Pilot Randomization to Protect MIMO Secret Key Generation Systems Against Injection Attacks

Thuy M. Pham, Miroslav Mitev, Arsenia Chorti, and Gerhard P. Fettweis

***Abstract*—In this paper, we investigate the problem of secret key generation under an injection attack, which refers to tampering of pilot signals over the air so that part of the shared randomness observed at the legitimate parties is controlled by the adversary. It has been shown that to launch such an attack, an adversary only needs one extra antenna, compared to the legitimate parties, in a single input single output (SISO) network. In this work, we generalize this result for the multiple input multiple output (MIMO) case. Furthermore, we propose pilot randomization as a means to protect against injection attacks by reducing them to jamming attacks that constitute a less serious threat. Finally, we derive a closed-form expression for the secret key rate of the investigated MIMO setting.**

***Index Terms*—secret key generation, jamming attack, injection attack, randomization, MIMO.**

I. Introduction

Physical layer security (PLS) covers a multitude of technologies ranging from the keyless transmission of confidential messages, to secret key generation (SKG) / distribution and authentication using physical unclonable functions and RF fingerprints, among others [1], [2]. The key premise of PLS is to exploit specific properties of wireless channels to build adaptive, low latency and low footprint security controls for sixth generation (6G) systems. However, in addition to studying performance bounds, proper threat models need to be investigated before PLS can be incorporated in 6G security standards.

In this paper, we focus on injection attacks against SKG systems, which fall in the general category of man-in-the-middle attack. An injection attack takes place during the pilot exchange phase [3]–[5] of SKG, i.e., advantage distillation. As shown for single input single output systems (SISO), an attacker can under certain conditions inject the same signal to both legitimate parties, and thus partially control their observations. The impact of such an attack can be detrimental as the adversary can control part of the observed shared randomness and therefore be able to launch a brute force attack at the *input* of the privacy amplification stage. An injection attack in essence reduces the effective size of the key space; note that such attacks have been experimentally demonstrated [6].

Previous studies of injection attacks either consider SISO settings for the legitimate users or reduce the multiple input multiple output (MIMO) case to multiple single cases [7]. In this work, we study a general MIMO setting, and propose a pilot randomization scheme to alleviate this attack and derive the achievable secret key rate. Additionally, we investigate the impact of the number of antennas and the power of the injected signals on the system performance. Our results show that a MIMO setting under an injection attack can still provide high key rates by exploiting MIMO diversity.

The rest of the paper is organized as follows: Section II introduces the system model and explains the attack, Section III presents the proposed countermeasures and performance analysis along with numerical results, while Section IV discusses the paper's findings and conclusions.

Notation: Bold lower and upper case letters represent vectors and matrices, respectively. $\mathbf{I}$ and $\mathbf{0}$ denote an identity matrix and null matrix, respectively; $\mathbf{G}^T$ and $\mathbf{G}^H$ denote the transpose and transpose conjugate of $\mathbf{G}$; $\mathbb{E}(\cdot)$ denotes the expectation of a random variable, $\otimes$ denotes the Kronecker product and $\mathbb{C}$ stands for the set of complex numbers, $|\mathbf{G}|$ stands for the determinant of $\mathbf{G}$, while all logarithms are assumed base 2; $\mathrm{tr}(\mathbf{G})$ denotes the trace of $\mathbf{G}$.

II. System and Threat Model

A. System Model

We follow the convention in which the system consists of two legitimate users – referred to as Alice and Bob – and a malicious node, referred to as Mallory (man-in-the-middle). We assume Alice and Bob, each has N antennas, while Mallory has $M > N$ antennas. In the following, we use the superscript $(\cdot)^A$ and $(\cdot)^B$ to relate a quantity to that of Alice and Bob, respectively. The legitimate users transmit pilot signals $\mathbf{Q}^A$ and $\mathbf{Q}^B$, respectively, to enact advantage distillation for SKG. If no active attacks are considered, the estimated channel matrices at Alice and Bob are given by:

$$\mathbf{Y}^B = \mathbf{H}\mathbf{Q}^A + \mathbf{Z}^B, \tag{1}$$

$$\mathbf{Y}^A = (\mathbf{Q}^B)^T\mathbf{H} + \mathbf{Z}^A, \tag{2}$$

where $\mathbf{Q}^A, \mathbf{Q}^B \in \mathbb{C}^{N\times N}$ are pilots transmitted by Alice and Bob, respectively, $\mathbf{Z}^A, \mathbf{Z}^B \in \mathbb{C}^{N\times N}$ denote noise matrices at the corresponding users and $\mathbf{H} \in \mathbb{C}^{N\times N}$ is the channel matrix between Alice and Bob. In the following, we assume that $\mathbf{H}$ is a full-rank channel matrix following complex standard normal distribution, i.e., $\mathbf{H} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$, thus the marginal distribution of each element is $f_X(h) = \frac{1}{\sqrt{\pi}} e^{-|h|^2}$. Note that the secret key rate for this scenario has been investigated in [8]. In this paper, we focus on the case in which Mallory injects the same signal in the links to Alice and Bob as depicted in Fig. 1; the conditions that render this attack feasible are discussed in the following subsection. In the presence of an injection attack, the observed channel matrices at Alice and Bob can be expressed

The authors are with Barkhausen Institut gGmbH, Dresden, Germany (email: {minhthuy.pham, miroslav.mitev, arsenia.chorti, gerhard.fettweis}@barkhauseninstitut.org). A. Chorti is also with ETIS UMR 8051, CY Paris Cergy University, ENSEA, CNRS, France. This work is financed on the basis of the budget passed by the Saxon State Parliament. The work of A. Chorti is supported by INEX funding project PHEBE.

