Cfr discussion eq 10 + notes mais il semblairait qu'Eve puisse tjs avoir le même SINR que Bob et que le schéma soit donc tjs possible avec un SR = 0 —> unsecure.
Il faut pseudo inverse He et non inverse He, pour y arriver, quelque soit N_e N_t, N_b

If H is MIMO Rayleigh or Rician channel models, then HH* (or H*H) results in a Wishart random matrix

Secrecy Capacity Analysis of Artificial Noisy MIMO Channels—An Approach Based on Ordered Eigenvalues of Wishart Matrices

N_A > N_E is assumed

Yiliang Liu, Hsiao-Hwa Chen, *Fellow, IEEE*, and Liangmin Wang, *Member, IEEE*

AN scheme where messages are encoded in s (which is a variable) strongest eigen-subchannels based on ordered eigenvalues of Wishart matrices, while AN signals are generated in remaining t – s spaces.

Uncorrelation between antennas.

Send AN in part of the non null space and part of the null space of the MIMO channel —> special car va y avoir du leakage @B qu'il supprime en post process la data reçue —> more processing needed

***Abstract*— Artificial noise (AN) can be used to confuse eavesdroppers in a physical layer security system. One of the main issues concerned in AN schemes is how to improve secrecy capacities. Most existing AN schemes were proposed based on an assumption that the number of transmit antennas t is larger than that of receiver antennas r, such that they can utilize all r eigen-subchannels of a multiple-output multiple-input (MIMO) system to send messages, and use remaining $t-r$ null spaces for transmitting AN signals. These AN signals null out legitimate receivers and degrade eavesdropper channels. However, transmitting messages in all eigen-subchannels is not always a good strategy. In particular, when the number of transmit antennas is constrained or even smaller than those of receivers, the secrecy capacities of legitimate receivers will be impaired significantly if using all eigen-subchannels for message transmission. To improve secrecy capacity, we propose an AN scheme where messages are encoded in s (which is a variable) strongest eigen-subchannels based on ordered eigenvalues of Wishart matrices, while AN signals are generated in remaining $t-s$ spaces. We derive the average secrecy capacity of a single-user MIMO wiretap channel in the presence of an eavesdropper with multiple antennas. We show that the numerical results are in a good agreement with simulation results. The secrecy capacity of the proposed AN scheme can be improved by approximately 20% ∼ 40% if compared with existing AN schemes.**

***Index Terms*— Artificial noise, secrecy capacity, physical layer security, MIMO wiretap channel, Wishart matrix.**

I. Introduction

WIRELESS communication is vulnerable to eavesdroppers due to the broadcasting nature of wireless channels. In order to deal with security threats, traditional security approaches employ symmetric or asymmetrical cryptographic algorithms to achieve communication confidentiality [1]. Recently, physical layer security has attracted a lot of attention due to its potential to offer stronger security protection. Compared to cryptographic technologies implemented on upper layers, physical layer security explores the randomness nature of physical layer media with the help of appropriate coding and precoding schemes to resist against brute force attacks or analytical attacks [2]. In addition, physical layer security can be implemented with channel coding without the need to use complicated protocols or algorithms, such as complex elliptic curve [3] or bilinear pairing algorithms [4], etc.

Several physical layer security schemes with MIMO technologies were proposed in the literature as an effort to achieve information confidentiality, which can be traced back to Wyner's information theoretic secrecy analysis and secrecy capacity definition [5]. In particular, Yan *et al.* presented an antenna selection scheme that selects two strongest antennas and then encodes messages with Alamouti coding [6], where the signals from the two antennas are complex orthogonal but carry the same information [7]. As these two antennas transmit the same message in a slot, this scheme can not work well in a multiplex system to improve secrecy capacities. Huang and Xiong *et al.* in their research efforts [8], [9] selected a strongest transmit eigen-subchannel based on an optimal instantaneous signal to noise ratio of all eigen-subchannels, which in fact used maximal ratio combining (MRC) for achieving secrecy capacity without a MIMO multiplexing gain. These physical layer security schemes [6], [8]–[11] are sensitive to channel conditions, and may fail to provide good secrecy capacities in bad channel conditions, where, for instance, the main and wiretap channels follow a similar distribution or even wiretap channels are better.

Goel *et al.* in their work [12] introduced an AN scheme to tackle the aforementioned problem that eavesdroppers happen to have better channels than legitimate receivers. In [12], AN signals were generated in multiple spaces, and transmitters can steer their beams such that only eavesdropper channels are degraded. In this way, a positive secrecy capacity can always be guaranteed. Also, Ng and Khandaker *et al.* used the strongest eigen-subchannel for transmitting messages, and used remaining $t-1$ spaces for sending AN signals [13], [14]. In their proposed schemes, a positive secrecy capacity was achievable without multiplexing gains. Tsai *et al.* in their work [15] used r eigen-subchannels to transmit messages and used $t-r$ null spaces for sending AN signals. The scheme

Manuscript received June 22, 2016; revised September 22, 2016; accepted October 18, 2016. Date of publication November 9, 2016; date of current version January 18, 2017. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Lifeng Lai. *(Corresponding author: Hsiao-Hwa Chen.)*

Y. Liu is with the Department of Electronics Information Engineering, Harbin Institute of Technology, Harbin, China (e-mail: alanliuyiliang@gmail.com).

H.-H. Chen is with the Department of Engineering Science, National Cheng Kung University, Tainan 70101, Taiwan (e-mail: hshwchen@mail.ncku.edu.tw).

L. Wang is with the Department of Internet of Things Engineering, Jiangsu University, Zhenjiang, China (e-mail: jasonwanglm@gmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIFS.2016.2627219

offers a multiplexing gain, which improves secrecy capacity significantly. Liu *et al.* derived a close-form average secrecy capacity of Tsai's scheme with the help of non-ordered eigenvalues of a Wishart matrix [16]. However, these schemes are optimal only under the condition of $t \gg r$. When the number of transmit antennas is constrained or even smaller than that of receive antennas, the methods of using all eigen-subchannels of a MIMO system for secure transmission degrade secrecy capacities. This motivates us to design a better AN scheme to improve the secrecy capacity via selecting proper numbers of message-sending eigen-subchannels (denoted by s) and AN-sending spaces (denoted by d).

In this paper, our goal is to address the aforementioned issues in the existing AN schemes. In particular, we will propose an AN scheme for guaranteeing a positive average secrecy capacity. In this proposed scheme, the message-sending eigen-subchannels are chosen based on the first to the kth largest eigenvalues of a Wishart matrix, instead of selecting all its positive eigenvalues. In this way, the number of message-sending eigen-subchannels is not limited by r but by a variable s, which can be leveraged to achieve an optimal secrecy capacity. In addition, we do not assume $t > r$ for secure communications. We also derive an average secrecy capacity of the proposed AN scheme using marginal probability distribution functions (pdf) of the first to the kth largest eigenvalues of the Wishart matrix, by which an optimal average secrecy capacity can be calculated. In addition, the precoding matrices of the proposed scheme can be broadcasted to both legitimate receivers and eavesdroppers. Hence, we do not need to worry about the leakage of key precoding messages for its resistance against analysis attacks, such as brute force attacks and known-plaintext attacks [17].

Specifically, our contributions in this work can be summarized as follows:

1) We present an AN scheme for Rayleigh MIMO scenarios, where the message-sending eigen-subchannels are chosen based on the first to the kth largest eigenvalues of a Wishart matrix (see Section III).
2) We derive the theoretical average secrecy capacity of the proposed AN scheme, which is calculated by the marginal pdfs of the first to the kth largest eigenvalues of the Wishart matrix (see Section IV).

We explain the notations used in this paper as follows. Bold uppercase letters, such as $\mathbf{A}$, denote matrices, and bold lowercase letters, such as $\mathbf{x}$, denote column vectors. $\mathbf{A}^\dagger$ represents a Hermitian transpose of $\mathbf{A}$. $\mathbf{I}_a$ denotes an identity matrix with its rank being a. $\mathbb{C}^{M\times N}$ denotes an M-by-N dimensional complex matrix set. $\mathbb{R}^{M\times N}$ denotes an M-by-N dimensional real matrix set. $\mathrm{E}[\cdot]$ denotes the expectation operator. $[\mathbf{A}]_{i,j}$ denotes the ith row and the jth column element of $\mathbf{A}$. $[\mathbf{A}]_{(i\sim u),(j\sim v)}$ denotes a sub-matrix of $\mathbf{A}$, including the ith to the uth rows and the jth to the vth columns of $\mathbf{A}$. $\mathrm{Re}(x)$ denotes the real part of x. $H(x)$ is the entropy of a random variable x. $I(x;y)$ is the mutual information between random variables x and y. $I(x;y|z) = \mathrm{E}_z[I(x;y)]$ is the conditional mutual information between x and y for given z. $\exp(x)$ denotes the exponential function of x. $\det[\mathbf{A}]$ is the determinant of $\mathbf{A}$. $\mathrm{Rank}(\mathbf{A})$ is to calculate the rank of $\mathbf{A}$. The plain letter $\mathrm{e} \approx 2.72$ is the Euler's number, which is to avoid the confusion with the number of an eavesdropper's antennas e. $[x]^+ = \max(x, 0)$.

The remainders of this paper can be outlined as follows. Section II introduces the preliminaries, focusing mainly on the Wishart matrix distribution. Section III is to discuss the details on the system model and the proposed AN scheme. Section IV is dedicated to secrecy capacity derivations of the proposed scheme. Section V gives numerical analysis and simulations for the proposed scheme and the existing schemes, followed by the conclusions made in Section VI.

II. Preliminaries

The spaces of a MIMO channel denote the orthonormal bases of a MIMO channel matrix. For $\mathbf{H} \in \mathbb{C}^{r\times t}$, let $\mathbf{H}^\dagger\mathbf{H} = \mathbf{U}\Lambda\mathbf{U}^\dagger$ be singular value decomposition (SVD) of $\mathbf{H}^\dagger\mathbf{H}$, where each column vector (also called an orthonormal basis or an eigenvector) in a unitary matrix $\mathbf{U}$ is a space of a MIMO channel, such that the number of spaces is the number of transmit antennas t. Λ is a diagonal matrix, whose elements are eigenvalues of $\mathbf{H}^\dagger\mathbf{H}$, and these eigenvalues are ranked from the largest to the smallest in the diagonal line. The eigenvectors corresponding to positive eigenvalues in Λ are the eigen-subchannels of a MIMO channel. The number of eigen-subchannels denotes the degrees of freedom, which equals to $\mathrm{Rank}(\mathbf{H})$ [18]. The eigenvectors corresponding to zero eigenvalues in Λ are the null spaces of a MIMO channel. The eigen-subchannels used for transmitting messages are defined as message-sending eigen-subchannels (It is impossible to use null spaces to transmit messages). The spaces used for transmitting AN signals are named as AN-sending spaces (both eigen-subchannels and null spaces can be selected for transmitting AN signals).

In this work, we use a part of eigen-subchannels and all remaining spaces to transmit AN signals, while existing AN schemes used all eigen-subchannels to transmit messages and all null spaces to transmit AN signals.

Definition 1 (Uncorrelated central Wishart matrix [19]): Each element of a complex Gaussian random matrix $\mathbf{H} \in \mathbb{C}^{r\times t}$ is a complex variable with its real and imaginary parts obeying normal distribution $\mathcal{CN}(0, 1/2)$. Let us define $m = \max(t, r)$ and $n = \min(t, r)$. Then, the Hermitian matrix $\mathbf{W} \in \mathbb{C}^{n\times n}$ is defined as

$$\mathbf{W} = \begin{cases} \mathbf{H}\mathbf{H}^\dagger, & t \geq r, \\ \mathbf{H}^\dagger\mathbf{H}, & t < r, \end{cases} \tag{1}$$

where $\mathbf{W}$ is called an uncorrelated central Wishart matrix and denoted as $\mathbf{W} \sim W_n(m, \mathbf{0}_n, \mathbf{I}_n)$. The matrix $\mathbf{H}$ is used to represent a Rayleigh MIMO fading channel, whose antennas are uncorrelated with each other.

Theorem 1: The marginal cumulative distribution function (cdf) of the kth largest eigenvalue λ_k of an uncorrelated central Wishart matrix $\mathbf{W} \sim W_n(m, \mathbf{0}_n, \mathbf{I}_n)$ is defined as

$$F_{\lambda_k}(x) = K^{-1} \sum_{i=1}^{k} \sum_{1} \det\left[\Theta(\boldsymbol{\mu}, i; x)\right], \tag{2}$$

where $K = \prod_{i=1}^{n}(t-i)!(r-i)!$, $\sum_1$ denotes the summation over the combination of two sets, i.e.,

$(\mu_1 < \mu_2 < \cdots < \mu_{k-1})$ and $(\mu_k < \mu_{k+1} < \cdots < \mu_n)$, and $(\mu_1, \ldots, \mu_n)$ is a permutation of $(1, \ldots, n)$. An $(n \times n)$ real matrix $\Theta(\boldsymbol{\mu}, i; x)$ is defined as

$$\left[\Theta(\boldsymbol{\mu}, i; x)\right]_{\mu_u, \upsilon} = \begin{cases} \Gamma(m - n + \mu_u + \upsilon - 1, x), & u = 1, \ldots, k-1, \\ \gamma(m - n + \mu_u + \upsilon - 1, x), & u = k, \ldots, n, \end{cases} \tag{3}$$

for $u, \upsilon = 1, \ldots, n$, where $\Gamma(\cdot, \cdot)$ and $\gamma(\cdot, \cdot)$ are the upper and lower incomplete gamma functions defined as

$$\Gamma(a, x) = \int_x^{\infty} \exp(-z) z^{a-1} dz, \tag{4}$$

$$\gamma(a, x) = \int_0^{x} \exp(-z) z^{a-1} dz, \tag{5}$$

respectively.

