,i}||^2 \end{array}\right\}\right)^{\frac{1}{2}}}, \quad (180)$$

and ν are constants which do not depend P. $\mathbf{P}_t, \mathbf{P}_c$ are chosen as in Section V-G, with $d = \frac{N+1}{2}$. The eavesdropper's received signal is the same as in (176). Similar to (92)-(100), we have

$$I(\mathbf{X}_t;\mathbf{Y}_e) \leq N. \quad (181)$$

The receiver employs the decoding scheme in Sections V-C and V-E. Following similar steps as in Sections V-C and V-E, we have

$$I(\mathbf{X}_t;\mathbf{Y}_r) \geq \frac{(1-\epsilon)N}{2+\epsilon}\log P + o(\log P). \quad (182)$$

Using (181), (182), (71), and (5), the s.d.o.f. is lower bounded as $D_s \geq \frac{(1-\epsilon)N}{2+\epsilon}$, and since $\epsilon > 0$ is arbitrarily small, the s.d.o.f. of $\frac{N}{2}$ is achievable for this case.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2738011, IEEE Transactions on Information Theory

I. Case 9: $N_e > N$, $N_e < N_c \le N + N_e$, and $N + N_c - N_e$ is even

In Sections V-G and V-H, we observe that the flat s.d.o.f. range extends to $N_c = N_e$, and not $N_c = N$ as in Sections V-B and V-C. Achieving the alignment of information and cooperative jamming at the eavesdropper requires that $N_c > N_e$ in order for the cooperative jammer to begin sending some jamming signals invisible to the legitimate receiver. For this case, in addition to choosing its precoding matrix jointly with the transmitter to satisfy the alignment condition, the cooperative jammer chooses its precoder to send $N_c - N_e$ jamming streams invisible to the receiver. The s.d.o.f. here is integer valued, for which we utilize Gaussian streams.

The transmitted signals are given by (72) with $d = l = \frac{N+N_c-N_e}{2}$, and $\mathbf{U}_t, \mathbf{V}_c$ are defined as in Section V-F. Let $\mathbf{P}_t = [\mathbf{P}_{t,1}\ \mathbf{P}_{t,2}]$, and $\mathbf{P}_c = [\mathbf{P}_{c,1}\ \mathbf{P}_{c,2}]$, where $\mathbf{P}_{t,1} \in \mathbb{C}^{N\times g}$, $\mathbf{P}_{t,2} \in \mathbb{C}^{N\times(N_c-N_e)}$, $\mathbf{P}_{c,1} \in \mathbb{C}^{N_c\times g}$, $\mathbf{P}_{c,2} \in \mathbb{C}^{N_c\times(N_c-N_e)}$, and $g = \frac{N_e+N-N_c}{2}$. The matrices $\mathbf{P}_t$ and $\mathbf{P}_c$ are chosen as follows. Let $\mathbf{G} = [\mathbf{G}_t\ -\mathbf{G}_c] \in \mathbb{C}^{N_e\times(N+N_c)}$, and let $\mathbf{G}' \in \mathbb{C}^{(N_e+N)\times(N+N_c)}$ be expressed as

$$\mathbf{G}' = \begin{bmatrix} \mathbf{G}_t & -\mathbf{G}_c \\ \mathbf{0}_{N\times N} & \mathbf{H}_c \end{bmatrix}. \tag{183}$$

Let $\mathbf{Q}' \in \mathbb{C}^{(N+N_c)\times(N_c-N_e)}$ be randomly chosen such that its columns span $\mathcal{N}(\mathbf{G}')$, and let the columns of the matrix $\mathbf{Q} \in \mathbb{C}^{(N+N_c)\times g}$ be randomly chosen as linearly independent vectors in $\mathcal{N}(\mathbf{G})$, and not in $\mathcal{N}(\mathbf{G}')$. Write the matrix $\mathbf{Q}$ as $\mathbf{Q} = \left[\mathbf{Q}_1^T\ \mathbf{Q}_2^T\right]^T$, and the matrix $\mathbf{Q}'$ as $\mathbf{Q}' = \left[\mathbf{Q}_1'^T\ \mathbf{Q}_2'^T\right]^T$, where $\mathbf{Q}_1 \in \mathbb{C}^{N\times g}$, $\mathbf{Q}_2 \in \mathbb{C}^{N_c\times g}$, $\mathbf{Q}_1' \in \mathbb{C}^{N\times(N_c-N_e)}$, and $\mathbf{Q}_2' \in \mathbb{C}^{N_c\times(N_c-N_e)}$. Set $\mathbf{P}_{t,1} = \mathbf{Q}_1$, $\mathbf{P}_{t,2} = \mathbf{Q}_1'$, $\mathbf{P}_{c,1} = \mathbf{Q}_2$, and $\mathbf{P}_{c,2} = \mathbf{Q}_2'$.

This choice of $\mathbf{P}_t$ and $\mathbf{P}_c$ results in $\mathbf{G}_t\mathbf{P}_t = \mathbf{G}_c\mathbf{P}_c$ and $\mathbf{H}_c\mathbf{P}_{c,2} = \mathbf{0}_{N\times(N_c-N_e)}$. Thus, the received signals at the receiver and eavesdropper are given by

$$\mathbf{Y}_r = \begin{bmatrix}\mathbf{H}_t\mathbf{P}_t & \mathbf{H}_c\mathbf{P}_{c,1}\end{bmatrix}\begin{bmatrix}\mathbf{U}_t \\ \mathbf{V}_{c1}^g\end{bmatrix} + \mathbf{Z}_r \tag{184}$$

$$\mathbf{Y}_e = \mathbf{G}_c\mathbf{P}_c(\mathbf{U}_t + \mathbf{V}_c) + \mathbf{Z}_e. \tag{185}$$

Using (185), and similar to going from (79) to (84), we have

$$I(\mathbf{X}_t; \mathbf{Y}_e) \le \frac{N + N_c - N_e}{2}. \tag{186}$$

Because of the assumption of randomly generated channel gains, each of $\mathbf{P}_t$ and $\mathbf{P}_c$ is full column rank a.s. Using Lemma 1, we have the matrix $[\mathbf{H}_t\mathbf{P}_t\ \ \mathbf{H}_c\mathbf{P}_{c,1}]$ is full column rank a.s., and hence, using (184), we have

$$I(\mathbf{X}_t; \mathbf{Y}_r) \ge \frac{N + N_c - N_e}{2}\log P + o(\log P). \tag{187}$$

Thus, using (186), (187), (71), and (5), the s.d.o.f. is lower bounded as $D_s \ge \frac{N+N_c-N_e}{2}$.

J. Case 10: $N_e > N$, $N_e < N_c \le N + N_e$, and $N + N_c - N_e$ is odd

The s.d.o.f. for this case is not an integer, and we have $N_c > N_e$, and hence, we utilize here precoding as in Section V-I, and signaling and decoding scheme as in Section V-H; $\mathbf{U}_t, \mathbf{V}_c$ are defined as in Section V-H, and $\mathbf{P}_t, \mathbf{P}_c$ are chosen as in Section V-I, with $d = \frac{N+N_c-N_e+1}{2}$ and $g = \frac{N_e+N-N_c+1}{2}$. Using the same decoding scheme as in Section V-H, we obtain that the s.d.o.f. is lower bounded as $D_s \ge \frac{N+N_c-N_e}{2}$ for this case, which completes the achievability proof of (70). Thus, we have completed the proof for Theorem 1.

VI. Extending to the General Case: Theorem 2

The converse and achievability proofs for Theorem 2 involve the same techniques as those utilized for Theorem 1. However, one needs to carefully handle the antenna configurations when $N_t \neq N_r$. In the following, we summarize how to extend the main ideas presented in Sections IV and V in order to prove Theorem 2.

A. Converse

The converse proof for Theorem 2 follows similar steps as in Section IV. In particular, we derive the following two upper bounds which hold for two different ranges of N_c.