Proof: See Appendix I.A.

The marginal pdf of the kth largest eigenvalue can be derived readily from Theorem 1, which will be illustrated in the following corollary.

Corollary 1: The marginal pdf of the kth largest eigenvalue λ_k of an uncorrelated central Wishart matrix $\mathbf{W} \sim W_n(m, \mathbf{0}_n, \mathbf{I}_n)$ is given by

$$f_{\lambda_k}(x) = K^{-1} \sum_{i=1}^{k} \sum_{1} \sum_{j=1}^{n} \det\left[\Omega(\boldsymbol{\mu}, i, j; x)\right], \tag{6}$$

where $K = \prod_{i=1}^{n} (t-i)!(r-i)!$, $\sum_1$ denotes the summation over the combination of two sets $(\mu_1 < \mu_2 < \cdots < \mu_{k-1})$ and $(\mu_k < \mu_{k+1} < \cdots < \mu_n)$, and $(\mu_1, \ldots, \mu_n)$ is a permutation of $(1, \ldots, n)$. The $(n \times n)$ real matrix $\Omega(\boldsymbol{\mu}, i, j; x)$ is defined as

$$\left[\Omega(\boldsymbol{\mu}, i, j; x)\right]_{\mu_u, \upsilon} = \begin{cases} \Gamma(m - n + \mu_u + \upsilon - 1, x), & u = 1, \ldots, k-1,\ \mu_u \neq j, \\ -\exp(-x) x^{m-n+\mu_u+\upsilon-2}, & u = 1, \ldots, k-1,\ \mu_u = j, \\ \gamma(m - n + \mu_u + \upsilon - 1, x), & u = k, \ldots, n,\ \mu_u \neq j, \\ \exp(-x) x^{m-n+\mu_u+\upsilon-2}, & u = k, \ldots, n,\ \mu_u = j, \end{cases} \tag{7}$$

where $u, \upsilon = 1, \ldots, n$.

Proof: See Appendix I. B.

Fig. 1 shows the probability distribution functions of λ_1, λ_2, λ_3, and λ_4 of an uncorrelated central Wishart matrix $\mathbf{W}_4(7, \mathbf{0}_4, \mathbf{I}_4)$. We can show that the means of λ_1, λ_2, λ_3, and λ_4 are equal approximately to 14.46, 7.98, 4.03, and 1.53, respectively. The variances of λ_1, λ_2, λ_3, and λ_4 are equal approximately to 10.52, 3.88, 1.58, and 0.51, respectively. A great difference exists between λ_1 and λ_4, which implies that weaker eigen-subchannels give smaller gains. These results are matched to Monte Carlo simulation results of the Wishart matrix very well.

The eigenvalue distribution of Wishart matrices is important for deriving information-theoretic capacity of MIMO channels [20], [21]. For instance, the pdf of the largest eigenvalue can be used to analyze the performance of MIMO-MRC systems [22], and the pdf of the smallest eigenvalue can be used in selection of MIMO antennas [23].

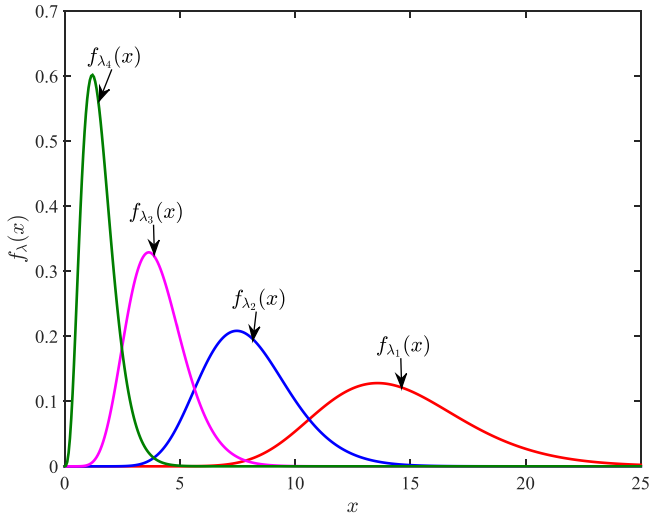


Fig. 1. The probability distribution functions of λ_1, λ_2, λ_3, and λ_4 of a uncorrelated central Wishart matrix $\mathbf{W} \sim W_4(7, \mathbf{0}_4, \mathbf{I}_4)$.

In the literature, the marginal pdf of the kth eigenvalue of Wishart matrices was investigated first in [24]. The works in [25] extended it to other Wishart matrices, including uncorrelated non-central Wishart matrices and correlated non-central (central) Wishart matrices. The authors of [19] and [26] proposed a general framework for deriving the pdf of Wishart matrix arguments to analyze the performance of MIMO singular value decomposition (SVD) systems.

A knows B CSI and E CDI

III. System Model and Proposed AN Scheme

Let us consider a MIMO communication system, which consists of a transmitter (Alice) with t transmit antennas, a legitimate receiver (Bob) with r receive antennas, and an eavesdropper (Eve) with e receive antennas, as shown in Fig. 2. We also assume that Alice has full channel state information (CSI) of Bob via a broadcast feedback channel, but knows Eve's channel distribution information (CDI) only. In general, the main channel between Alice and Bob, and the wiretap channel between Alice and Eve are defined as complex Gaussian matrices $\mathbf{H} \in \mathbb{C}^{r \times t}$ and $\mathbf{H}_e \in \mathbb{C}^{e \times t}$ with each element obeying distribution $\mathcal{CN}(0, 1)$, respectively. Here, we assume $t > e$ for secure transmission, and define $m = \max(t, r)$ and $n = \min(t, r)$.

Uncorrelated channels

In our scheme, there are s $(s \leq t)$ message-sending eigen-subchannels, which are selected by Alice based on the CSI feedback from Bob. More specifically, Alice performs SVD of $\mathbf{H}^{\dagger}\mathbf{H} \in \mathbb{C}^{t \times t}$ in a pre-processor, whose output is a unitary matrix $\mathbf{U} \in \mathbb{C}^{t \times t}$, its Hermitian transpose form $\mathbf{U}^{\dagger} \in \mathbb{C}^{t \times t}$, and a diagonal matrix $\Lambda \in \mathbb{R}^{t \times t}$, which consists of positive and zero eigenvalues of $\mathbf{H}^{\dagger}\mathbf{H}$. Then, Alice generates a message precoding matrix $\mathbf{B} \in \mathbb{C}^{t \times s}$, whose columns are the eigenvectors corresponding to the first to the sth largest eigenvalues of $\mathbf{H}^{\dagger}\mathbf{H}$, and an AN precoding matrix $\mathbf{Z} \in \mathbb{C}^{t \times d}$ $(s + d = t)$, whose columns are the eigenvectors of remaining eigenvalues of $\mathbf{H}^{\dagger}\mathbf{H}$.

Lemma 1: $[\mathbf{HB}]^{\dagger}\mathbf{HZ} = \mathbf{0}$, $[\mathbf{HB}]^{\dagger}\mathbf{H}_e\mathbf{Z} \neq \mathbf{0}$, and $[\mathbf{H}_e\mathbf{B}]^{\dagger}\mathbf{H}_e\mathbf{Z} \neq \mathbf{0}$.

Proof: See Appendix II. A.

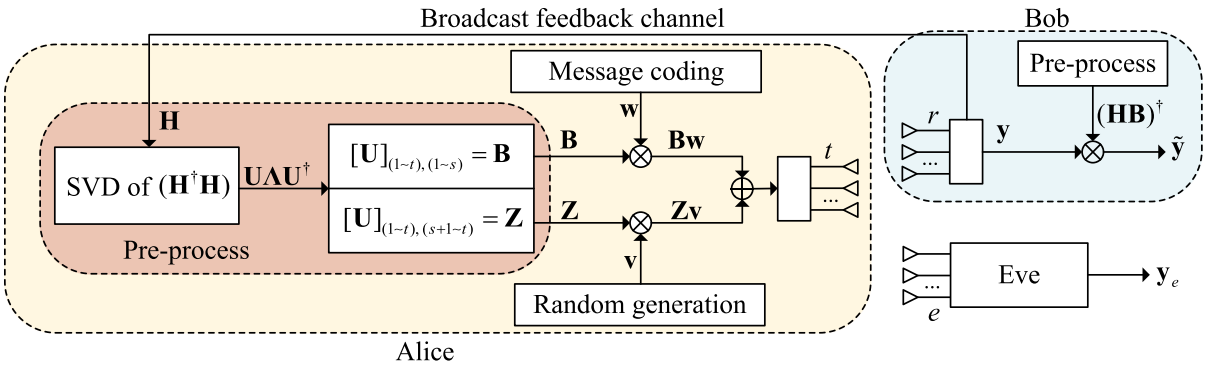


Fig. 2. Illustration of a MIMO wiretap channel model with the proposed AN scheme, where the main channel between Alice and Bob, and the wiretap channel between Alice and Eve are assumed to be Rayleigh channels. Alice has t transmit antennas, Bob has r receive antennas, and Eve has e receive antennas. Assume that Alice, Bob, and Eve have the same pre-processor. Alice performs SVD of $\mathbf{H}^\dagger\mathbf{H}$ in the pre-processor, whose output is a unitary matrix $\mathbf{U}$, and then generates a message precoding matrix $\mathbf{B} = [\mathbf{U}_{(1\sim t),(1\sim s)}]$ and an AN precoding matrix $\mathbf{Z} = [\mathbf{U}_{(1\sim t),(s+1\sim t)}]$. Bob can eliminate the AN signals using the same pre-processor, while Eve is incapable to eliminate these AN signals.

Alice transmits $\mathbf{Bw} + \mathbf{Zv}$ via s message-sending eigen-subchannels and d AN-sending spaces, respectively. It means that each antenna transmits a combination of message components and AN components, but the AN components can be eliminated by the pre-processor at Bob. In this way, we create a capcity difference between the main channels and wiretap channels. Note that $\mathbf{B}$ and $\mathbf{Z}$ are fixed semi-unitary matrices derived from $\mathbf{H}$. Hence, we have $\mathbf{B}^\dagger\mathbf{B} = \mathbf{I}_s$ and $\mathbf{Z}^\dagger\mathbf{Z} = \mathbf{I}_d$.

According to CSI matrix $\mathbf{H}$ and pre-processing method used, the received signals at Bob and Eve can be expressed as

$$\mathbf{y} = \mathbf{HBw} + \mathbf{HZv} + \mathbf{n}, \tag{8a}$$

$$\mathbf{y}_e = \mathbf{H}_e\mathbf{Bw} + \mathbf{H}_e\mathbf{Zv} + \mathbf{n}_e, \tag{8b}$$

respectively. Here, $\mathbf{w}$ is a transmitted signal of the desired user, and $\mathbf{v}$ is a random AN signal. Both $\mathbf{w}$ and $\mathbf{v}$ are circularly symmetric complex Gaussian vectors[1] with zero-means and its covariance matrices $P/t\mathbf{I}_s$ and $P/t\mathbf{I}_d$, respectively, where P is the average transmit power constraint. For simplification, we distribute total power to each antenna evenly. Reference [28] showed that an equal power distribution approach offers a channel capacity approximate to optimal power allocation. $\mathbf{n}$ and $\mathbf{n}_e$ are additive white Gaussian noise (AWGN) vectors with their covariance matrices $\mathbf{I}_r$ and $\mathbf{I}_e$, respectively.

Bob can eliminate the AN signal $\mathbf{v}$ by pre-processing ($[\mathbf{HB}]^\dagger\mathbf{HZ} = \mathbf{0}$) the received signal $\mathbf{y}$ via

Bob must know B (the pre processing matrix)

$$\tilde{\mathbf{y}} = \left[\mathbf{HB}\right]^\dagger\mathbf{y} = \Lambda_s\mathbf{w} + \tilde{\mathbf{n}}, \tag{9}$$

where $\tilde{\mathbf{n}} = [\mathbf{HB}]^\dagger\mathbf{n} \in \mathbb{C}^{s\times 1}$ is an AWGN vector with its distribution $\mathcal{CN}(\mathbf{0}, \Lambda_s)$. $\Lambda_s \in \mathbb{R}^{s\times s}$ is a diagonal matrix formed by the first to the sth eigenvalues of $\mathbf{H}^\dagger\mathbf{H}$. In the elimination process, the received signal multiplied by a fixed matrix will not change its capacity. Even we consider the worst case that Eve has the knowledge of $\mathbf{H}$, $\mathbf{H}_e$, $\mathbf{B}$, and $\mathbf{Z}$, the AN signal still degrades Eve's channel because Eve can not eliminate the AN signal as $[\mathbf{HB}]^\dagger\mathbf{H}_e\mathbf{Z} \neq \mathbf{0}$ and $[\mathbf{H}_e\mathbf{B}]^\dagger\mathbf{H}_e\mathbf{Z} \neq \mathbf{0}$. Pq elle pourrait pas faire une pseudo inverse si t < e ?