1) $0 \le N_c \le N_e$: When $N_t \neq N_r$, the range of N_c for which the first upper bound holds is the same as in the case when $N_t = N_r = N$ in Section IV-A. However, unlike in Section IV-A, when $N_t \neq N_r$, this range of N_c is further subdivided into two ranges. The first upper bound on the s.d.o.f. we derive here is again $D_s \le [N_t + N_c - N_e]^+$, yet, the maximum s.d.o.f. for the channel is equal to $\min\{N_t, N_r\}$. Hence, for the case $N_r < N_t + N_c - N_e$, the maximum s.d.o.f., N_r, is reached at an N_c that is smaller than N_e. In particular, using similar analysis as in Section IV-A, we have

$$R_s \le C_s(P) = \rho\log P + o(\log P), \tag{188}$$

where, for $0 \le N_c \le [N_e - [N_t - N_r]^+]^+$, $\rho = [N_c + N_t - N_e]^+$. Since $[N_c + N_t - N_e]^+ \le N_r$ for $[N_e - [N_t - N_r]^+]^+ \le N_c \le N_e$, we have, for $0 \le N_c \le N_e$,

$$D_s \le \min\{N_r, [N_c + N_t - N_e]^+\}. \tag{189}$$

2) $N_r + [N_e - N_t]^+ < N_c \le 2\min\{N_t, N_r\} + N_e - N_t$: Following similar steps as in Section IV-B, where the two cases we consider here are $N_e \le N_t$ and $N_e > N_t$, the s.d.o.f. for this range of N_c is upper bounded as

$$D_s \le \frac{N_c + N_t - N_e}{2}. \tag{190}$$

It easy to see that, when $N_t = N_r = N$, the range of N_c for which the second upper bound in (190) holds is reduced to the range $\max\{N, N_e\} < N_c \le N + N_e$ in Section IV-B. However, when $N_t \neq N_r$, the range of N_c is different. In particular, we have that $N_c > N_r + [N_e - N_t]^+$ because, when $N_e > N_t$, (190) holds only when $N_c > N_r + N_e - N_t$ so that the number of antennas at the cooperative jammer in the modified channel, c.f. (59), is greater than N_r. We also have that $N_c \le 2\min\{N_t, N_r\} + N_e - N_t$. This this because, when $N_t < N_r$, we have $\frac{N_c+N_t-N_e}{2} = N_t$ at $N_c = N_t + N_e$, and when $N_t > N_r$, we have $\frac{N_c+N_t-N_e}{2} = N_r$ at $N_c = 2N_r + N_e - N_t$.

3) Obtaining the upper bound: For each of the following cases, we use the two bounds in (189) and (190) to obtain the upper bound for the s.d.o.f.

i) $N_t \geq N_r + N_e$
For this case, we use the trivial bound for the s.d.o.f., $D_s \leq N_r$ for all the values of N_c.

ii) $N_r \geq N_t \geq N_e$ and $N_r \geq N_t + N_e$
Using the bound in (189), we have

$$D_s \leq N_c + N_t - N_e, \text{ for } 0 \leq N_c \leq N_e,$$

where at $N_c = N_e$, we have $D_s \leq N_t$, which is the maximum achievable s.d.o.f. for this case.

iii) $N_t \geq N_e$ and $N_t - N_e < N_r < N_t + N_e$
Combining the bounds in (189) and (190), as in Section IV-C, yields

$$D_s \leq \begin{cases} N_c + N_t - N_e, & \text{if } 0 \leq N_c \leq \frac{N_r+N_e-N_t}{2} \\ \frac{N_r+N_t-N_e}{2}, & \text{if } \frac{N_r+N_e-N_t}{2} < N_c \leq N_r \\ \frac{N_c+N_t-N_e}{2}, & \\ \quad \text{if } N_r < N_c \leq 2\min\{N_t, N_r\} + N_e - N_t. & \end{cases} \tag{191}$$

iv) $N_e > N_t$ and $N_r \geq 2N_t$
Using the bound in (189), we have

$$D_s \leq [N_c + N_t - N_e]^+, \text{ for } 0 \leq N_c \leq N_e.$$

v) $N_e > N_t$ and $N_r < 2N_t$
By combining the bounds in (189) and (190), we have

$$D_s \leq \begin{cases} [N_c + N_t - N_e]^+, & \text{if } 0 \leq N_c \leq \frac{N_r}{2} + N_e - N_t \\ \frac{N_r}{2}, & \text{if } \frac{N_r}{2} + N_e - N_t < N_c \leq N_r + N_e - N_t \\ \frac{N_c+N_t-N_e}{2}, & \text{if } \begin{array}{c} N_r + N_e - N_t < N_c \leq \\ 2\min\{N_t, N_r\} + N_e - N_t \end{array} \end{cases}. \tag{192}$$

One can easily verify that the cases cited above cover all possible combinations of number of antennas at various terminals. By merging the upper bounds for these cases in one expression, we obtain (7) as the upper bound for the s.d.o.f. of the channel.

B. Achievability

The s.d.o.f. for the channel when N_t is not equal to N_r, given in (7), is achieved using techniques similar to what we presented in Section V. There are few cases, of the number of antennas, where the achievability is straightforward. One such case is when $N_t \geq N_r + N_e$, where the transmitter can send N_r Gaussian information streams invisible to the eavesdropper, and the maximum possible s.d.o.f. of the channel, i.e., N_r, is achieved without the help of the cooperative jammer, i.e., $N_c = 0$. Another case is when $N_r \geq N_t + \min\{N_t, N_e\}$, where the receiver's signal space is sufficient for decoding the information and jamming streams, at high SNR, for all $0 \leq N_c \leq N_e$, arriving at the s.d.o.f. of N_t (the maximum possible s.d.o.f.) at $N_c = N_e$. Thus, there is no constant period in the s.d.o.f. characterization for this case where the s.d.o.f. keeps increasing by increasing N_c, and Gaussian signaling and cooperative jamming are sufficient to achieve the s.d.o.f. of the channel.

We consider the five cases of the number of antennas at the different terminals listed in Section VI-A3. In the following, we summarize the achievable schemes for these cases. Let d and l denote the number of information and cooperative jamming streams. $\mathbf{P}_t, \mathbf{P}_c$ are the precoding matrices at the transmitter and the cooperative jammer.

i) $N_t \geq N_r + N_e$
The transmitter sends N_r Gaussian information streams over $\mathcal{N}(\mathbf{G}_t)$. $D_s = N_r$ is achievable at $N_c = 0$.

ii) $N_r \geq N_t \geq N_e$ and $N_r \geq N_t + N_e$
For $0 \leq N_c \leq N_e$, $d = N_c + N_t - N_e$ and $l = N_c$ Gaussian streams are transmitted. Choose $\mathbf{P}_t$ to send $N_t - N_e$ information streams over $\mathcal{N}(\mathbf{G}_t)$ and align the remaining information streams over cooperative jamming streams at the eavesdropper. $D_s = N_c + N_t - N_e$.

iii) $N_t \geq N_e$ and $N_t - N_e < N_r < N_t + N_e$:
1) For $0 \leq N_c \leq \frac{N_r+N_e-N_t}{2}$:
The same scheme as in case (ii) is utilized. $D_s = N_c + N_t - N_e$.
2) For $\frac{N_r+N_e-N_t}{2} < N_c \leq N_r$ and $N_r + N_t - N_e$ is even:
The same scheme as in case (iii-1), with $d = \frac{N_r+N_t-N_e}{2}$ and $l = \frac{N_r+N_e-N_t}{2}$, is utilized. The cooperative jammer uses only $\frac{N_r+N_e-N_t}{2}$ of its N_c antennas. $D_s = \frac{N_r+N_t-N_e}{2}$.
3) For $\frac{N_r+N_e-N_t}{2} < N_c \leq N_r$ and $N_r + N_t - N_e$ is odd:
$d = \frac{N_r+N_t-N_e+1}{2}$ and $l = \frac{N_r+N_e-N_t+1}{2}$ structured streams, as defined in Section V-C, are transmitted. The cooperative jammer uses only $\frac{N_r+N_e-N_t+1}{2}$ of its N_c antennas. $\mathbf{P}_t$ is chosen as in case (ii). The legitimate receiver uses the projection and cancellation technique, as in Section V-C. $D_s = \frac{N_r+N_t-N_e}{2}$.
4) For $N_r < N_c \leq 2\min\{N_t, N_r\} + N_e - N_t$ and $N_c + N_t - N_e$ is even:
$d = \frac{N_c+N_t-N_e}{2}$ and $l = \frac{N_c+N_e-N_t}{2}$ Gaussian streams are transmitted. The cooperative jammer chooses $\mathbf{P}_c$ to send $N_c - N_r$ cooperative jamming streams over $\mathcal{N}(\mathbf{H}_c)$. $\mathbf{P}_t$ is chosen as in case (ii). $D_s = \frac{N_c+N_t-N_e}{2}$.
5) For $N_r < N_c \leq 2\min\{N_t, N_r\} + N_e - N_t$ and $N_c + N_t - N_e$ is odd:
$d = \frac{N_c+N_t-N_e+1}{2}$ and $l = \frac{N_c+N_e-N_t+1}{2}$ structured streams are transmitted. $\mathbf{P}_t, \mathbf{P}_c$ are chosen as in case (iii-4). The legitimate receiver uses the projection and cancellation technique. $D_s = \frac{N_c+N_t-N_e}{2}$.

iv) $N_e > N_t$ and $N_r \geq 2N_t$
For $0 \leq N_c \leq N_e$, $d = l = [N_c + N_t - N_e]^+$ Gaussian streams are transmitted. Both $\mathbf{P}_t, \mathbf{P}_c$ are chosen to align the information streams over the cooperative jamming streams at the eavesdropper as in Section V-F. $D_s = [N_c + N_t - N_e]^+$.

v) $N_e > N_t$ and $N_r < 2N_t$:
1) For $0 \leq N_c \leq \frac{N_r}{2} + N_e - N_t$:
The same scheme as in case (iv) is utilized. $D_s = [N_c + N_t - N_e]^+$.
2) For $\frac{N_r}{2} + N_e - N_t < N_c \leq N_r + N_e - N_t$ and N_r is

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2738011, IEEE Transactions on Information Theory

even:
$d = l = \frac{N_r}{2}$ Gaussian streams are transmitted. $\mathbf{P}_t, \mathbf{P}_c$ are chosen as in case (iv). $D_s = \frac{N_r}{2}$.

3) For $\frac{N_r}{2} + N_e - N_t < N_c \leq N_r + N_e - N_t$ and N_r is odd:
$d = l = \frac{N_r+1}{2}$ structured streams are transmitted. $\mathbf{P}_t, \mathbf{P}_c$ are as in case (iv). The legitimate receiver uses the projection and cancellation technique. $D_s = \frac{N_r}{2}$.

4) For $N_r + N_e - N_t < N_c \leq 2\min\{N_t, N_r\} + N_e - N_t$ and $N_c + N_t - N_e$ is even:
$d = l = \frac{N_c+N_t-N_e}{2}$ Gaussian streams are transmitted. Both $\mathbf{P}_t, \mathbf{P}_c$ are chosen to align the information and the cooperative jamming streams at the eavesdropper. $\mathbf{P}_c$ is also chosen to send $N_c - N_r$ cooperative jamming streams over $\mathcal{N}(\mathbf{H}_c)$ as in Section V-I. $N_c > N_r + N_e - N_t$ achieves the above two conditions. $D_s = \frac{N_c+N_t-N_e}{2}$.

5) For $N_r + N_e - N_t < N_c \leq 2\min\{N_t, N_r\} + N_e - N_t$ and $N_c + N_t - N_e$ is odd:
$d = l = \frac{N_c+N_t-N_e+1}{2}$ structured streams are transmitted. $\mathbf{P}_t, \mathbf{P}_c$ are chosen as in case (v-4). The receiver uses the projection and cancellation technique. $D_s = \frac{N_c+N_t-N_e}{2}$.

Using the achievable schemes described above for the different cases of the number of antennas, and their analysis as in Section V, we have (7) as the achievable s.d.o.f., which completes the proof for Theorem 2.

VII. Discussion

At this point, it is useful to discuss the results and the implications of this work. Theorem 1, c.f. (6), shows the behavior of the s.d.o.f., for an $(N \times N \times N_e)$ multi-antenna Gaussian wire-tap channel with an N_c-antenna cooperative jammer, associated with increasing N_c form 0 to $N + N_e$. The s.d.o.f. first increases linearly by increasing N_c from 0 to $N_e - \lceil \frac{\min\{N,N_e\}}{2} \rceil$, that is to say adding one antenna at the cooperative jammer provided the system to have one additional degrees of freedom. The s.d.o.f. remains constant in the N_c range of $N_e - \lfloor \frac{\min\{N,N_e\}}{2} \rfloor$ to $\max\{N, N_e\}$, and starts to increase again for N_c from $\max\{N, N_e\}$ to $N + N_e$, until the s.d.o.f. arrives at its maximum value, N, at $N_c = N + N_e$. This behavior transpires both when the eavesdropper antennas are fewer or more than that of the legitimate receiver.

The reason for the flat s.d.o.f. range is as follows: At high SNR, achieving the secrecy constraint requires i) the entropy of the cooperative jamming signal, $\mathbf{X}_c^n$, to be greater than or equal to that of the information signal visible to the eavesdropper, and ii) $\mathbf{X}_c^n$ to completely cover the information signal, $\mathbf{X}_t^n$, at the eavesdropper. For $N_e \leq N$, part of $\mathbf{X}_t^n$ can be sent invisible to the eavesdropper, and the information signal visible to the eavesdropper can be covered by jamming for all N_c. For $0 \leq N_c \leq \frac{N_e}{2}$, the spatial resources at the receiver are sufficient, at high SNR, for decoding information and jamming signals which satisfy the above constraints. Thus, increasing the possible entropy of $\mathbf{X}_c^n$ by increasing N_c from 0 to $\lfloor \frac{N_e}{2} \rfloor$ allows for increasing the entropy of $\mathbf{X}_t^n$, and hence, the achievable secrecy rate and the s.d.o.f. increase. At $N_c = \lceil \frac{N_e}{2} \rceil$, the possible entropy of $\mathbf{X}_c^n$ and, correspondingly, the maximum possible entropy of $\mathbf{X}_t^n$, result in information and jamming signals which completely occupy the receiver's signal space. Thus, increasing the possible uncertainty of $\mathbf{X}_c^n$ by increasing N_c from $\lceil \frac{N_e}{2} \rceil$ to N is useless, since, in this range, $\mathbf{X}_c^n$ is totally observed by the receiver which has its signal space already full at $N_c = \lceil \frac{N_e}{2} \rceil$.

Increasing N_c over N increases the possible entropy of $\mathbf{X}_c^n$ and simultaneously increases the part of $\mathbf{X}_c^n$ that can be transmitted invisible to the receiver, leaving more space for $\mathbf{X}_t^n$ at the receiver. This allows for increasing the secrecy rate, and hence, the s.d.o.f. starts to increase again. For $N_e > N$, the s.d.o.f. is equal to zero for all $0 \leq N_c \leq N_e - N$, where $\mathbf{X}_c^n$ cannot cover the information at the eavesdropper for this case. The s.d.o.f. starts to increase again, after the flat range, at $N_c > N_e$, since sending jamming signals invisible to the receiver while satisfying the covering condition at the eavesdropper requires that $N_c > N_e$.

The difference in the slope for the increase in the s.d.o.f. in the ranges before and after the flat range, for both $N_e \leq N$ and $N_e > N$, can be explained as follows. For $0 \leq N_c \leq N_e - \frac{\min\{N,N_e\}}{2}$, each additional antenna at the cooperative jammer allows for utilizing two more spatial dimensions at the receiver; one spatial dimension is used for the jamming signal and the other is used for the information signal. By contrast, for $\max\{N, N_e\} < N_c \leq N + N_e$, each additional antenna at the cooperative jammer sets one spatial dimension at the receiver free from jamming, and this spatial dimension is shared between the extra cooperative jamming and information streams.