Remark 1: For secure transmission, Alice is required to equip more antennas than Eve, i.e., $t > e$; otherwise, Eve can eliminate the AN signal by a left inverse matrix of $\mathbf{H}_e$, denoted by $\mathbf{H}_e^{-1}$. The elimination process can be expressed as

assez contraignant car en realite
Alice ne peut pas estimer N_E

$$\begin{aligned}\ddot{\mathbf{y}} &= \left[\mathbf{HB}\right]^\dagger\mathbf{H}\mathbf{H}_e^{-1}\mathbf{y}_e \\ &= \left[\mathbf{HB}\right]^\dagger\mathbf{H}\mathbf{H}_e^{-1}\mathbf{H}_e\mathbf{Bw} + \left[\mathbf{HB}\right]^\dagger\mathbf{H}\mathbf{H}_e^{-1}\mathbf{H}_e\mathbf{Zv} + \mathbf{Q}_e \\ &= \Lambda_s\mathbf{w} + \left[\mathbf{HB}\right]^\dagger\mathbf{HZv} + \mathbf{Q}_e,\end{aligned} \tag{10}$$

where $\left[\mathbf{HB}\right]^\dagger\mathbf{HZ} = \mathbf{0}$, and $\mathbf{Q}_e = \left[\mathbf{HB}\right]^\dagger\mathbf{H}\mathbf{H}_e^{-1}\mathbf{n} \in \mathbb{C}^{s\times 1}$ is an AWGN vector with its distribution $\mathcal{CN}(\mathbf{0}, \Lambda_s)$. Hence, $\ddot{\mathbf{y}} = \tilde{\mathbf{y}}$, which means Eve receives the message signal that has the same quality with Bob.

cfr notes mais du coup ici Eve peut tjs avoir le meme SINR que Bob —> SR = 0

IV. Secrecy Capacity of the Proposed Scheme

In this section, we derive an average secrecy capacity expression for the proposed AN scheme using random matrix statistics.

In the MIMO wiretap channel model, while Bob has the knowledge of $\mathbf{H}$, and Eve has the knowledge of $\mathbf{H}$ and $\mathbf{H}_e$, the instantaneous secrecy capacity is

$$C_s = \max_{p(\mathbf{w}),p(\mathbf{v})} \left\{I(\mathbf{w};\mathbf{y}) - I(\mathbf{w};\mathbf{y}_e)\right\}, \tag{11}$$

where the maximization is taken over all possible input distributions of $p(\mathbf{w})$ and $p(\mathbf{v})$ [5].

However, it is hard to find optimal distributions of $p(\mathbf{w})$ and $p(\mathbf{v})$ to maximize the instantaneous secrecy capacity. Hence, we follow the convention in [16] and use Gaussian input alphabets and Gaussian AN, i.e., both $\mathbf{w}$ and $\mathbf{v}$ are circularly symmetric complex Gaussian vectors. In this case, the instantaneous secrecy capacity can be expressed as

$$C_s = [C_m - C_w]^+, \tag{12}$$

where $[x]^+ = \max(x, 0)$. C_m and C_w are

$$\begin{aligned}C_m &= \log_2\det(\mathbf{I}_r + (P/t)\mathbf{H}_1\mathbf{H}_1^\dagger), \\ C_w &= \log_2\det\left(\mathbf{I}_e + \frac{(P/t)\mathbf{H}_2\mathbf{H}_2^\dagger}{(P/t)\mathbf{H}_3\mathbf{H}_3^\dagger + \mathbf{I}_e}\right),\end{aligned} \tag{13}$$

[1]From [20], the channel capacity is maximized when $\mathbf{w}$ is a circularly symmetric complex Gaussian vector. And if $\mathbf{v}$ is a circularly symmetric complex Gaussian vector, AN signal achieves indeterminacy maximization. Note that non-Gaussian AN was also used in the literature [27].

respectively.

$$\begin{aligned}\mathbf{H}_1 &= \mathbf{HB} \in \mathbb{C}^{r\times s},\\ \mathbf{H}_2 &= \mathbf{H}_e\mathbf{B} \in \mathbb{C}^{e\times s},\\ \mathbf{H}_3 &= \mathbf{H}_e\mathbf{Z} \in \mathbb{C}^{e\times d},\end{aligned} \tag{14}$$

respectively.

However, we can not derive the instantaneous secrecy capacity by Eqn. (12) in the absence of Eve's CSI. To measure the secrecy capacity of the proposed scheme, the average secrecy capacity is calculated with Eve's CDI, as a global security metric, assuming that the communication lasts longer enough to experience all channel states.

The full knowledge of $\mathbf{H}$ has no effect on optimizing the average secrecy capacity in our scheme. With a fixed s, Alice can calculate her average secrecy capacity in terms of the CDI of $\mathbf{H}$ and $\mathbf{H}_e$. Our task is to search for the optimal s. The average secrecy capacity is a function of s, which is embedded in $\mathbf{y}$ and $\mathbf{y}_e$. Hence, the main channel output is $(\mathbf{y}, \mathbf{H})$, and the wiretap channel output is $(\mathbf{y}_e, \mathbf{H}, \mathbf{H}_e)$. The average secrecy capacity $\tilde{C}_s$ is expressed as

$$\tilde{C}_s = \max_{p(\mathbf{w}),p(\mathbf{v})} \big\{ I(\mathbf{w};\mathbf{y}|\mathbf{H}) - I(\mathbf{w};\mathbf{y}_e|(\mathbf{H}_e,\mathbf{H}))\big\}, \tag{15}$$

where the maximization is taken over all possible input distributions of $p(\mathbf{w})$ and $p(\mathbf{v})$.

Theorem 2: If message signal $\mathbf{w}$ forms a circularly symmetric complex Gaussian vector, $I(\mathbf{w};\mathbf{y}|\mathbf{H}) - I(\mathbf{w};\mathbf{y}_e|(\mathbf{H}_e,\mathbf{H}))$ is its maximization if and only if $\mathbf{v}$ is also a circularly symmetric complex Gaussian vector. And the average secrecy capacity $\tilde{C}_s$ can be written as

Pas une approximation?!

$$\tilde{C}_s = \mathrm{E}_{\mathbf{H}}[C_m] - \mathrm{E}_{\mathbf{H}_e,\mathbf{H}}[C_w]. \tag{16}$$

Proof: See Appendix II. B.

We can calculate the average secrecy capacity by Monte Carlo simulations [29] using Eqn. (16). In this paper, we use AN signals to interfere with the wiretap channels, which make wiretap channels worse than the main channels. Hence, the average secrecy capacity of the proposed scheme is always positive. The proposed scheme is different from classic schemes [30], which view wiretap channel gains as variables, and make the variation range of wiretap channel smaller than the main channels to guarantee a positive average secrecy capacity. In order to go further to investigate the performance of the AN scheme, we should use the knowledge of the kth eigenvalue's pdf of Wishart matrices to derive an exact average secrecy capacity expression, as illustrated in Theorems 3 and 4.

A. Exact Secrecy Capacity Expression

Recall that the system parameter s is the number of the selected message-sending eigen-subchannels, t and r are the numbers of the transmit and receive antennas, respectively. We can express the exact average secrecy capacity as a function of $\mathbf{w}$, $\mathbf{v}$, $\mathbf{H}$, and $\mathbf{H}_e$.

Theorem 3 ($s \geq \min(t,r)$): For given CSI matrix $\mathbf{H}$ and CDI matrix $\mathbf{H}_e$, the average secrecy capacity $\tilde{C}_s(\mathbf{w},\mathbf{v},\mathbf{H},\mathbf{H}_e)$ is

$$\begin{aligned}&\tilde{C}_s(\mathbf{w},\mathbf{v},\mathbf{H},\mathbf{H}_e)\\ &\quad= C\big[\mathbf{H},(P/t)\big] + C\big[\mathbf{H}_3,(P/t)\big] - C\big[\mathbf{H}_4,(P/t)\big],\end{aligned} \tag{18}$$

where $\mathbf{H}_4 = [\mathbf{H}_2,\mathbf{H}_3] = \mathbf{H}_e\mathbf{U} \in \mathbb{C}^{e\times t}$.

For any $\mathbf{A} \in \mathbb{C}^{\alpha\times\beta}$, and

$$n = \min(\alpha,\beta), \tag{19}$$

$$m = \max(\alpha,\beta), \tag{20}$$

$C(\mathbf{A},\rho)$ is defined as [21, Th. 1]

$$\begin{aligned}&C(\mathbf{A},\rho)\\ &= n\int_0^\infty \log_2(1+\rho\lambda) f(\lambda) d\lambda\\ &= \frac{\exp(1/\rho)}{\ln(2)} \sum_{k=0}^{n-1}\sum_{l=0}^{k}\sum_{i=0}^{2l} \left\{ \frac{(-1)^i (2l)!(m-n+i)!}{2^{2k-i} l!\, i!\,(m-n+l)!}\right.\\ &\quad\left.\times \binom{2k-2l}{k-l}\binom{2l+2m-2n}{2l-i} \sum_{j=0}^{m-n+i} E_{j+1}(1/\rho)\right\},\end{aligned} \tag{21}$$

where $f(\lambda)$ is the pdf of a randomly selected eigenvalue λ of the Wishart matrix of $\mathbf{W}$ ($\mathbf{AA}^\dagger$ or $\mathbf{A}^\dagger\mathbf{A}$) distributed over $W_n(m,\mathbf{0}_n,\mathbf{I}_n)$, and $E_\tau(z)$ is the exponential integral of order τ as

$$E_\tau(z) = \int_1^\infty \exp(-zx)x^{-\tau}dx, \quad \tau = 0,1,\ldots,\mathrm{Re}\{z\} > 0. \tag{22}$$

Actually, for a given $\mathbf{H}$, we have $m = \max(t,r)$ and $n = \min(t,r)$. For a given $\mathbf{H}_4$, we get $m = \max(t,e)$ and $n = \min(t,e)$. For a given $\mathbf{H}_3$, we have $m = \max(d,e)$ and $n = \min(d,e)$.

Proof: See Appendix II. C.

Remark 2: It is interesting to note from Theorem 3 that in the proposed AN scheme for MIMO channels, when $t \gg r = e$, the optimal choice of s equals to the rank of $\mathbf{H}$. It means that Alice can make a full use of all eigen-subchannels of an MIMO system, i.e., $s \geq \min(t,r)$. However, $s > \min(t,r)$ seems to be a bad choice because messages can not be decoded correctly at Bob, as it degrades the capacity of the main channels, and thus we limit $s = \min(t,r)$. When $r < e$, $s = \min(t,r)$ must be checked to see if it is the best choice because the wiretap channel is better than the main channel. In this case, we may allocate more spaces for transmitting AN signals, and s can be smaller than $\min(t,r)$. Moreover, in the case $t < r$, $s < \min(t,r)$ is a reality we have to deal with. Hence, we derive the average secrecy capacity for the case $s < \min(t,r)$ as follows:

Theorem 4 ($s < \min(t,r)$): For given CSI matrix $\mathbf{H}$ and CDI matrix $\mathbf{H}_e$, the average secrecy capacity function $\tilde{C}_s(\mathbf{w},\mathbf{v},\mathbf{H},\mathbf{H}_e)$ is

$$\begin{aligned}&\tilde{C}_s(\mathbf{w},\mathbf{v},\mathbf{H},\mathbf{H}_e)\\ &\quad= \Upsilon\big[\mathbf{H},(P/t),s\big] + C\big[\mathbf{H}_3,(P/t)\big] - C\big[\mathbf{H}_4,(P/t)\big],\end{aligned} \tag{23}$$

where $\Upsilon(\mathbf{H},(P/t),s)$ is defined in Eqn. (17), as shown at the top of the next page, $\sum_1$ denotes the summation over the combination of sets $(\mu_1 < \mu_2 < \cdots < \mu_{k-1})$ and

$$\Upsilon(\mathbf{H}, (P/t), s) = K^{-1} \sum_{k=1}^{s} \sum_{i=1}^{k} \sum_{1} \sum_{j=1}^{n} \int_0^{\infty} \log_2 \left(1 + (P/t)x\right) \det[\Omega(\boldsymbol{\mu}, i, j; x)] dx, \tag{17}$$

$(\mu_k < \mu_{k+1} < \cdots < \mu_n)$, and $(\mu_1, \ldots, \mu_n)$ is a permutation of $(1, \ldots, n)$. The $(n \times n)$ real matrix $\Omega(\boldsymbol{\mu}, i, j; x)]_{\mu_u, \upsilon}$ is defined in Eqn. (7) for $u, \upsilon = 1, \ldots, n$, and $C(\mathbf{H}_4, (P/t))$ and $C(\mathbf{H}_3, (P/t))$ are defined in Eqn. (21).

Proof: See Appendix II. D.

Remark 3: Theorem 4 is used to calculate the average secrecy capacity when the number of Alice's antennas is constrained (even smaller than r). In this case, the optimal choice of s no longer equals to the number of eigen-subchannels. A better method is to allocate stronger eigen-subchannels to transmit messages and weaker eigen-subchannels to transmit AN signals. It is obvious that the optimal s satisfies $s \leq \min(t, r)$, and the use of an enumeration method is possible since s is an integer taken from zero to $\min(t, r)$. As stated in Remark 2, when $r < e$, the value of s will further shrink. We define s^* and corresponding d^* as the optimal allocation parameters.