It is important to note that the result that suggests that increasing the cooperative jammer antennas is not useful in the range $N_e - \frac{\min\{N,N_e\}}{2} < N_c \leq \max\{N, N_e\}$ applies only to the prelog of the secrecy capacity, i.e., is specific to the high SNR behavior. This should not be taken to mean that additional antennas do not improve secrecy rate, but only the secrecy rate scaling with power in the high SNR.

Theorem 2 generalizes the results above to the case where the number of transmit antennas at the transmitter, N_t, is not equal to the number of receive antennas at the legitimate receiver, N_r. Although the maximum possible s.d.o.f. of the channel for this case is limited to $\min\{N_t, N_r\} = N_d$, increasing N_t over N_r, or increasing N_r over N_t, do change the behavior of the s.d.o.f. associated with increasing N_c until the maximum possible s.d.o.f. is reached. Let us start at $N_t = N_r = N_d$. For $N_t \geq N_e$, increasing N_t over $N_d = N_r$ increases the number of the information streams that can be sent invisible to the eavesdropper, and hence the s.d.o.f. without the help of the cooperative jammer, i.e., $N_c = 0$, increases. This results in increasing the range of N_c for which the s.d.o.f. remains constant by increasing N_c, since the receiver's signal space gets full at a smaller N_c and remains full until N_c is larger than $N_d = N_r$. In addition, increasing N_t over N_d, when $N_t \geq N_e$, results in decreasing the value of N_c at which the maximum s.d.o.f. of the channel, N_d, is achievable, arriving at $N_t \geq N_r + N_e$, where the s.d.o.f. of N_d is achievable without the help of the cooperative jammer. Fig. 5 illustrates this behavior. When $N_e > N_t$, increasing N_t over

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2738011, IEEE Transactions on Information Theory

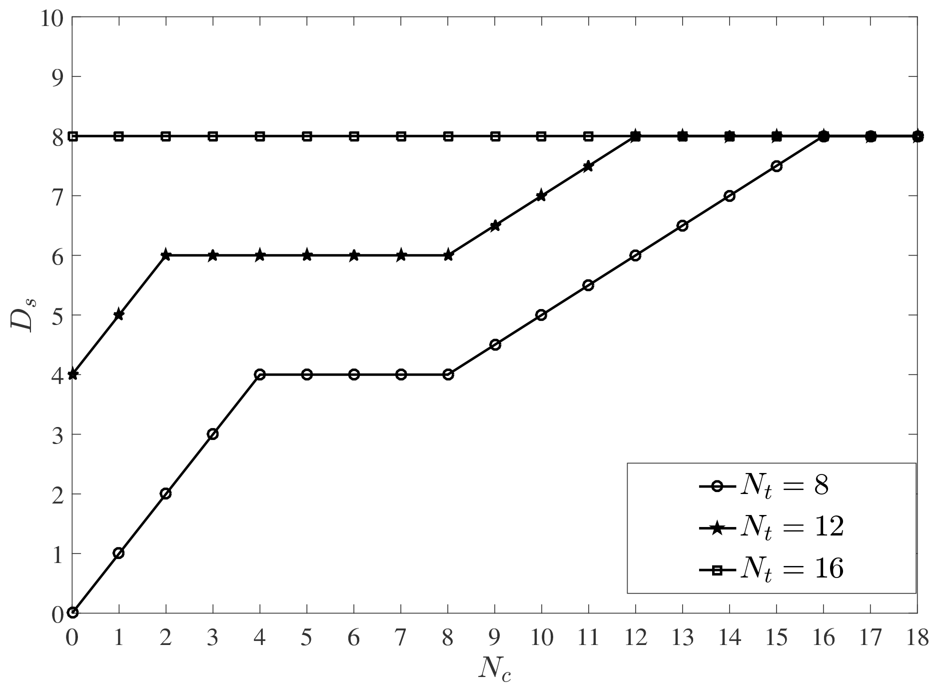


Fig. 5: D_s versus N_c when $N_r = N_e = 8$ and N_t increases from N_r to $N_r + N_e$.

N_d decreases the value of N_c at which the s.d.o.f. is positive, and decreases the value of N_c at which the s.d.o.f. of N_d is achievable, arriving at $N_t > N_e$, where the channel renders itself to the previous case. This behavior is demonstrated in Fig. 6. For both the cases $N_t \geq N_e$ and $N_t < N_e$, increasing N_r over $N_d = N_t$, results in increasing the available space at the receiver's signal space, and hence the constant period decreases, arriving at $N_r \geq N_t + N_e$ when $N_t \geq N_e$, or at $N_r \geq 2N_t$ when $N_e > N_t$, where the constant period vanishes. Fig. 7 illustrates the behavior of the s.d.o.f. curve associated with increasing N_r over N_t.

VIII. Conclusion

In this paper, we have studied the multi-antenna wire-tap channel with a N_c-antenna cooperative jammer, N_t-antenna transmitter, N_r-antenna receiver, and N_e-antenna eavesdropper. We have completely characterized the s.d.o.f. for this channel for all possible values of the number of antennas at the cooperative jammer, N_c. We have shown that when the s.d.o.f. of the channel is integer valued, it can be achieved by linear precoding at the transmitter and cooperative jammer, Gaussian signaling both for transmission and jamming, and linear processing at the legitimate receiver. By contrast, when the s.d.o.f. is not an integer, we have shown that a scheme which employs structured signaling both at the transmitter and the cooperative jammer, along with joint signal space and signal scale alignment achieves the s.d.o.f. of the channel. We have seen that, when $N_t \geq N_e$, the transmitter uses its precoder to send a part of its information signal invisible to the eavesdropper, and to align the remaining part over jamming at the eavesdropper, while the cooperative jammer uses its precoder to send a part of its jamming signal invisible to the receiver, whenever possible. When $N_e > N_t$, more intricate precoding at the transmitter and cooperative jammer is required, where both the transmitter and cooperative jammer choose their precoders to achieve the alignment of information and jamming at the eavesdropper, and simultaneously, the cooperative jammer designs its precoder, whenever possible, to send a part of the jamming signal invisible to the receiver.

The converse was established by allowing for full cooperation between the transmitter and cooperative jammer for a certain range of N_c, and by incorporating both the secrecy and reliability constraints, for the other values of N_c. We note that while this paper settles the degrees of freedom of this channel, its secrecy capacity is still open. Additionally, while the model considered here assumes channels to be known, universal secrecy as in [32] should be considered in the future.

Appendix A
Choice of $\mathbf{K}_t$ and $\mathbf{K}_c$

The covariance matrices $\mathbf{K}_t$ and $\mathbf{K}_c$ are chosen so that they are *positive definite*, i.e., $\mathbf{K}_t, \mathbf{K}_c \succ \mathbf{0}$, and hence non-singular, in order to guarantee the finiteness of $h(\tilde{\mathbf{Z}}_t)$ and $h(\tilde{\mathbf{Z}}_c)$ in (26). In addition, positive definite $\mathbf{K}_t$ and $\mathbf{K}_c$ result in positive definite $\mathbf{\Sigma}_{\tilde{\mathbf{Z}}_1}$ and $\mathbf{\Sigma}_{\tilde{\mathbf{Z}}_2}$, and hence, $h(\tilde{\mathbf{Z}}_1)$ and $h(\tilde{\mathbf{Z}}_2)$ in (28) are also finite.

For $\mathbf{I}_{N_e} - \mathbf{G}_t\mathbf{K}_t\mathbf{G}_t^H$ to be a valid covariance matrix for $\tilde{\mathbf{Z}}_e$ in (30), $\mathbf{K}_t$ has to satisfy $\mathbf{G}_t\mathbf{K}_t\mathbf{G}_t^H \preceq \mathbf{I}_{N_e}$, which is equivalent to

$$||\mathbf{K}_t^{\frac{1}{2}}\mathbf{G}_t^H|| \leq 1. \tag{193}$$

Recall that $||\mathbf{K}_t^{\frac{1}{2}}\mathbf{G}_t^H||$ is the induced norm for the matrix $\mathbf{K}_t^{\frac{1}{2}}\mathbf{G}_t^H$.

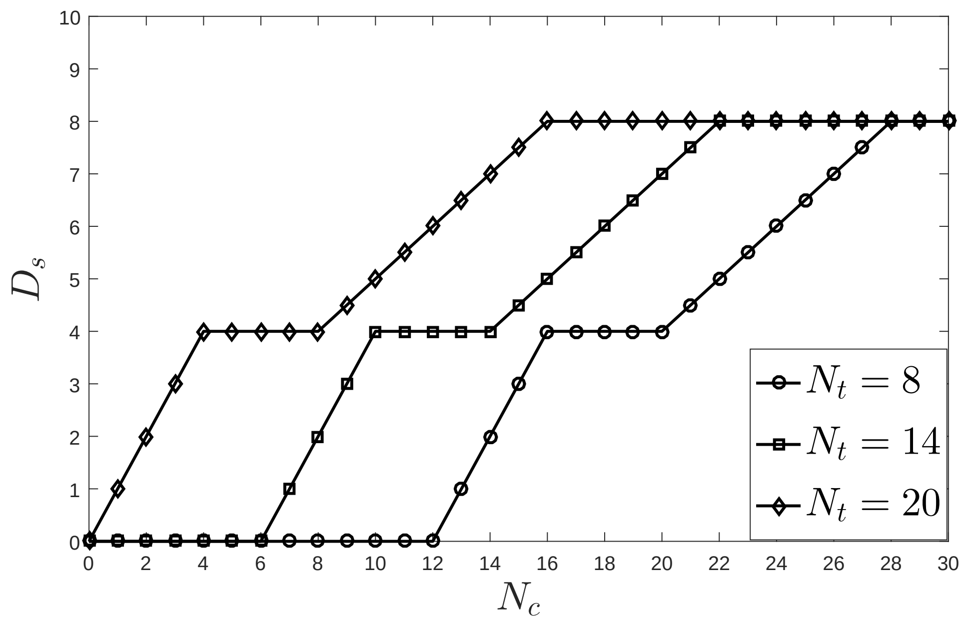


Fig. 6: D_s versus N_c when $N_r = 8$, $N_e = 20$ and N_t increases from N_r to N_e.

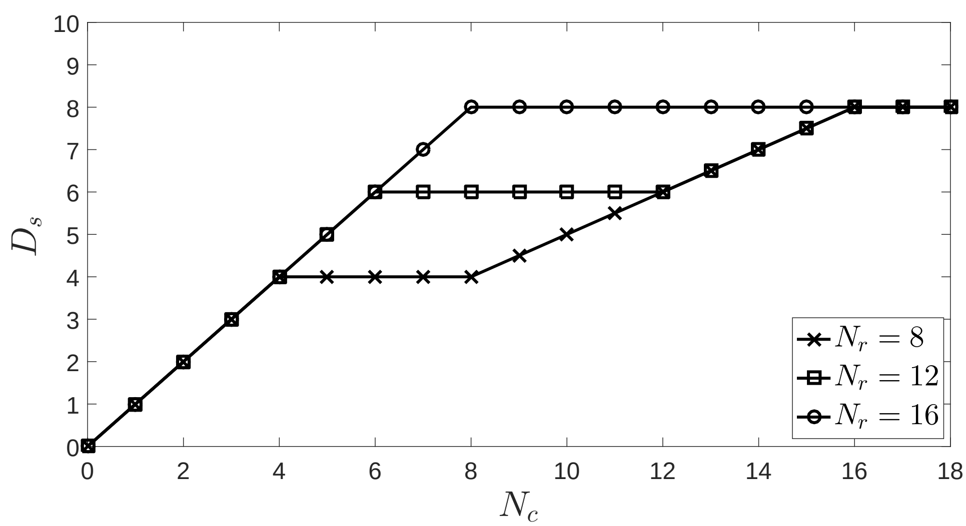


Fig. 7: D_s versus N_c when $N_t = N_e = 8$ and N_r increases from N_t to $N_t + N_e$.

Similarly, for $\mathbf{I}_N - \mathbf{H}_c\mathbf{K}_c\mathbf{H}_c^H$, $\mathbf{I}_{N_e} - \mathbf{G}_t\mathbf{K}_t\mathbf{G}_t^H - \mathbf{G}_c\mathbf{K}_c\mathbf{G}_c^H$, and $\mathbf{I}_N - \mathbf{H}'_{c_2}\mathbf{K}'_c\mathbf{H}'^H_{c_2}$ to be valid covariance matrices for $\tilde{\mathbf{Z}}_r, \tilde{\mathbf{Z}}'_e$, and $\tilde{\mathbf{Z}}'_r$, in (40), (53), (62), $\mathbf{K}_t, \mathbf{K}_c, \mathbf{K}'_c$ have to satisfy

$$\begin{aligned}&||\mathbf{K}_c^{\frac{1}{2}}\mathbf{H}_c^H|| \leq 1, \quad ||\mathbf{K}_t^{\frac{1}{2}}\mathbf{G}_t^H||^2 + ||\mathbf{K}_c^{\frac{1}{2}}\mathbf{G}_c^H||^2 \leq 1,\\ &\text{and } ||\mathbf{K}'^{\frac{1}{2}}_c\mathbf{H}'^H_{c_2}|| \leq 1.\end{aligned} \tag{194}$$

In order to satisfy the conditions (193) and (194), we choose $\mathbf{K}_t = \rho^2\mathbf{I}_N$, $\mathbf{K}_c = \rho^2\mathbf{I}_K$, where

$$0 < \rho \leq \frac{1}{\max\left\{\begin{array}{c}||\mathbf{G}_t^H||, ||\mathbf{H}_c^H||, ||\mathbf{H}'^H_{c_2}||,\\ \sqrt{||\mathbf{G}_t^H||^2 + ||\mathbf{G}_c^H||^2}\end{array}\right\}} \tag{195}$$

$$= \frac{1}{\max\left\{||\mathbf{H}_c^H||, \sqrt{||\mathbf{G}_t^H||^2 + ||\mathbf{G}_c^H||^2}\right\}}. \tag{196}$$

Appendix B
Derivation of (48),(49), and (66)

In order to upper bound $h(Y_{r,k}(i))$, for all $i = 1, 2, \cdots, n$ and $k = 1, 2, \cdots, N$, we first upper bound the variance of $Y_{r,k}(i)$, denoted by $\mathrm{Var}(Y_{r,k}(i))$. Let $\mathbf{h}_{t,k}^r$ and $\mathbf{h}_{c,k}^r$ denote the transpose of the kth row vectors of $\mathbf{H}_t$ and $\mathbf{H}_c$, respectively. Let $\mathbf{Z}_r(i) = [Z_{r,1}(i) \cdots Z_{r,N}(i)]^T$. Using (1), $Y_{r,k}(i)$ is expressed as

$$Y_{r,k}(i) = \mathbf{h}_{t,k}^{r^T}\mathbf{X}_t(i) + \mathbf{h}_{c,k}^{r^T}\mathbf{X}_c(i) + Z_{r,k}(i). \tag{197}$$

Thus, $\mathrm{Var}\left(Y_{r,k}(i)\right)$ can be bounded as

$$\mathrm{Var}\left(Y_{r,k}(i)\right) \leq \mathrm{E}\left(Y_{r,k}(i)Y^*_{r,k}(i)\right) \tag{198}$$

$$= \mathrm{E}\left(|\mathbf{h}^{r^T}_{t,k}\mathbf{X}_t(i)|^2\right) + \mathrm{E}\left(|\mathbf{h}^{r^T}_{c,k}\mathbf{X}_c(i)|^2\right) + \mathrm{E}\left(|Z_{r,k}(i)|^2\right) \tag{199}$$

$$\leq ||\mathbf{h}^r_{t,k}||^2\, \mathrm{E}\left(||\mathbf{X}_t(i)||^2\right) + ||\mathbf{h}^r_{c,k}||^2\, \mathrm{E}\left(||\mathbf{X}_c(i)||^2\right) + 1 \tag{200}$$

$$\leq 1 + \left(||\mathbf{h}^r_{t,k}||^2 + ||\mathbf{h}^r_{c,k}||^2\right) P, \tag{201}$$

where (200) follows from Cauchy-Schwarz inequality and monotonicity of expectation, and (201) follows from the power constraints at the transmitter and cooperative jammer.