Closed forms only for high/low SNR regimes

B. Secrecy Capacity Analysis

The derived average secrecy capacity expressions in Eqns. (18) and (23) are not in closed forms. We can simplify these expressions for high and low SNR regions, to show the impacts of power P and the numbers of antennas t, r, and e on the secrecy capacity.

$C[\mathbf{A}, (P/t)]$ and $\Upsilon\left[\mathbf{A}, (P/t), s\right]$ are central items in Eqns. (18) and (23), which can be expressed as

$$C[\mathbf{A}, (P/t)] = \sum_{i=1}^{n} \mathrm{E}[\log_2(1 + (P/t)\lambda_i(\mathbf{A}))], \tag{24}$$

$$\Upsilon\left[\mathbf{A}, (P/t), s\right] = \sum_{i=1}^{s} \mathrm{E}[\log_2(1 + (P/t)\lambda_i(\mathbf{A}))], \tag{25}$$

respectively, where $\lambda_1(\mathbf{A}) > \lambda_2(\mathbf{A}) > \ldots > \lambda_n(\mathbf{A})$ are the ordered eigenvalues of $\mathbf{A}\mathbf{A}^{\dagger}$. $\mathbf{A} \in \mathbb{C}^{\alpha \times \beta}$, and n is defined in Eqn. (19).

As $\{(P/t)\lambda_i(\mathbf{A})\}$ is the SNR of an eign-subchannel of $\mathbf{A}$, let us look at approximate average secrecy capacity expressions in low and high SNR regions.

Theorem 5: Eqns. (18) and (23) in high SNR and low SNR regions can approximately be expressed as

$$\tilde{C}_s^H = (x + y - e)\log_2(P/t) + \chi_1 + \chi_2 - \chi_3, \tag{26}$$

$$\tilde{C}_s^L = \Big(\frac{\sum_{i=1}^{x} \mathrm{E}[\lambda_i(\mathbf{H})] - s * e}{t}\Big) P \log_2(\mathrm{e}), \tag{27}$$

respectively, where

$$\chi_1 = \sum_{i=1}^{x} \log_2(\mathrm{E}[\lambda_i(\mathbf{H})]), \tag{28}$$

$$\chi_2 = \log_2 \frac{z!}{(z - y)!}, \tag{29}$$

$$\chi_3 = \log_2 \frac{t!}{(t - e)!}, \tag{30}$$

$$x = \begin{cases} r, & s = \min(t, r), \\ s, & s < \min(t, r), \end{cases}$$
$$y = \min(e, d),$$
$$z = \max(e, d). \tag{31}$$

Proof: See Appendix II. E.

Obviously, it is seen that Eqn. (26) grows logarithmically with P, and Eqn. (27) grows lineally with P when t, r, e, and s are fixed. Note that the optimal s^* can be derived by enumerating s in Eqns. (18) and (23). Increasing e alone needs more spaces for sending AN signals, which decreases message-sending eigen-subchannels s. And $\{\sum_{i=1}^{x} \mathrm{E}[\lambda_i(\mathbf{H})]\}$ decreases with a decreasing s. Thus, Eqns. (26) and (27) decreases with e faster than linearly.

However, when the numbers of antennas are variables, Eqns. (26) and (27) look like more complicated. For example, in the case where $r = e$, an increasing number of receive antennas provides more eigen-subchannels to transmit messages, but the case of more eavesdropper antennas requires more spaces for AN signals. These two opposite effects imply that the impact of increasing r (equals to e) on the average secrecy capacity is not clear. Here, let us consider a special case with an adequate number of transmit antennas.

Case 1: $t : r = \kappa \geq 2$ and $e = r$.

From the discussion in Remark 2, when $t : r = \kappa \geq 2$ and $e = r$, we have $x = r$, $y = e$, and $z = t - r$.

In the case of high SNR regions, $x = r$, $y = e$, and $z = t - r$, we get $\tilde{C}_s^H = C[\mathbf{H}_3, (P/t)]$, where $\mathbf{H}_3$ is a $r \times r(\kappa - 1)$ matrix. From [20], we know that $C[\mathbf{H}_3, (P/t)]$ increases with r (equals to e and t/κ). Hence, the average secrecy capacity increases with the numbers of Alice, Bob and Eve's antennas.

In the case of low SNR regions, $x = r$, $y = e$, and $z = t - r$, Eqn. (27) can be expressed as

$$\tilde{C}_s^L = \Big(\frac{\sum_{i=1}^{x} \mathrm{E}[\lambda_i(\mathbf{H})] - s * e}{t}\Big) P \log_2(\mathrm{e})$$
$$= \Big((1 - \frac{1}{\kappa})r\Big) P \log_2(\mathrm{e}). \tag{32}$$

In low SNR regions, with a constant κ, the average secrecy capacity also increases with the numbers of Alice, Bob and Eve's antennas.

Then, let us consider the second special case with an adequate number of receive antennas and diversity technologies (s=1).

Case 2: $t : r = \kappa \ll 1$, $s = 1$ and $e \leq t - 1$.

From Eqn. (31), when $t : r = \kappa \ll 1$, $s = 1$ and $e \leq t - 1$, we get $x = 1$, $y = e$, and $z = t - 1$.

TABLE I

IMPACTS OF POWER AND ANTENNAS

Conditions	Impact of power	Impact of antennas
t and r fixed	$\tilde{C}_s \uparrow$ with $P \uparrow$	$\tilde{C}_s \downarrow$ with $e \uparrow$
$t : r = \kappa \geq 2$ $e = r$	$\tilde{C}_s \uparrow$ with $P \uparrow$	$\tilde{C}_s \uparrow$ with $t \uparrow$
$t : r = \kappa \ll 1$ $s = 1$, $e \leq t - 1$	$\tilde{C}_s \uparrow$ with $P \uparrow$	$\tilde{C}_s \rightarrow$ with $t \uparrow$

In the case of high SNR regions, $x = 1$, $y = e$, and $z = t-1$, Eqn. (26) can be expressed as

$$\begin{aligned}\tilde{C}_s^H &= \mathrm{E}[\log_2(1 + (P/t)\lambda_1(\mathbf{H}))] \\ &\quad + \log_2 \frac{(t-1)!}{(t-1-e)!} - \log_2 \frac{(t)!}{(t-e)!} \\ &\leq \log_2(1 + \mathrm{E}[(P/t)\lambda_1(\mathbf{H})]) + \log_2(1 - \frac{e}{t}). \end{aligned} \quad (33)$$

Using the limit theorem on the distribution of the eigenvalues of large dimensional random matrices [31, Th. 2.13], we get

$$\lim_{t \to +\infty} \frac{\lambda_1(\mathbf{H})}{t} = (\sqrt{1/\kappa} + 1)^2. \quad (34)$$

With a constant κ, $\log_2(1+\mathrm{E}[(P/t)\lambda_1(\mathbf{H})])$ and $\log_2(1-e/t)$ converge to two deterministic constants for a large t. We make sure that the average secrecy capacity converges to a constant as the numbers of Alice and Bob's antennas increases.

In the case of low SNR regions, $x = 1$, $y = e$, and $z = t-1$, Eqn. (27) can be expressed as

$$\begin{aligned}\tilde{C}_s^L &= \Big(\frac{\sum_{i=1}^{x} \mathrm{E}[\lambda_i(\mathbf{H})] - s * e}{t}\Big) P \log_2(\mathrm{e}) \\ &= \frac{\mathrm{E}(\lambda_1(\mathbf{H})) - e}{t} P \log_2(\mathrm{e}) \\ &= \big((\sqrt{1/\kappa} + 1)^2 - \frac{e}{t}\big) P \log_2(\mathrm{e}). \end{aligned} \quad (35)$$

Given a constant κ, Eqn. (35) also converges to a deterministic constant for a large t.

In conclusion, we use Table I to show the impacts of the power P and the numbers of antennas t, r and e on the average secrecy capacity. We use $\uparrow$, $\downarrow$ and $\rightarrow$ to represent "increase", "decrease", and "constant", respectively. For example, "$\tilde{C}_s \uparrow$ with $P \uparrow$" means that "the average secrecy capacity increases with increasing transmit power". We also see that, increasing transmit power has a strong effect on the average secrecy capacities in low SNR regions.

V. NUMERICAL AND SIMULATE RESULTS

In this section, numerical and simulations results are provided to investigate joint impacts of the number of antennas, the number of selected message-sending eigen-subchannels, and transmit power on the secrecy capacities. In addition, we will also examine whether numerical and simulation results are consistent or not. As shown in all the figures given below, the numerical results are in a very good agreement with the Monte Carlo simulation results, verifying the accuracy of our derivations. Table II gives the means and variances of eigenvalues of

TABLE II

THE MEANS AND VARIANCES OF EIGENVALUES OF WISHART MATRICES $\mathbf{W} \sim W_n(m, \mathbf{0}_n, \mathbf{I}_n)$

Matrix	Means and variances	λ_1	λ_2	λ_3	λ_4	λ_5
$W_5(9, \mathbf{0}_5, \mathbf{I}_5)$	$\mathrm{E}(\lambda_i)$	19.58	12.20	7.43	4.07	1.73
	$\mathrm{V}(\lambda_i)$	13.24	5.54	2.74	1.30	0.51
$W_5(7, \mathbf{0}_5, \mathbf{I}_5)$	$\mathrm{E}(\lambda_i)$	16.46	9.65	5.42	2.61	0.86
	$\mathrm{V}(\lambda_i)$	11.50	4.49	2.01	0.81	0.22
$W_4(12, \mathbf{0}_4, \mathbf{I}_4)$	$\mathrm{E}(\lambda_i)$	21.58	13.66	8.36	4.39	
	$\mathrm{V}(\lambda_i)$	14.82	6.46	3.37	1.73	
$W_4(8, \mathbf{0}_4, \mathbf{I}_4)$	$\mathrm{E}(\lambda_i)$	15.94	9.13	4.88	2.04	
	$\mathrm{V}(\lambda_i)$	11.42	4.40	1.92	0.72	
$W_4(7, \mathbf{0}_4, \mathbf{I}_4)$	$\mathrm{E}(\lambda_i)$	14.46	7.98	4.03	1.53	
	$\mathrm{V}(\lambda_i)$	10.52	3.88	1.58	0.51	
$W_3(5, \mathbf{0}_3, \mathbf{I}_3)$	$\mathrm{E}(\lambda_i)$	9.52	4.16	1.32		
	$\mathrm{V}(\lambda_i)$	7.58	2.18	0.53		

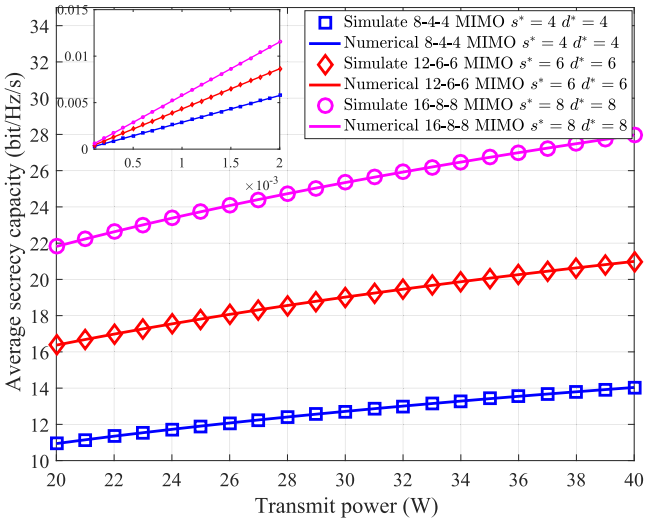


Fig. 3. Average secrecy capacities of numerical analysis and simulations with the optimal choices of (s, d), where we assume $t : r : e = 2 : 1 : 1$.

Wishart matrices $\mathbf{W} \sim W_n(m, \mathbf{0}_n, \mathbf{I}_n)$, including $W_5(9, \mathbf{0}_5, \mathbf{I}_5)$, $W_5(7, \mathbf{0}_5, \mathbf{I}_5)$, $W_4(12, \mathbf{0}_4, \mathbf{I}_4)$, $W_4(8, \mathbf{0}_4, \mathbf{I}_4)$, $W_4(7, \mathbf{0}_4, \mathbf{I}_4)$, and $W_3(5, \mathbf{0}_3, \mathbf{I}_3)$. These Wishart matrices are used in MIMO physical layer security analysis and simulations, which help to rank the eigen-subchannels for each corresponding MIMO channel according to their strengths.

Fig. 3 illustrates the impact of the number of antennas on the secrecy capacity with the optimal choices of (s, d), where we assume $t : r : e = 2 : 1 : 1$, representing the scenarios with abundant transmit antennas. As shown in the figure, an increasing number of transmit antennas improves achievable average secrecy capacities. We can also see that the average secrecy capacity in high SNR regions grows logarithmically with P, and grows lineally with P in low SNR regions. Here, the sub-diagram shows the performance in low SNR regions, which also applies to the simulations.

Fig. 4 compares the average secrecy capacities in terms of the optimal number of selected message-sending eigen-subchannels s^* and other choices of s. Here, we set $t : r = 2 : 1$ and $r < e$. Note that the work in [16] used all eigen-subchannels to transmit messages with $s = r$ and $d = t - r$. We derived an optimal s^* and d^* via Theorems 3 and 4, and then found that s^* and d^* are 3 and 5 in 8-4-6 MIMO

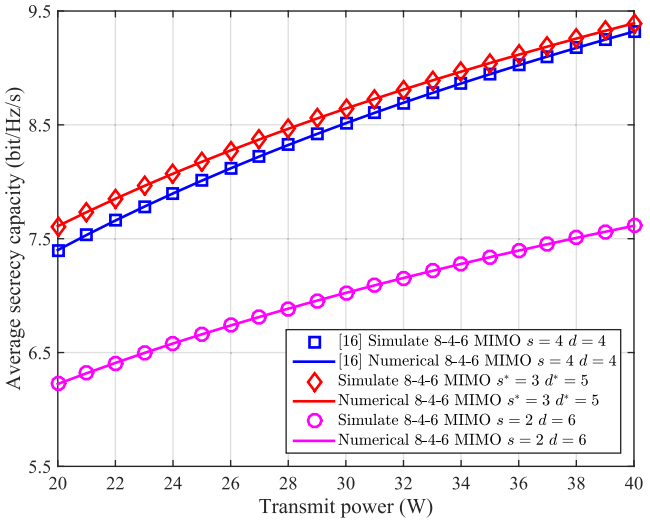


Fig. 4. Average secrecy capacities based on numerical analysis and simulations in terms of different choices of (s, d), where we assume $t : r = 2 : 1$ and $r < e < t$.

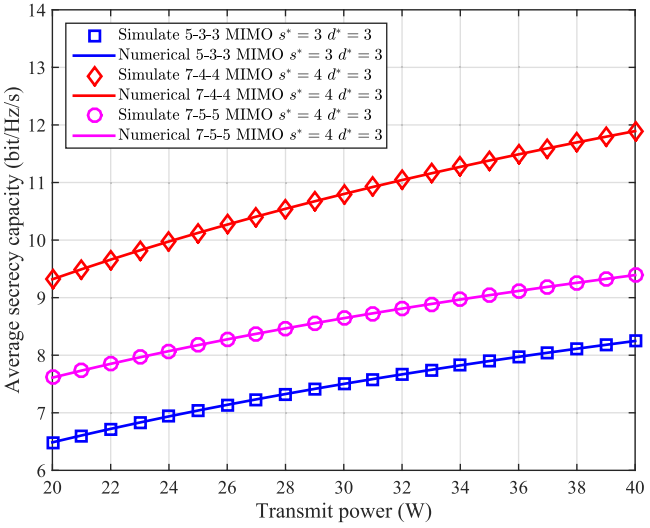


Fig. 5. Average secrecy capacities based on numerical analysis and simulations with the optimal choices of (s, d), where we assume $t : r : e \leq 2 : 1 : 1$ and $t > r$.

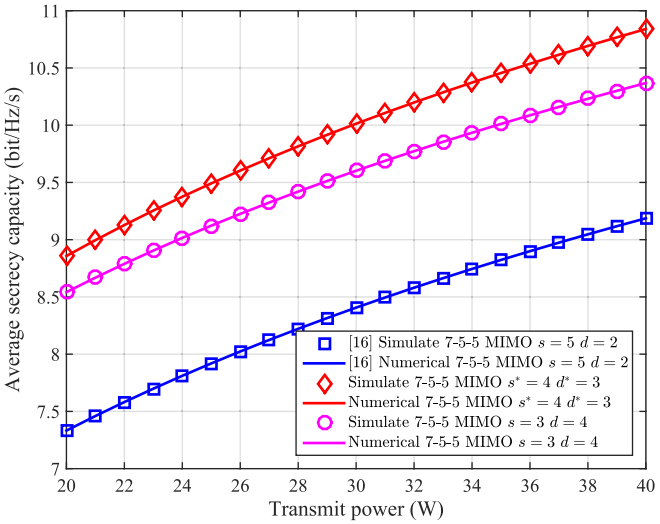


Fig. 6. Average secrecy capacities based on numerical analysis and simulations in terms of different choices of (s, d), where we assume $t : r : e \leq 2 : 1 : 1$ and $t > r$.

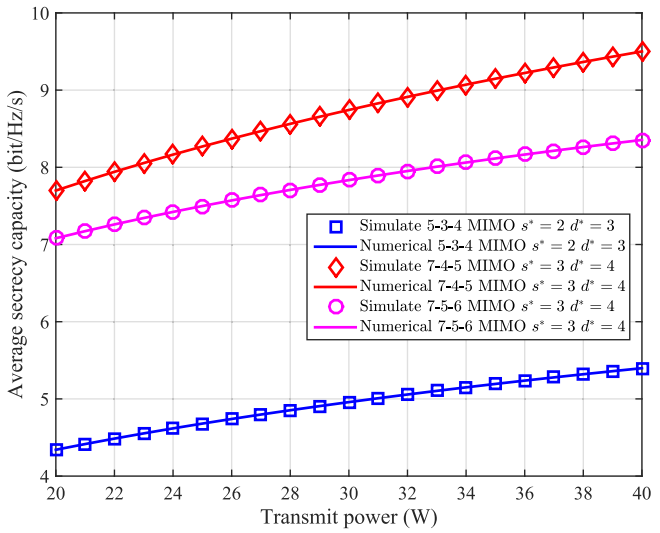


Fig. 7. Average secrecy capacities based on numerical analysis and simulations with the optimal choices of (s, d), where we assume $t : r \leq 2 : 1$ and $r < e < t$.

scenario. It means that when Eve uses more antennas than Bob, an optimal strategy is to allocate more eigen-subchannels for AN signals than the scheme proposed in [16]. Note that our optimal scheme may not help much if abundant transmit antennas are available.

Fig. 5 shows the average secrecy capacities in the scenarios with a small number of transmit antennas, i.e., $t : r : e \leq 2 : 1 : 1$. We considered three scenarios, such as 5-3-3, 7-4-4, and 7-5-5 MIMO scenarios. It also shows that increasing the number of transmit antennas improves secrecy capacities. However, if Bob and Eve use the same number of antennas, increasing the number of receive antennas simultaneously will degrade secrecy capacities as the gains in both main and wiretap channels seem to be enhanced. The reason is that Alice does not have enough spaces to confuse Eve.

Fig. 6 compares the average secrecy capacities with the optimal number of selected message-sending eigen-subchannels and other choices. With comparison to Fig. 4, the results from numerical analysis and simulations in Fig. 6 were obtained for $t : r : e \leq 2 : 1 : 1$. In this case, $s = r$ is no longer the best choice even Bob and Eve use the same number of antennas. The results show that the optimal s^* equals to four instead of $s = 5$, as given in [16]. We also see a massive gap between our scheme and the existing scheme with an approximate 20% difference.

Figs. 7 and 8 show the results in the scenarios with a limited number of transmit antennas with more antennas available at Eve. In this case, we observe that it is almost impossible to make a full use of all eigen-subchannels for sending messages. Especially, in Fig. 8, a comparison of the optimal result with the results from [16] shows that if we insist on using $s = r$, the performance degradation will be approximately 40% compared to the optimal choice.

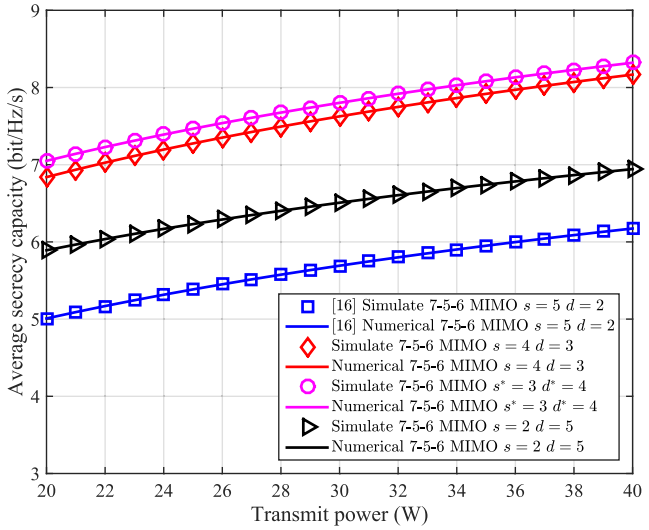


Fig. 8. Average secrecy capacities based on numerical analysis and simulations in terms of different choices of (s, d), where we assume $t : r \leq 2 : 1$ and $r < e < t$.

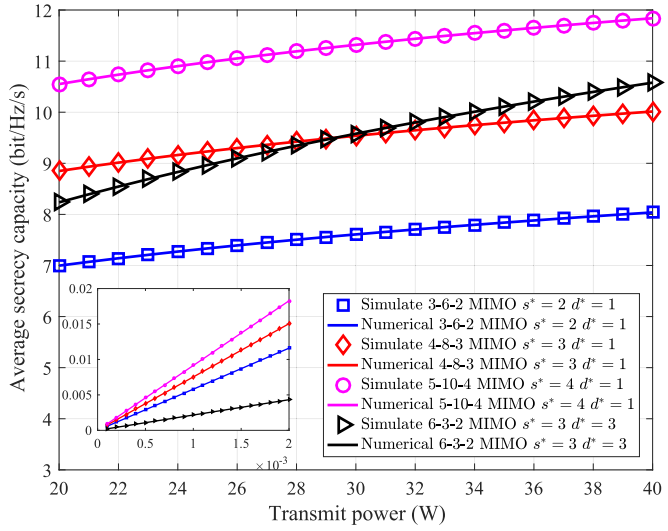


Fig. 9. Average secrecy capacities based on numerical analysis and simulations with the optimal choices of (s, d), where we assume $e = t - 1$ and $r : t = 1 : 2$.

Fig. 9 focuses on the scenarios where the number of transmit antennas is smaller than that of Bob. In this case, the existing schemes [13], [15], [16] are all incapable to ensure security. However, in the proposed scheme, a positive average secrecy capacity is still guaranteed, which approaches to about 25 percent of that with abundant transmit antennas. Especially, in the case of 3-6-2 scenario, compared to the 6-3-2 MIMO scenario, the AN-sending spaces are reduced with a decreasing number of transmit antennas. We can also see that the average secrecy capacities are stable with increasing power in high SNR regions, which means that when $t < r$, increasing transmit power can not bring in much benefit for security. However, the black curve, showing a steeper shape, indicates that transmit power contributes more significantly to the improvement of secrecy capacity than the case of $t < r$.

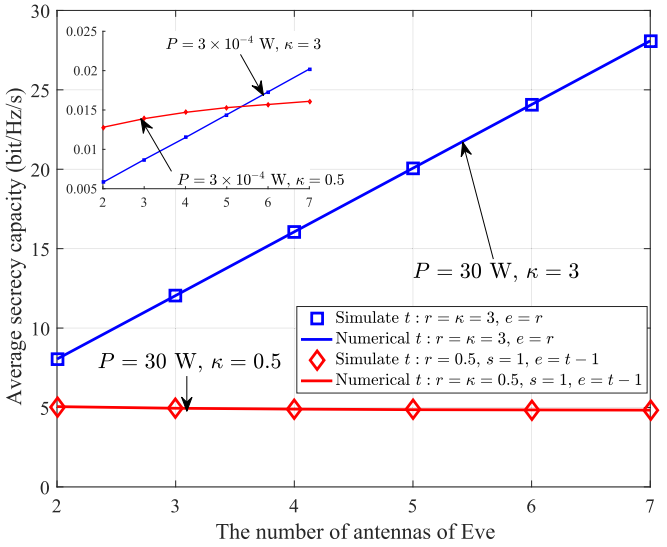


Fig. 10. Average secrecy capacities based on numerical analysis and simulations in terms of the number of Eve' antennas.

Fig. 10 shows the relationship between the average secrecy capacity and the number of Eve's antennas. In the case of adequate transmit antennas, i.e., $t : r = \kappa = 3$ and $e = r$, the average secrecy capacity increases with an increasing number of Eve's antennas (equals to r and $t/2$) in both high and low SNR regions. In the case with a small number of transmit antennas and adequate receive antennas (here we set $t : r = 1/2$, $s = 1$, and $e = t - 1$), we find that the average secrecy capacity converges to a deterministic constant when e becomes large. It means that an increasing number of antennas brings in no benefit when Alice uses less antennas than Bob and selects a transmit diversity scheme. The results conform to the analysis given in Section IV. B.

VI. Conclusions and Further Works

In this paper, we proposed an AN scheme for secure MIMO communications, which improves the average secrecy capacity via searching for an optimal number of message-sending eigen-subchannels and AN-sending spaces. Specifically, we selected the number of message-sending eigen-subchannels based on the first to the kth largest eigenvalues of Wishart matrices, instead of using all valid spaces. We also derived a theoretical average secrecy capacity expression for the proposed AN scheme by marginal pdf of the first to the kth largest eigenvalues of Wishart matrices, and the analysis was done based on different system parameters. Based on the results given in the analysis and simulations, we can make the conclusions as follows.