Define $h^2 = \max_k \left(||\mathbf{h}^r_{t,k}||^2 + ||\mathbf{h}^r_{c,k}||^2\right)$. Since $h(Y_{r,k}(i))$ is upper bounded by the entropy of a complex Gaussian random variable with the same variance, we have, for all $i = 1, 2, \cdots, n$ and $k = 1, 2, \cdots, N$,

$$h(Y_{r,k}(i)) \leq \log 2\pi e \left(1 + \left(||\mathbf{h}^r_{t,k}||^2 + ||\mathbf{h}^r_{c,k}||^2\right) P\right) \tag{202}$$

$$\leq \log 2\pi e + \log(1 + h^2 P). \tag{203}$$

Similarly, we have

$$\bar{Y}_{r,k}(i) = \mathbf{h}^{r^T}_{t,k}\mathbf{X}_t(i) + \mathbf{h}'^{r^T}_{c,k}\mathbf{X}'_{c_2}(i) + Z_{r,k}(i), \tag{204}$$

where $\mathbf{h}'^r_{c,k}$ is the transpose of the k-th row vector of $\mathbf{H}'_{c_2}$. Thus, we have,

$$h(\bar{Y}_{r,k}(i)) \leq \log 2\pi e + \log(1 + \bar{h}^2 P), \tag{205}$$

where $\bar{h}^2 = \max_k \left(||\mathbf{h}^r_{t,k}||^2 + ||\mathbf{h}'^r_{c,k}||^2\right)$.

Next, we upper bound $h(\tilde{X}_{t,k}(i))$. The power constraint at the transmitter, for $i = 1, 2, \cdots, n$, is $\mathrm{E}\left(\mathbf{X}^H_t(i)\,\mathbf{X}_t(i)\right) = \sum_{k=1}^N \mathrm{E}\left(|X_{t,k}(i)|^2\right) \leq P$. Thus, $\mathrm{E}\left(|X_{t,k}(i)|^2\right) \leq P$ for all $i = 1, 2, \cdots, n$, and $k = 1, 2, \cdots, N$. Recall that $\tilde{X}_{t,k}(i) = X_{t,k}(i) + \tilde{Z}_{t,k}(i)$, where $X_{t,k}(i)$ and $\tilde{Z}_{t,k}(i)$ are independent, and the covariance matrix of $\tilde{\mathbf{Z}}_t$ is $\mathbf{K}_t = \rho^2\mathbf{I}_N$, where $0 < \rho \leq \min\left\{\frac{1}{||\mathbf{H}^H_c||}, \frac{1}{\sqrt{||\mathbf{G}^H_t||^2 + ||\mathbf{G}^H_c||^2}}\right\}$. Thus, $\mathrm{Var}\left(\tilde{X}_{t,k}(i)\right)$ is upper bounded as

$$\mathrm{Var}\left(\tilde{X}_{t,k}(i)\right) = \mathrm{Var}\left(X_{t,k}(i)\right) + \mathrm{Var}\left(\tilde{Z}_{t,k}(i)\right) \tag{206}$$

$$\leq \mathrm{E}\left(|X_{t,k}(i)|^2\right) + \rho^2 \leq P + \rho^2. \tag{207}$$

Thus, for $i = 1, 2, \cdots, n$ and $k = 1, 2, \cdots, N$, we have

$$h(\tilde{X}_{t,k}(i)) \leq \log 2\pi e + \log(\rho^2 + P). \tag{208}$$

Similarly, using the power constraint at the cooperative jammer, we have, for $i = 1, \cdots, n$ and $m = 1, \cdots, K$,

$$h(\tilde{X}_{c,m}(i)) \leq \log 2\pi e + \log(\rho^2 + P). \tag{209}$$

Appendix C
Proof of Lemma 1

Consider two matrices $\mathbf{Q} \in \mathbb{C}^{M\times K}$ and $\mathbf{W} \in \mathbb{C}^{K\times N}$ such that $\mathbf{Q}$ is full row-rank and $\mathbf{W}$ has all of its entries independently drawn from a continuous distribution, where $K > N, M$. Let $L = \min\{N, M\}$. We show that $\mathbf{QW}$ has a rank L a.s. The matrices $\mathbf{Q}$ and $\mathbf{W}$ can be written as

$$\mathbf{Q} = \begin{bmatrix}\mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_K\end{bmatrix}, \tag{210}$$

$$\mathbf{W} = \begin{bmatrix}\mathbf{w}_1 & \mathbf{w}_2 & \cdots & \mathbf{w}_N\end{bmatrix}, \tag{211}$$

where $\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_K$ are the K length-M column vectors of $\mathbf{Q}$, and $\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_N$ are the N length-K column vectors of $\mathbf{W}$.

Let $w_{m,i}$ denotes the entry in the mth row and ith column of $\mathbf{W}$. Let $\mathbf{QW} = [\mathbf{s}_1\ \mathbf{s}_2\ \cdots\ \mathbf{s}_N]$, where $\mathbf{s}_i$ is a length-M column vector, $i = 1, 2, \cdots, N$. When $M \geq N$, $\mathbf{QW} = [\mathbf{s}_1\ \mathbf{s}_2\ \cdots\ \mathbf{s}_L]$, and when $M < N$, $\{\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_L\}$ are the first L columns of $\mathbf{QW}$. In order to show that the matrix $\mathbf{QW}$ has rank L, we show that, in either case, $\{\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_L\}$ are a.s. linearly independent, i.e.,

$$\sum_{i=1}^{L} \lambda_i \mathbf{s}_i = \mathbf{0}_{M\times 1} \tag{212}$$

if and only if $\lambda_i = 0$ for all $i = 1, 2, \cdots, L$.

Each $\mathbf{s}_i$, for $i = 1, 2, \cdots, L$, can be viewed as a linear combination of the K columns of $\mathbf{Q}$ with coefficients that are the entries of the ith column of $\mathbf{W}$, i.e.,

$$\mathbf{s}_i = \sum_{m=1}^{K} w_{m,i}\mathbf{q}_m. \tag{213}$$

Using (213), we can rewrite (212) as

$$\sum_{m=1}^{K} \varphi_m \mathbf{q}_m = \mathbf{0}_{M\times 1} \tag{214}$$

where, for $m = 1, 2, \cdots, K$,

$$\varphi_m = \sum_{i=1}^{L} \lambda_i w_{m,i}. \tag{215}$$

The K columns of $\mathbf{Q}$ are linearly dependent since each of them is of length M and $K > M$. Thus, equation (214) has infinitely many solutions for $\{\varphi_m\}_{m=1}^K$.

Each of these solutions for φ_m's constitutes a system of K linear equations $\left\{\varphi_m = \sum_{i=1}^L \lambda_i w_{m,i}, m = 1, 2, \cdots, K\right\}$. The number of unknowns in this system, i.e. λ's, is L. Since the number of equations in this system, K, is greater than the number of unknowns, L, this system has a solution for $\{\lambda_i\}_{i=1}^L$ only if the elements $\{w_{m,i} : m = 1, 2, \cdots, K,$ and $i = 1, 2, \cdots, L\}$ are dependent. Since the entries of $\mathbf{W}$ are all randomly and independently drawn from some continuous distribution, the probability that these entries are dependent is zero.

Moreover, consider the set with infinite cardinality, where each element in this set is a structured $\mathbf{W}$ that causes the system of equations in (215) to have a solution for $\{\lambda_i\}_{i=1}^L$ for one of the infinitely many solutions of $\{\varphi_m\}_{m=1}^K$ to (214). This set with infinite cardinality has a measure zero in the space $\mathbb{C}^{K\times L}$, since this set is a subspace of $\mathbb{C}^{K\times L}$ with a dimension strictly less than $K \times L$. We conclude that (212) a.s. has no non-zero solution for $\{\lambda_i\}_{i=1}^L$. Thus, $\mathbf{QW}$ has rank L a.s.

If $\mathbf{QW}$ has rank L a.s. , then so does $(\mathbf{QW})^T = \mathbf{W}^T\mathbf{Q}^T$.

Setting $\mathbf{E}_1 = \mathbf{W}^T$ and $\mathbf{E}_2 = \mathbf{Q}^T$, we have $\mathbf{E}_1 \in \mathbb{C}^{N\times K}$ has all of its entries independently drawn from some continuous distribution, $\mathbf{E}_2 \in \mathbb{C}^{K\times M}$ is full column-rank, $K > N, M$, and $\mathbf{E}_1\mathbf{E}_2$ has rank $L = \min\{N, M\}$ a.s. Thus, Lemma 1 is proved.

APPENDIX D
DERIVATION OF (89) AND (90)

The power constraints at the transmitter and cooperative jammer are $\mathrm{E}\left(\mathbf{X}_t^H\mathbf{X}_t\right) \le P$ and $\mathrm{E}\left(\mathbf{X}_c^H\mathbf{X}_c\right) \le P$. Using (72), we have

$$\mathrm{E}\left(\mathbf{X}_t^H\mathbf{X}_t\right) = \mathrm{E}\left(\mathbf{U}_t^H\mathbf{P}_t^H\mathbf{P}_t\mathbf{U}_t\right) \tag{216}$$

$$= \sum_{i=1}^{d}\sum_{m=1}^{d} \mathbf{p}_{t,m}^H\mathbf{p}_{t,i}\mathrm{E}\left(U_m^* U_i\right) \tag{217}$$

$$= \sum_{i=1}^{d} ||\mathbf{p}_{t,i}||^2\mathrm{E}\left(|U_i|^2\right) \tag{218}$$

$$= ||\mathbf{p}_{t,1}||^2\mathrm{E}\left(|U_1|^2\right) + \sum_{i=2}^{d} ||\mathbf{p}_{t,i}||^2\left(\mathrm{E}\left(U_{i,\mathrm{Re}}^2\right) + \mathrm{E}\left(U_{i,\mathrm{Im}}^2\right)\right) \tag{219}$$

$$\le \left(||\mathbf{p}_{t,1}||^2 + 2\sum_{i=2}^{d} ||\mathbf{p}_{t,i}||^2\right) a^2Q^2, \tag{220}$$

where (218) follows since U_i and U_m, for $i \neq m$, are independent, and (220) follows since $\mathrm{E}\left(U_1^2\right)$, $\mathrm{E}\left(U_{i,\mathrm{Re}}^2\right)$, $\mathrm{E}\left(U_{i,\mathrm{Im}}^2\right) \le a^2Q^2$, for $i = 2, 3, \cdots, d$.

Similarly, using (72) and (87), we have

$$\mathrm{E}\left(\mathbf{X}_c^H\mathbf{X}_c\right) = \mathrm{E}\left(\mathbf{V}_c^H\mathbf{P}_c^H\mathbf{P}_c\mathbf{V}_c\right) = \sum_{i=1}^{l}\mathrm{E}\left(|V_i|^2\right) \tag{221}$$

$$= \mathrm{E}\left(V_1^2\right) + \sum_{i=2}^{l}\left(\mathrm{E}\left(V_{i,\mathrm{Re}}^2\right) + \mathrm{E}\left(V_{i,\mathrm{Im}}^2\right)\right) \tag{222}$$

$$\le (2l-1)a^2Q^2. \tag{223}$$

From (220) and (223), in order to satisfy the power constraints, we need that

$$a^2Q^2 \le \gamma^2 P, \tag{224}$$

where,

$$\gamma^2 = \frac{1}{\max\left\{2l-1, ||\mathbf{p}_{t,1}||^2 + 2\sum_{i=2}^{d}||\mathbf{p}_{t,i}||^2\right\}}. \tag{225}$$

Let us choose the integer Q as

$$Q = \left\lfloor P^{\frac{1-\epsilon}{2+\epsilon}}\right\rfloor = P^{\frac{1-\epsilon}{2+\epsilon}} - \nu, \tag{226}$$

where ν is a constant which does not depend on the power P. Thus,

$$a = \gamma P^{\frac{3\epsilon}{2(2+\epsilon)}}, \tag{227}$$

satisfies the condition in (224). Thus, the power constraints at the transmitter and cooperative jammer are satisfied.

REFERENCES

[1] M. Nafea and A. Yener, "Secure degrees of freedom for the MIMO wiretap channel with a multiantenna cooperative jammer," *IEEE Information Theory Workshop*, pp. 626–630, November 2014.

[2] ——, "Secure degrees of freedom of $N\times N\times M$ wiretap channel with a K-antenna cooperative jammer," *IEEE International Conference on Communications*, pp. 4169–4174, June 2015.

[3] A. D. Wyner, "The wire-tap channel," *Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1387, 1975.

[4] I. Csiszar and J. Korner, "Broadcast channels with confidential messages," *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 339–3487, 1978.

[5] S. Leung-Yan-Cheong and M. Hellman, "The Gaussian wire-tap channel," *IEEE Transactions on Information Theory*, vol. 24, no. 4, pp. 451–456, 1978.

[6] R. Liu, I. Maric, P. Spasojevic, and R. D. Yates, "Discrete memoryless interference and broadcast channels with confidential messages: Secrecy rate regions," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2493–2507, 2008.

[7] R. Liu and H. V. Poor, "Secrecy capacity region of a multiple-antenna Gaussian broadcast channel with confidential messages," *IEEE Transactions on Information Theory*, vol. 55, no. 3, pp. 1235–1249, 2009.

[8] E. Ekrem and S. Ulukus, "The secrecy capacity region of the Gaussian MIMO multi-receiver wiretap channel," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 2083–2114, 2011.

[9] A. Khandani, G. Bagherikaram, and A. Motahary, "The secrecy capacity region of the Gaussian MIMO broadcast channel," *IEEE Transactions on Information Theory*, vol. 59, no. 5, pp. 2673–2682, 2013.

[10] E. Tekin, S. Serbetli, and A. Yener, "On secure signaling for the Gaussian multiple access wire-tap channel," *Asilomar Conference on Signals, Systems, and Computers*, pp. 1747–1751, November 2005.

[11] E. Tekin and A. Yener, "The general Gaussian multiple-access and two-way wiretap channels: Achievable rates and cooperative jamming," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2735–2751, 2008.

[12] Y. Liang and H. V. Poor, "Multiple-access channels with confidential messages," *IEEE Transactions on Information Theory*, vol. 54, no. 3, pp. 976–1002, 2008.

[13] M. R. Bloch and J. N. Laneman, "Strong secrecy from channel resolvability," *IEEE Transactions on Information Theory*, vol. 59, no. 12, pp. 8077–8098, 2013.

[14] P. K. Gopala, L. Lai, and H. El Gamal, "On the secrecy capacity of fading channels," *IEEE Transactions on Information Theory*, vol. 54, no. 10, pp. 4687–4698, 2008.

[15] A. Khisti, A. Tchamkerten, and G. W. Wornell, "Secure broadcasting over fading channels," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2453–2469, 2008.

[16] L. Lai and H. El Gamal, "The relay-eavesdropper channel: Cooperation for secrecy," *IEEE Transactions on Information Theory*, vol. 54, no. 9, pp. 4005–4019, 2008.

[17] Y. Oohama, "Capacity theorems for relay channels with confidential messages," *IEEE International Symposium on Information Theory*, pp. 926–930, June 2007.