1) In scenarios with abundant transmit antennas (such as the scenario of $t : r \geq 2 : 1$), an optimal strategy is to take full advantage of all eigen-subchannels. But in scenarios with a limited number of transmit antennas, a better way is to do a search using Theorems 3 and 4.
2) If Alice finds that Bob has more antennas and selects a transmit diversity scheme, it is useless to add more transmit antennas when the ratio of antenna numbers for Alice and Bob is a constant.

Moreover, we compared the secrecy capacities of the proposed scheme with the other schemes in terms of different transmission space allocations. The proposed scheme can improve the secrecy capacity by 20∼40% compared to other schemes, and guarantee a positive secrecy capacity even if the number of transmit antennas is smaller than that of Bob.

In this work, we considered Rayleigh channel only. However, it may also be necessary to study security performance in other channels. In this work, we assumed that transmit and receive antennas are uncorrelated. Nevertheless, real MIMO channels may not be perfect, especially if the separation distances of antennas are much shorter than the wavelength. In addition, we should also exploit to use other Wishart matrices to extend it to the works considering Rican channels and correlated antennas.

APPENDIX I

A. Proof of Theorem 1

We prove Theorem 1 by a similar way shown in [24]. The following two lemmas are needed to complete the proof.

Lemma 2 (Proved in [25, eqs. (4.20) and (4.21)]): For any $f_i(\lambda)$, $i = 1, \ldots, n$, which are probability density functions, we have

$$\sum_i^{\sim} \int_{D_1} \prod_{k=1}^{n} f_{i_k}(\lambda_k) d\lambda_k = \prod_{k=1}^{n} \left[\int_0^x f_k(\lambda) d\lambda \right],$$
$$\sum_i^{\sim} \int_{D_2} \prod_{k=1}^{n} f_{i_k}(\lambda_k) d\lambda_k = \prod_{k=1}^{n} \left[\int_x^{\infty} f_k(\lambda) d\lambda \right], \quad (36)$$

where $D_1 = \{0 < \lambda_1 < \cdots < \lambda_n < x\}$ and $D_2 = \{x < \lambda_1 < \cdots < \lambda_n < \infty\}$, k is the index of an independent variable λ in D_1 or D_2, and $\sum_i^{\sim}$ denotes the summation over all permutations $(i_1, \ldots, i_n)$ of $(1, \ldots, n)$.

Lemma 3 (See in [32, eq. (95)]): The joint pdf of ordered eigenvalues $\lambda_1 > \lambda_2 > \cdots > \lambda_n \geq 0$ of a uncorrelated central Wishart matrix $\mathbf{W} \sim W_n(m, \mathbf{0}_n, \mathbf{I}_n)$ is

$$f_{\boldsymbol{\lambda}}(\boldsymbol{\lambda}) = K^{-1} \prod_{i<j}^{n} (\lambda_i - \lambda_j)^2 \prod_{i=1}^{n} \lambda_i^{m-n} \exp(-\lambda_i), \quad (37)$$

where $K = \prod_{i=1}^{n} (t-i)!(r-i)!$.

Next, we begin to prove Theorem 1 as follows. First, we have

$$\begin{aligned} F_{\lambda_k}(x) &= P(\lambda_k \leq x) \\ &= P(\lambda_{k-1} \leq x) \\ &\quad + P(\lambda_n < \cdots < \lambda_k < x < \lambda_{k-1} < \cdots < \lambda_1) \\ &= P(\lambda_{k-1} \leq x) + p. \end{aligned} \quad (38)$$

Let the domain as $D_3 = \{\lambda_n < \cdots < \lambda_k < x < \lambda_{k-1} < \cdots < \lambda_1\}$. Integrating Eqn. (37), we can get the probability p as

$$p = K^{-1} \int_{D_3} \prod_{i<j}^{n} (\lambda_i - \lambda_j)^2 \prod_{i=1}^{n} \lambda_i^{m-n} \exp(-\lambda_i) d\lambda_i. \quad (39)$$

Note that we have

$$\begin{aligned} &\prod_{i<j}^{n} (\lambda_i - \lambda_j)^2 \\ &= \det \begin{bmatrix} 1 & \cdots & 1 \\ \lambda_1 & \cdots & \lambda_n \\ \vdots & & \vdots \\ \lambda_1^{n-1} & \cdots & \lambda_n^{n-1} \end{bmatrix} \times \det \begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_n & \cdots & \lambda_n^{n-1} \end{bmatrix} \\ &= \det \begin{bmatrix} \sum_{q=1}^{n} \lambda_q^0 & \sum_{q=1}^{n} \lambda_q^1 & \cdots & \sum_{q=1}^{n} \lambda_q^{n-1} \\ \sum_{q=1}^{n} \lambda_q^1 & \sum_{q=1}^{n} \lambda_q^2 & \cdots & \sum_{q=1}^{n} \lambda_q^n \\ \vdots & \vdots & & \vdots \\ \sum_{q=1}^{n} \lambda_q^{n-1} & \sum_{q=1}^{n} \lambda_q^r & \cdots & \sum_{q=1}^{n} \lambda_q^{2n-2} \end{bmatrix} \\ &= \sum_q^{\sim} \det \begin{bmatrix} \lambda_{q_1}^0 & \lambda_{q_2}^1 & \cdots & \lambda_{q_n}^{n-1} \\ \lambda_{q_1}^1 & \lambda_{q_2}^2 & \cdots & \lambda_{q_n}^n \\ \vdots & \vdots & & \vdots \\ \lambda_{q_1}^{n-1} & \lambda_{q_2}^n & \cdots & \lambda_{q_n}^{2n-2} \end{bmatrix} \\ &= \sum_q^{\sim} \sum_{\sigma}^{\sim} (-1)^{\mathrm{per}(\sigma_1, \ldots, \sigma_n)} \prod_{i=1}^{n} \lambda_{q_i}^{i+\sigma_i-2}, \end{aligned} \quad (40)$$

where $\sum_q^{\sim}$ denotes the summation over all permutations $(q_1, \ldots, q_n)$ of $(1, \ldots, n)$, $\sum_{\sigma}^{\sim}$ denotes the summation over all permutations $(\sigma_1, \ldots, \sigma_n)$ of $(1, \ldots, n)$, and $\mathrm{per}(\sigma_1, \ldots, \sigma_n)$ is either 0 or 1, corresponding to even or odd value of the permutation $(\sigma_1, \ldots, \sigma_n)$. Then, Eqn. (39) can be written as

$$\begin{aligned} p &= K^{-1} \int_{D_3} \prod_{i<j}^{n} (\lambda_i - \lambda_j)^2 \prod_{i=1}^{n} \lambda_i^{m-n} \exp(-\lambda_i) d\lambda_i \\ &= K^{-1} \sum_q^{\sim} \sum_{\sigma}^{\sim} \int_{D_3} (-1)^{\mathrm{per}(\sigma_1, \ldots, \sigma_n)} \\ &\quad \times \prod_{i=1}^{n} \lambda_{q_i}^{m-n+i+\sigma_i-2} \exp(-\lambda_{q_i}) d\lambda_{q_i} \\ &= K^{-1} \sum_1 \sum_{\sigma}^{\sim} (-1)^{\mathrm{per}(\sigma_1, \ldots, \sigma_n)} I_1(\mu, \sigma) I_2(\mu, \sigma), \end{aligned} \quad (41)$$

where

$$\sum_q^{\sim} = \sum_1 \sum_{q_{\mu_\psi}}^{\sim} \sum_{q_{\mu_\omega}}^{\sim}. \quad (42)$$

Here, $\sum_{q_{\mu_\psi}}^{\sim}$ denotes the summation over the permutations $(q_{\mu_1}, \ldots, q_{\mu_{k-1}})$ of $(1, \ldots, k-1)$, $\sum_{q_{\mu_\omega}}^{\sim}$ denotes the summation over the permutations $(q_{\mu_k}, \ldots, q_{\mu_n})$ of $(k, \ldots, n)$, $\sum_1$ denotes the summation over the combination of sets $(\mu_1 < \mu_2 < \cdots < \mu_{k-1})$ and $(\mu_k < \mu_{k+1} < \cdots < \mu_n)$, and $(\mu_1, \ldots, \mu_n)$ is a permutation of $(1, \ldots, n)$. From Lemma 1, we can obtain

$$\begin{aligned} I_1(\mu, \sigma) &= \sum_{q_{\mu_\psi}}^{\sim} \int_{D_4} \prod_{i=1}^{k-1} \lambda_{q_{\mu_i}}^{m-n+\mu_i+\sigma_i-2} \exp(-\lambda_{q_{\mu_i}}) d\lambda_{q_{\mu_i}} \\ &= \prod_{i=1}^{k-1} \int_x^{\infty} \lambda_{\mu_i}^{m-n+\mu_i+\sigma_i-2} \exp(-\lambda_{\mu_i}) d\lambda_{\mu_i}, \end{aligned}$$

$$I_2(\mu,\sigma) = \sum_{q_{\mu\omega}}^{\sim} \int_{D_5} \prod_{i=k}^{n} \lambda_{q_{\mu_i}}^{m-n+\mu_i+\sigma_i-2} \exp(-\lambda_{q_{\mu_i}}) d\lambda_{q_{\mu_i}} = \prod_{i=k}^{n} \int_0^x \lambda_{\mu_i}^{m-n+\mu_i+\sigma_i-2} \exp(-\lambda_{\mu_i}) d\lambda_{\mu_i}, \quad (43)$$

where $D_4 = \{x < \lambda_{k-1} < \cdots < \lambda_1 < \infty\}$ and $D_5 = \{0 < \lambda_n < \cdots < \lambda_k < x\}$. σ_i is the ith position after re-ordering $(\sigma_1, \ldots, \sigma_n)$, which can be viewed as the column position of the determinant of an $(n \times n)$ matrix. μ_i is the row position of the determinant of the $(n \times n)$ matrix dependent on k. Hence, we can re-define the order index numbers of rows and columns as μ_u and v. Finally, we get

$$p = K^{-1} \sum\nolimits_1 \det\big[\Theta(\boldsymbol{\mu}, k, x)\big], \quad (44)$$

where a $(n \times n)$ real matrix $\Theta(\boldsymbol{\mu}, k; x)$ is defined in Eqn. (3), and the marginal cdf of the kth largest eigenvalue is derived in Eqn. (2). This proof is completed. ■

B. Proof of Corollary 1

The following lemma is required for proving Corollary 1.

Lemma 4 (Derivative of a determinant [33]): The determinant of an $(n \times n)$ matrix $\mathbf{A}(x)$ is give by

$$\frac{d}{dx} \det\big[\mathbf{A}(x)\big] = \sum_{t=1}^{n} \det\big[\mathbf{A}(t; x)\big], \quad (45)$$

where $\mathbf{A}(t; x)$ coincides with $\mathbf{A}(x)$, except that every entry in the tth row (or column) is differentiated with respect to x.

The marginal pdf of the kth largest eigenvalue can be obtained from its marginal cdf as

$$f_{\lambda_k}(x) = \frac{d}{dx}\left\{ K^{-1} \sum_{i=1}^{k} \sum_{1} \det\big[\Theta(\boldsymbol{\mu}, i; x)\big] \right\} = K^{-1} \sum_{i=1}^{k} \sum_{1} \sum_{j=1}^{n} \det\big[\Omega(\boldsymbol{\mu}, i, j; x)\big], \quad (46)$$

where the $(n \times n)$ real matrix $\Omega(\boldsymbol{\mu}, i, j; x)$ is defined in Eqn. (7). This proof is completed. ■

APPENDIX II

C. Proof of Lemma 1

We first prove $[\mathbf{HB}]^{\dagger}\mathbf{HZ} = \mathbf{0}$. Note that we have

$$[\mathbf{HB}]^{\dagger}\mathbf{HZ} = \mathbf{B}^{\dagger}\mathbf{H}^{\dagger}\mathbf{HZ} = \mathbf{B}^{\dagger}\mathbf{U}\Lambda\mathbf{U}^{\dagger}\mathbf{Z}, \quad (47)$$

where $\Lambda \in \mathbb{R}^{t\times t}$ is a diagonal matrix, $\mathbf{U} \in \mathbb{C}^{t\times t}$ and $\mathbf{U}^{\dagger} \in \mathbb{C}^{t\times t}$ are unitary matrices. Since $\mathbf{B} = [\mathbf{U}]_{(1\sim t),(1\sim s)} \in \mathbb{C}^{t\times s}$ denotes a sub-matrix of $\mathbf{U}$, including the first to the tth rows and the first to the sth columns of $\mathbf{U}$, according to the property of a unitary matrix, different columns of $\mathbf{U}$ are complex orthogonal with each other such that

$$[\mathbf{U}^{\dagger}]_{j,(1\sim t)}[\mathbf{U}]_{(1\sim t),i} = 0, \quad j \neq i, \quad (48)$$

and $[\mathbf{U}^{\dagger}]_{i,(1\sim t)}[\mathbf{U}]_{(1\sim t),i} = 1$. Then, we have $\mathbf{B}^{\dagger} = [\mathbf{U}^{\dagger}]_{(1\sim s),(1\sim t)}$, and get a $s \times t$ matrix as

$$\mathbf{B}^{\dagger}\mathbf{U}\Lambda = [\mathbf{I}_s; \mathbf{0}_{s\times d}]\Lambda = [\Lambda_s; \mathbf{0}_{s\times d}], \quad (49)$$

where $\mathbf{I}_s$ is a $s \times s$ identity matrix, $\mathbf{0}_{s\times d}$ is a $s \times d$ zero matrix, $\Lambda_s \in \mathbb{R}^{s\times s}$ is a diagonal matrix formed by the first to the sth eigenvalues of $\mathbf{H}^{\dagger}\mathbf{H}$. Similarly, we have a $t \times d$ matrix

$$\mathbf{U}^{\dagger}\mathbf{Z} = \begin{bmatrix} \mathbf{0}_{s\times d} \\ \mathbf{I}_d \end{bmatrix}, \quad (50)$$

where $\mathbf{Z} = [\mathbf{U}]_{(1\sim t),(s+1\sim t)} \in \mathbb{C}^{t\times d}$, $\mathbf{I}_d$ is a $d \times d$ identity matrix, and $\mathbf{0}_{s\times d}$ is a $s \times d$ zero matrix. Obviously, we get

$$[\mathbf{HB}]^{\dagger}\mathbf{HZ} = [\Lambda_s; \mathbf{0}_{s\times d}] \times \begin{bmatrix} \mathbf{0}_{s\times d} \\ \mathbf{I}_d \end{bmatrix} = \mathbf{0}_{s\times d}. \quad (51)$$

Hence, Eqn. (47) is a zero matrix no matter if $t \geq r$ or $t < r$. Since $\mathbf{H}$ and $\mathbf{H}_e$ have some different elements, we get $[\mathbf{HB}]^{\dagger}\mathbf{H}_e\mathbf{Z} \neq \mathbf{0}$, and $[\mathbf{H}_e\mathbf{B}]^{\dagger}\mathbf{H}_e\mathbf{Z} \neq \mathbf{0}$. This proof is completed. ■

D. Proof of Theorem 2

Lemma 5 (Proved in [34]): For a given complex Gaussian random matrix $\mathbf{H}_e$ obeying distribution $\mathcal{CN}(\mathbf{0}, \mathbf{I})$ and two fixed unitary matrices $\mathbf{B}$ and $\mathbf{Z}$, we carry through $\mathbf{H}_2 = \mathbf{H}_e\mathbf{B}$ and $\mathbf{H}_3 = \mathbf{H}_e\mathbf{Z}$. Then, both $\mathbf{H}_2$ and $\mathbf{H}_3$ are mutually independent complex Gaussian random matrices with each element obeying distribution $\mathcal{CN}(0, 1)$.

We begin to prove Theorem 2 as follows. We have

$$\begin{aligned} I(\mathbf{w}; \mathbf{y}|\mathbf{H}) &= H(\mathbf{w}|\mathbf{H}) - H(\mathbf{w}|\mathbf{y}, \mathbf{H}) \\ &= H(\mathbf{w}) - H(\mathbf{w}|\mathbf{y}) \\ &= H(\mathbf{y}) - H(\mathbf{y}|\mathbf{w}) \\ &= H(\mathbf{y}) - H\big[(\mathbf{H}_1\mathbf{w} + \mathbf{HZv} + \mathbf{n})|(\mathbf{H}_1\mathbf{w})\big] \\ &= H(\tilde{\mathbf{y}}) - H\big[(\Lambda_s\mathbf{w} + \tilde{\mathbf{n}})|(\Lambda_s\mathbf{w})\big]. \end{aligned} \quad (52)$$

Since $\Lambda_s\mathbf{w}$ and $\tilde{\mathbf{n}}$ are independent, Eqn. (52) can be expressed as

$$\begin{aligned} I(\mathbf{w}; \mathbf{y}|\mathbf{H}) &= H(\tilde{\mathbf{y}}) - H\big[(\Lambda_s\mathbf{w} + \tilde{\mathbf{n}})|(\Lambda_s\mathbf{w})\big] \\ &= H(\tilde{\mathbf{y}}) - H(\tilde{\mathbf{n}}) \\ &\leq \mathrm{E}\left[\log 2 \det\Big(\pi \mathrm{e}\big((P/t)\Lambda_s^2 + \Lambda_s\big)\Big)\right] \\ &\quad - \mathrm{E}\big[\log 2 \det(\pi \mathrm{e}\Lambda_s)\big] \\ &= \mathrm{E}\big[\log_2 \det\big(\mathbf{I}_s + (P/t)\mathbf{H}_1^{\dagger}\mathbf{H}_1\big)\big] \\ &= \mathrm{E}\big[\log_2 \det\big(\mathbf{I}_r + (P/t)\mathbf{H}_1\mathbf{H}_1^{\dagger}\big)\big], \end{aligned} \quad (53)$$

which gives a desirable value if and only if $\mathbf{w}$ is a circularly symmetric complex Gaussian vector [20, Le. 2]. Here, e is the Euler's number.

Similarly, we have

$$\begin{aligned} &I\big[\mathbf{w}; \mathbf{y}_e|(\mathbf{H}, \mathbf{H}_e)\big] \\ &\quad = H(\mathbf{y}_e) - H\big[\mathbf{y}_e|\mathbf{w}\big] \\ &\quad = H(\mathbf{y}_e) - H\big[(\mathbf{H}_2\mathbf{w} + \mathbf{H}_3\mathbf{v} + \mathbf{n}_e)|(\mathbf{H}_2\mathbf{w})\big]. \end{aligned} \quad (54)$$

Since $\mathbf{H}_2$, $\mathbf{H}_3$, and $\mathbf{n}_e$ are mutually independent, $\mathbf{H}_2\mathbf{w}+\mathbf{H}_3\mathbf{v}+\mathbf{n}_e$ and $\mathbf{H}_2\mathbf{w}$ are independent [20, Les. 3 and 4], Eqn. (54) can be expressed as

$$I\big[\mathbf{w};\mathbf{y}_e|(\mathbf{H},\mathbf{H}_e)\big] = H(\mathbf{y}_e) - H(\mathbf{H}_3\mathbf{v}+\mathbf{n}_e) \geq \mathrm{E}\bigg[\log_2\det\bigg(\mathbf{I}_e + \frac{(P/t)\mathbf{H}_2\mathbf{H}_2^\dagger}{(P/t)\mathbf{H}_3\mathbf{H}_3^\dagger+\mathbf{I}_e}\bigg)\bigg], \tag{55}$$

which yields a desirable value if and only if $\mathbf{v}$ is a circularly symmetric complex Gaussian vector [20, Le. 2]. $\mathbf{w}$ is reserved as a circularly symmetric complex Gaussian vector (for entropy maximization). The AN signal $\mathbf{v}$ is an artificial vector, each element of which has real and imaginary parts obeying a standard normal distribution and these elements are uncorrelated with each other.

Then, we can derive

$$\begin{aligned}\tilde{C}_s &\leq I(\mathbf{w};\mathbf{y}|\mathbf{H}) - I\big[\mathbf{w};\mathbf{y}_e|(\mathbf{H},\mathbf{H}_e)\big]\\ &= \mathrm{E}\big[\log_2\det(\mathbf{I}_r + (P/t)\mathbf{H}_1\mathbf{H}_1^\dagger)\big]\\ &\quad - \mathrm{E}\bigg[\log_2\det\bigg(\mathbf{I}_e + \frac{(P/t)\mathbf{H}_2\mathbf{H}_2^\dagger}{(P/t)\mathbf{H}_3\mathbf{H}_3^\dagger+\mathbf{I}_e}\bigg)\bigg],\end{aligned} \tag{56}$$

which holds if and only if $\mathbf{v}$ is a circularly symmetric complex Gaussian vector. And it is an inherent premise that $\mathbf{w}$ is a circularly symmetric complex Gaussian vector. This completes the proof. ■

E. Proof of Theorem 3

Recalling the average secrecy capacity, we have

$$\begin{aligned}\tilde{C}_s &= \mathrm{E}\big[\log_2\det(\mathbf{I}_r + (P/t)\mathbf{H}_1\mathbf{H}_1^\dagger)\big]\\ &\quad - \mathrm{E}\bigg[\log_2\det\bigg(\mathbf{I}_e + \frac{(P/t)\mathbf{H}_2\mathbf{H}_2^\dagger}{(P/t)\mathbf{H}_3\mathbf{H}_3^\dagger+\mathbf{I}_e}\bigg)\bigg].\end{aligned} \tag{57}$$

The second term in the above expression can be rewritten as

$$\begin{aligned}&\mathrm{E}\bigg[\log_2\det\bigg(\mathbf{I}_e + \frac{(P/t)\mathbf{H}_2\mathbf{H}_2^\dagger}{(P/t)\mathbf{H}_3\mathbf{H}_3^\dagger+\mathbf{I}_e}\bigg)\bigg]\\ &= \mathrm{E}\bigg[\log_2\det\bigg(\frac{(P/t)\mathbf{H}_3\mathbf{H}_3^\dagger + (P/t)\mathbf{H}_2\mathbf{H}_2^\dagger+\mathbf{I}_e}{(P/t)\mathbf{H}_3\mathbf{H}_3^\dagger+\mathbf{I}_e}\bigg)\bigg]\\ &= \mathrm{E}_{\mathbf{H}_4}\big[\log_2\det\big(\mathbf{I}_e + (P/t)\mathbf{H}_4\mathbf{H}_4^\dagger\big)\big]\\ &\quad - \mathrm{E}_{\mathbf{H}_3}\big[\log_2\det\big(\mathbf{I}_e + (P/t)\mathbf{H}_3\mathbf{H}_3^\dagger\big)\big]\\ &= C(\mathbf{H}_4,(P/t)) - C(\mathbf{H}_3,(P/t)),\end{aligned} \tag{58}$$

where $\mathbf{H}_4 = [\mathbf{H}_2,\mathbf{H}_3] = \mathbf{H}_e\mathbf{U}$. And we have

$$\mathbf{H}_2\mathbf{H}_2^\dagger + \mathbf{H}_3\mathbf{H}_3^\dagger = \mathbf{H}_4\mathbf{H}_4^\dagger. \tag{59}$$

From the conclusion made in [21, Th. 1], we can obtain $C(\mathbf{H}_4,(P/t))$ and $C(\mathbf{H}_3,(P/t))$, which are defined in Eqn. (21).

Similarly, since $s \geq \min(t,r)$ and $\mathbf{H}_1 = \mathbf{HB}$, we can also get the first term of Eqn. (57) as

$$\mathrm{E}[\log_2\det(\mathbf{I}_r + (P/t)\mathbf{H}_1\mathbf{H}_1^\dagger)] = C(\mathbf{H},(P/t)). \tag{60}$$

Its close-form expression can also be derived from [21, Th. 1], which is defined in Eqn. (21). Thus, we get

$$\begin{aligned}&\tilde{C}_s(\mathbf{w},\mathbf{v},\mathbf{H},\mathbf{H}_e)\\ &\quad = C(\mathbf{H},(P/t)) + C(\mathbf{H}_3,(P/t)) - C(\mathbf{H}_4,(P/t)).\end{aligned} \tag{61}$$

The proof is completed. ■

F. Proof of Theorem 4

Let us focus on the derivation of $\Upsilon(\mathbf{H},(P/t),s)$. We have

$$\Upsilon(\mathbf{H},(P/t),s) = \sum_{k=1}^{s}\int_0^\infty \log_2(1+(P/t)x)\,f_{\lambda_k}(x)dx. \tag{62}$$

For any k, the marginal pdf $f_{\lambda_k}(x)$ of the kth largest eigenvalue λ_k of an uncorrelated central Wishart matrix $\mathbf{W} \sim W_n(m,\mathbf{0}_n,\mathbf{I}_n)$ is defined in Corollary 1.

Hence, if we chose the first to the sth largest eigenvalues, we have

$$\begin{aligned}&\Upsilon(\mathbf{H},(P/t),s)\\ &= \int_0^\infty \log_2(1+(P/t)x)\,f_{\lambda_1}(x)\\ &\quad + \cdots + \int_0^\infty \log_2(1+(P/t)x)\,f_{\lambda_s}(x)\\ &= K^{-1}\sum_{k=1}^{s}\int_0^\infty \log_2(1+(P/t)x)\\ &\quad \times \sum_{i=1}^{k}\sum_{1}\sum_{j=1}^{n}\det[\Omega(\boldsymbol{\mu},i,j;x)]dx\\ &= K^{-1}\sum_{k=1}^{s}\sum_{i=1}^{k}\sum_{1}\sum_{j=1}^{n}\\ &\quad \times \int_0^\infty \log_2(1+(P/t)x)\Omega(\boldsymbol{\mu},i,j;x)]dx,\end{aligned} \tag{63}$$

where $\sum_1$ denotes the summation over the combinations $(\mu_1 < \mu_2 < \cdots < \mu_{k-1})$ and $(\mu_k < \mu_{k+1} < \cdots < \mu_n)$, and $(\mu_1,\ldots,\mu_n)$ is a permutation of $(1,\ldots,n)$. The $(n\times n)$ real matrix $\boldsymbol{\Omega}(\boldsymbol{\mu},i,j;x)$ is defined in Eqn. (7), where $u,v = 1,\ldots,n$. And $C_w(\mathbf{H}_4,(P/t))$ and $C_w(\mathbf{H}_3,(P/t))$ are defined in Eqn. (21). The proof is completed. ■

G. Proof of Theorem 5

We have

$$C[\mathbf{A},(P/t)] = \textstyle\sum_{i=1}^{n}\mathrm{E}[\log_2(1+(P/t)\lambda_i(\mathbf{A}))], \tag{64}$$

$$\Upsilon\big[\mathbf{A},(P/t),s\big] = \textstyle\sum_{i=1}^{s}\mathrm{E}[\log_2(1+(P/t)\lambda_i(\mathbf{A}))], \tag{65}$$

where $\lambda_1(\mathbf{A}) > \lambda_2(\mathbf{A}) > \ldots > \lambda(\mathbf{A})_n$ are the ordered eigenvalues of $\mathbf{AA}^\dagger$.