[18] X. He and A. Yener, "Cooperation with an untrusted relay: A secrecy perspective," *IEEE Transactions on Information Theory*, vol. 56, no. 8, pp. 3807–3827, 2010.

[19] E. Ekrem and S. Ulukus, "Secrecy in cooperative relay broadcast channels," *IEEE Transactions on Information Theory*, vol. 57, no. 1, pp. 137–155, 2011.

[20] Y. Liang, G. Kramer, H. V. Poor, and S. Shamai, "Compound wiretap channels," *EURASIP Journal on Wireless Communications and Networking, Special Issue Wireless Physical Layer Security*, March 2009.

[21] A. Khisti, "Interference alignment for the multiantenna compound wiretap channel," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 2976–2993, 2011.

[22] X. He and A. Yener, "K-user interference channels: Achievable secrecy rate and degrees of freedom," *IEEE Information Theory Workshop*, pp. 336–340, June 2009.

[23] ——, "Providing secrecy with structured codes: Two-user Gaussian channels," *IEEE Transactions on Information Theory*, vol. 60, no. 4, pp. 2121–2138, 2014.

[24] J. Xie and S. Ulukus, "Secure degrees of freedom of one-hop wireless networks," *IEEE Transactions on Information Theory*, vol. 60, no. 6, pp. 3359–3378, 2014.

[25] ——, "Secure degrees of freedom of K-user Gaussian interference channels: A unified view," *IEEE Transactions on Information Theory*, vol. 61, no. 5, pp. 2647–2661, 2015.
[26] ——, "Secure degrees of freedom regions of multiple access and interference channels: The polytope structure," *IEEE Transactions on Information Theory*, vol. 62, no. 4, pp. 2044–2069, 2016.
[27] E. Tekin and A. Yener, "Achievable rates for the general Gaussian multiple access wire-tap channel with collective secrecy," *Allerton Conference On Communication, Control, and Computing*, pp. 809–816, September 2006.
[28] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas-Part II: The MIMOME wiretap channel," *IEEE Transactions on Information Theory*, vol. 56, no. 11, pp. 5515–5532, 2010.
[29] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 4961–4972, 2011.
[30] S. Shafiee, N. Liu, and S. Ulukus, "Towards the secrecy capacity of the Gaussian MIMO wire-tap channel: The 2-2-1 channel," *IEEE Transactions on Information Theory*, vol. 55, no. 9, pp. 4033–4039, 2009.
[31] T. Liu and S. Shamai, "A note on the secrecy capacity of the multiple-antenna wiretap channel," *IEEE Transactions on Information Theory*, vol. 55, no. 6, pp. 2547–2553, 2009.
[32] X. He and A. Yener, "MIMO wiretap channels with unknown and varying eavesdropper channel states," *IEEE Transactions on Information Theory*, vol. 60, no. 11, pp. 6844–6869, 2014.
[33] X. He, A. Khisti, and A. Yener, "MIMO multiple access channel with an arbitrarily varying eavesdropper: Secrecy degrees of freedom," *IEEE Transactions on Information Theory*, vol. 59, no. 8, pp. 4733–4745, 2013.
[34] M. Nafea and A. Yener, "Degrees of freedom of the single antenna Gaussian wiretap channel with a helper irrespective of the number of antennas at the eavesdropper," *IEEE GlobalSIP Symposium on Cyber-Security and Privacy*, pp. 273–276, December 2013.
[35] ——, "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," *Allerton Conference on Communication, Control, and Computing*, pp. 774–780, October 2013.
[36] P. Mukherjee, J. Xie, and S. Ulukus, "Secure degrees of freedom of one-hop wireless networks with no eavesdropper CSIT," *IEEE Transactions on Information Theory*, vol. 63, no. 3, pp. 1898–1922, 2017.
[37] M. A. Maddah-Ali, "On the degrees of freedom of the compound MIMO broadcast channels with finite states," 2009, arXiv preprint arXiv:0909.5006.
[38] D. Kleinbock, "Baker-Sprindzhuk conjectures for complex analytic manifolds," 2002, arXiv preprint math/0210369.
[39] V. Sprindzuk, "On Mahler's conjecture," *Doklady Akademii Nauk SSSR*, vol. 154, pp. 783–786, 1964, (in Russian); English translation in Soviet Math. Dokl. 5, (1964), 183-186.
[40] ——, "More on Mahler's conjecture," *Doklady Akademii Nauk SSSR*, vol. 155, pp. 54–56, 1964, (in Russian); English translation in Soviet Math. Dokl. 5, (1964), 361-363.
[41] T. M. Cover and J. A. Thomas, *Elements of information theory 2nd edition*. New York, NY, USA: Wiley, 2006.
[42] A. Yener, "New directions in information theoretic security: Benefits of bidirectional signaling," *IEEE Information Theory Workshop*, pp. 1–5, April 2014.
[43] A. S. Motahari, S. O. Gharan, M.-A. Maddah-Ali, and A. K. Khandani, "Real interference alignment: Exploiting the potential of single antenna systems," *IEEE Transactions on Information Theory*, vol. 60, no. 8, pp. 4799–4810, 2014.
[44] W. M. Schmidt, *Diophantine approximation*. Berlin, Germany: Springer-Verlag, 1980.

Mohamed Nafea (S'11) received the B.Sc. degree in electrical engineering from Alexandria University, Alexandria, Egypt, in 2010; the M.Sc. degree in wireless communications from Wireless Intelligent Networks Center (WINC), Nile University, Giza, Egypt, in 2012, and the M.A. degree in mathematics from the Department of Mathematics, The Pennsylvania State University, University Park, PA, USA, in 2017. He is currently pursuing his Ph.D. degree in electrical engineering, and has been a Graduate Research Assistant with the Wireless Communications and Networking (WCAN) Laboratory, at The Pennsylvania State University, University Park, PA, USA, since 2012. His research interests include network information theory, information theoretic security, communication theory, statistical learning, mathematical logic and modeling theory.

Aylin Yener (S'91-M'01-SM'14-F'15) received the B.Sc. degree in electrical and electronics engineering and the B.Sc. degree in physics from Bogazici University, Istanbul, Turkey, and the M.S. and Ph.D. degrees in electrical and computer engineering from Wireless Information Network Laboratory (WINLAB), Rutgers University, New Brunswick, NJ, USA. She is a Professor of Electrical Engineering at The Pennsylvania State University, University Park, PA, USA, since 2010, where she joined the faculty as an Assistant Professor in 2002. Since 2017, she is a Deans Fellow in the College of Engineering at The Pennsylvania State University. She is currently also a Visiting Professor at the Department of Electrical Engineering, Stanford University, Stanford, CA, USA. From 2008 to 2009, she was a Visiting Associate Professor with the same department. Her research interests include information theory, communication theory, and network science, with recent emphasis on green communications and information security. She received the NSF CAREER award in 2003, the Best Paper Award in Communication Theory in the IEEE International Conference on Communications in 2010, the Penn State Engineering Alumni Society (PSEAS) Outstanding Research Award in 2010, the IEEE Marconi Prize Paper Award in 2014, the PSEAS Premier Research Award in 2014, and the Leonard A. Doggett Award for Outstanding Writing in Electrical Engineering at Penn State in 2014.

Dr. Yener is currently a member of the Board of Governors of the IEEE Information Theory Society, where she was previously the treasurer (2012-2014). She served as the student committee chair for the IEEE Information Theory Society 2007-2011, and was the co-founder of the Annual School of Information Theory in North America co-organizing the school in 2008, 2009, and 2010. She was a technical (co)-chair for various symposia/tracks at the IEEE ICC, PIMRC, VTC, WCNC, and Asilomar (2005-2014). She served as an editor for the IEEE TRANSACTIONS ON COMMUNICATIONS (2009-2012), an editor and an editorial advisory board member for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS (2001-2012), and a guest editor for the IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY (2011) and the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS (2015). Currently, she serves on the editorial board of the IEEE TRANSACTIONS ON MOBILE COMPUTING and as a senior editor for the IEEE JOURNAL ON SELECTED AREAS ON COMMUNICATIONS.