The approximation equations [18, eq. (5.15)] are

$$\begin{cases}\log_2(1+\upsilon) \approx \upsilon\log_2(\mathrm{e}) & \upsilon \approx 0,\\ \log_2(1+\upsilon) \approx \log_2\upsilon & \upsilon \gg 1,\end{cases} \tag{66}$$

where e is the Euler's number.

From [35, eq. (23)], we get

$$\sum_{i=1}^{n} \log_2(\mathrm{E}[\lambda_i(\mathbf{A})]) \approx \log_2 \frac{m!}{(m-n)!}. \quad (67)$$

In high SNR regions, we use Eqns. (66) and (67) to simplify $C[\mathbf{A}, (P/t)]$ as

$$\begin{aligned} C[\mathbf{A}, (P/t)] &= \sum_{i=1}^{n} \mathrm{E}[\log_2(1 + (P/t)\lambda_i(\mathbf{A}))] \\ &\approx n * \log_2(P/t) + \sum_{i=1}^{n} \mathrm{E}[\log_2 \lambda_i(\mathbf{A})] \\ &\leq n * \log_2(P/t) + \sum_{i=1}^{n} \log_2(\mathrm{E}[\lambda_i(\mathbf{A})]) \\ &= n * \log_2(P/t) + \log_2 \frac{m!}{(m-n)!}. \end{aligned} \quad (68)$$

$C[\mathbf{A}, (P/t)]$ increases with n [20]. Then, with the help of Eqn. (66) again, $\Upsilon\big[\mathbf{A}, (P/t), s\big]$ can be simplified as

$$\begin{aligned} \Upsilon\big[\mathbf{A}, (P/t), s\big] &= \sum_{i=1}^{s} \mathrm{E}[\log_2(1 + (P/t)\lambda_i(\mathbf{A}))] \\ &\approx s * \log_2(P/t) + \sum_{i=1}^{s} \mathrm{E}[\log_2 \lambda_i(\mathbf{A})]. \end{aligned} \quad (69)$$

Finally, we get $\tilde{C}_s^H$ as

$$\tilde{C}_s^H = (x + y - e)\log_2(P/t) + \chi_1 + \chi_2 - \chi_3, \quad (70)$$

where χ_1, χ_2, χ_3, x, and y are defined in Eqns. (28)-(31).

In low SNR regions, we use Eqn. (66) to simplify $C[\mathbf{A}, (P/t)]$ and $\Upsilon\big[\mathbf{A}, (P/t), s\big]$ as

$$\begin{aligned} C[\mathbf{A}, (P/t)] &= \sum_{i=1}^{n} \mathrm{E}[\log_2(1 + (P/t)\lambda_i(\mathbf{A}))] \\ &\approx \sum_{i=1}^{n} (P/t)\mathrm{E}[\lambda_i(\mathbf{A})]\log_2(\mathrm{e}) \\ &= (P/t)\mathrm{E}[\mathrm{Tr}(\mathbf{A}\mathbf{A}^{\dagger})]\log_2(\mathrm{e}) \\ &= \frac{mnP}{t}\log_2(\mathrm{e}), \end{aligned} \quad (71)$$

and

$$\begin{aligned} \Upsilon\big[\mathbf{A}, (P/t), s\big] &= \sum_{i=1}^{s} \mathrm{E}[\log_2(1 + (P/t)\lambda_i(\mathbf{A}))] \\ &\approx \sum_{i=1}^{s} (P/t)\mathrm{E}[\lambda_i(\mathbf{A})]\log_2(\mathrm{e}), \end{aligned} \quad (72)$$

respectively. Then, $\tilde{C}_s^L$ can be expressed as

$$\begin{aligned} \tilde{C}_s^L &= \Big(\frac{\sum_{i=1}^{x} \mathrm{E}[\lambda_i(\mathbf{H})] + e(t-s) - e*t}{t}\Big) P\log_2(\mathrm{e}) \\ &= \Big(\frac{\sum_{i=1}^{x} \mathrm{E}[\lambda_i(\mathbf{H})] - s*e}{t}\Big) P\log_2(\mathrm{e}). \end{aligned} \quad (73)$$

The proof is completed. ■

REFERENCES

[1] B. Schneier, *Applied Cryptography: Protocols, Algorithms, and Source Code in C*. Indianapolis, IN, USA: Wiley, 1996.

[2] B. Schneier, "Description of a new variable-length key, 64-bit block cipher (Blowfish)," in *Proc. Fast Softw. Encryption Cambridge Secur. Workshop*, 1994, pp. 191–204.

[3] D. Johnson, A. Menezes, and S. Vanstone, "The elliptic curve digital signature algorithm (ECDSA)," *Int. J. Inf. Secur.*, vol. 1, no. 1, pp. 36–63, Aug. 2001.

[4] F. Zhang, N. R. Safavi, and W. Susilo, "An efficient signature scheme from bilinear pairings and its applications," in *Proc. Int. Workshop Public Key Cryptogr. (PKC)*, vol. 2947. 2004, pp. 277–290.

[5] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1367, Oct. 1975.

[6] S. Yan, N. Yang, R. Malaney, and J. Yuan, "Transmit antenna selection with alamouti coding and power allocation in MIMO wiretap channels," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1656–1667, Mar. 2014.

[7] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1988.

[8] Y. Huang, F. S. Al-Qahtani, T. Q. Duong, and J. Wang, "Secure transmission in MIMO wiretap channels using general-order transmit antenna selection with outdated CSI," *IEEE Trans. Commun.*, vol. 63, no. 8, pp. 2959–2971, Aug. 2015.

[9] J. Xiong, Y. Tang, D. Ma, P. Xiao, and K.-K. Wong, "Secrecy performance analysis for TAS-MRC system with imperfect feedback," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 8, pp. 1617–1629, Aug. 2015.

[10] N. S. Ferdinand, D. B. da Costa, and M. Latva-Aho, "Physical layer security in MIMO OSTBC line-of-sight wiretap channels with arbitrary transmit/receive antenna correlation," *IEEE Wireless Commun. Lett.*, vol. 2, no. 5, pp. 467–470, Oct. 2013.

[11] J. Hu, Y. Cai, N. Yang, and W. Yang, "A new secure transmission scheme with outdated antenna selection," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 11, pp. 2435–2446, Nov. 2015.

[12] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2180–2189, Jun. 2008.

[13] D. W. K. Ng, E. S. Lo, and R. Schober, "Secure resource allocation and scheduling for OFDMA decode-and-forward relay networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 10, pp. 3528–3540, Oct. 2011.

[14] M. R. A. Khandaker and K.-K. Wong, "Masked beamforming in the presence of energy-harvesting eavesdroppers," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 1, pp. 40–54, Jan. 2015.

[15] S.-H. Tsai and H. V. Poor, "Power allocation for artificial-noise secure MIMO precoding systems," *IEEE Trans. Signal Process.*, vol. 62, no. 13, pp. 3479–3493, Jul. 1, 2014.

[16] S. Liu, Y. Hong, and E. Viterbo, "Artificial noise revisited," *IEEE Trans. Inf. Theory*, vol. 61, no. 7, pp. 3901–3911, Jul. 2015.

[17] M. Schulz, A. Loch, and M. Hollick, "Practical known-plaintext attacks against physical layer security in wireless MIMO systems," in *Proc. Symp. NDSS*, San Diego, CA, USA, 2014, pp. 1–13.

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

[19] A. Zanella, M. Chiani, and M. Z. Win, "On the marginal distribution of the eigenvalues of Wishart matrices," *IEEE Trans. Commun.*, vol. 57, no. 4, pp. 1050–1060, Apr. 2009.

[20] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 586–595, 1999.

[21] H. Shin and J. H. Lee, "Closed-form formulas for ergodic capacity of MIMO Rayleigh fading channels," in *Proc. IEEE Conf. Commun.*, Anchorage, AK, USA, May 2003, pp. 2996–3000.

[22] T. K. Y. Lo, "Maximum ratio transmission," in *Proc. IEEE Conf. Commun.*, Vancouver, BC, Canada, 1999, pp. 1310–1314.

[23] C. S. Park and K. B. Lee, "Statistical transmit antenna subset selection for limited feedback MIMO systems," in *Proc. Asia–Pacific Conf. Commun.*, Busan, South Korea, Aug. 2006, pp. 1–5.

[24] C. G. Khatri, "Distribution of the largest or the smallest characteristic root under null hypothesis concerning complex multivariate normal populations," *Ann. Math. Statist.*, vol. 35, no. 4, pp. 1807–1810, 1964.

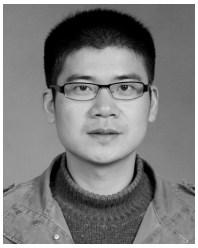
[25] T. Ratnarajah, "Topics in complex random matrices and information theory," M.S. thesis, Dept. Mathematics Stat., Univ. Ottawa, Ottawa, ON, Canada, May 2003.

[26] L. G. Ordóñez, D. P. Palomar, and J. R. Fonollosa, "Ordered eigenvalues of a general class of hermitian random matrices with application to the performance analysis of MIMO systems," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 672–689, Feb. 2009.

[27] S. Liu, Y. Hong, and E. Viterbo, "Practical secrecy using artificial noise," *IEEE Commun. Lett.*, vol. 17, no. 7, pp. 1483–1486, Jul. 2013.
[28] M.-A. Khalighi, J. Brossier, G. Jourdain, and K. Raoof, "Water filling capacity of Rayleigh MIMO channels," in *Proc. IEEE Int. Symp. Pers., Indoor Mobile Radio Commun.*, San Diego, CA, USA, Sep. 2001, pp. 155–158.
[29] W. Tranter, K. Shanmugan, T. Rappaport, and K. Kosbar, *Principles of Communication Systems Simulation With Wireless Application*. Englewood Cliffs, NJ, USA: Prentice-Hall, 2003.
[30] P. K. Gopala, L. Lai, and H. El Gamal, "On the secrecy capacity of fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4687–4698, Oct. 2008.
[31] R. Couillet and M. Debbah, *Random Matrix Methods for Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2011.
[32] A. T. James, "Distributions of matrix variates and latent roots derived from normal samples," *Ann. Math. Statist.*, vol. 35, no. 2, pp. 475–501, 1964.
[33] J. G. Christiano and J. E. Hall, "On the n-th derivative of a determinant of the j-th Order," *Math. Mag.*, vol. 37, no. 4, pp. 215–217, 1964.
[34] E. Lukacs and E. P. King, "A property of the normal distribution," *Ann. Math. Statist.*, vol. 25, no. 2, pp. 389–394, 1954.
[35] H. Shin and J. H. Lee, "Capacity of multiple-antenna fading channels: Spatial fading correlation, double scattering, and keyhole," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2636–2647, Oct. 2003.

Yiliang Liu received the B.E. and M.Sc. degrees in computer science and communication engineering from Jiangsu University, Zhenjiang, China, in 2012 and 2015, respectively. He is currently pursuing the Ph.D. degree with the Communication Research Center, Harbin Institute of Technology, China. He was a Visiting Research Student with the Department of Engineering Science, National Cheng Kung University, Tainan, Taiwan, from 2014 to 2015. His research interests include security for wireless networks, and vehicular ad hoc network.

Hsiao-Hwa Chen (S'89–M'91–SM'00–F'10) received the B.Sc. and M.Sc. degrees from Zhejiang University, China, in 1982 and 1985, respectively, and the Ph.D. degree from the University of Oulu, Finland, in 1991. He is currently a Distinguished Professor with the Department of Engineering Science, National Cheng Kung University, Taiwan. He is the founding Editor-in-Chief of *Security and Communication Networks Journal* (Wiley). He is a Fellow of the IET and an elected Member at Large of the IEEE ComSoc. He is a recipient of the 2016 IEEE Jack Neubauer Memorial Award. He served as the Editor-in-Chief of IEEE WIRELESS COMMUNICATIONS from 2012 to 2015.



Liangmin Wang received the B.S. degree in computational mathematics from Jilin University, Changchun, China, in 1999, and the Ph.D. degree in cryptology from Xidian University, Xi'an, China, in 2007. He is currently a Full Professor with the School of Computer Science and Communication Engineering, Jiangsu University, Zhenjiang, China. He has authored over 60 technical papers at premium international journals and conferences. His research interests include security protocols and Internet of Things. He has been honored as a Wan-Jiang Scholar of Anhui Province since 2013. He is a member of the ACM and a Senior Member of Chinese Computer Federation. He has severed as the TPC of many IEEE conferences, such as IEEE ICC, IEEE HPCC, and IEEE TrustCOM. He is currently an Associate Editor of *Security and Communication Networks*.