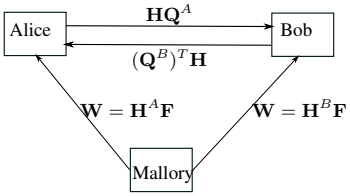

Fig. 1: System model.

as follows

$$\mathbf{Y}^B = \mathbf{HQ}^A + \mathbf{W} + \mathbf{Z}^B, \tag{3}$$

$$\mathbf{Y}^A = (\mathbf{Q}^B)^T\mathbf{H} + \mathbf{W} + \mathbf{Z}^A, \tag{4}$$

where $\mathbf{W}$ is the injected signal.

B. Feasibility of Injection Attack

Considering the severity of this type of attack, we prove its feasibility in the following lemma.

Lemma 1. If Mallory has $M > N$ antennas and can estimate the channel state information in the links Mallory – Alice and Mallory – Bob, she can always design a proper precoding matrix so that

$$\mathbf{W} = \mathbf{H}^A\mathbf{F} = \mathbf{H}^B\mathbf{F}, \tag{5}$$

where $\mathbf{H}^A, \mathbf{H}^B \in \mathbb{C}^{N\times M}$ are the channel matrices between Mallory and Alice and Bob, respectively, and $\mathbf{F} \in \mathbb{C}^{M\times N}$ is Mallory's precoding matrix. For the proof of this lemma, we refer the reader to Appendix A.

The key in the proof of this lemma, is to express the precoding matrix as $\mathbf{F} = [\bar{\mathbf{F}}^T \quad \hat{\mathbf{F}}^T]$ where $\hat{\mathbf{F}} \in \mathbb{C}^{(M-N)\times N}$ is fixed and $\bar{\mathbf{F}}$ is a function of $\hat{\mathbf{F}}$, i.e., $\bar{\mathbf{F}} = f(\hat{\mathbf{F}})$ which leads to the remark below.

Remark 1. Mallory needs at least $N + 1$ antennas to perform an injection attack.

From the above remark, we conclude that the feasibility of injection attacks is trivial (one extra antenna). Note that a similar observation was also reported in [5]. It could be straightforward for legitimate users to deploy more antennas than Mallory to shield the system from injection attacks. However, information on the number of the malicious nodes' antennas cannot be considered known in general, and more importantly, PLS cannot rely on placing hardware constraints on adversarial nodes. In the following, we investigate a pilot randomization approach to minimize the impact of such an attack and derive the achievable secret key rates in this scenario.

III. Countermeasures and Performance Analysis

In this section, we will describe the method to counter an injection attack and derive the corresponding secret key rate.

A. Pilot Randomization

SKG exploits the fact that Alice and Bob observations are correlated. With the injection of $\mathbf{W}$, the same is true for the adversary who now controls part of the shared randomness. As a countermeasure, we generalize the pilot randomization approach that was proposed in [3], [5], relying on pre- and post-multiplication with locally generated random pilot matrices $\mathbf{Q}^A$ and $\mathbf{Q}^B$ at Alice and Bob, respectively. This method will thus reduce an injection attack to a jamming attack and common signals used for SKG are legitimate-user-based signals. More specifically, we premultiply (3) by a random matrix $\mathbf{Q}_B^T$ to obtain an observation

$$\tilde{\mathbf{Y}}^B = (\mathbf{Q}^B)^T\mathbf{Y}^B = (\mathbf{Q}^B)^T\mathbf{HQ}^A + (\mathbf{Q}^B)^T\mathbf{W} + (\mathbf{Q}^B)^T\mathbf{Z}^B. \tag{6}$$

By postmultiplying (4) by $\mathbf{Q}_A$, we also achieve

$$\tilde{\mathbf{Y}}^A = \mathbf{Y}^A\mathbf{Q}^A = (\mathbf{Q}^B)^T\mathbf{HQ}^A + \mathbf{WQ}^A + \mathbf{Z}^A\mathbf{Q}^A. \tag{7}$$

The elements of the randomized pilots $\mathbf{Q}^A$ and $\mathbf{Q}^B$ are chosen to be zero-mean so that $(\mathbf{Q}^B)^T\mathbf{W}, \mathbf{WQ}^A, \mathbf{W}$ are uncorrelated. More specifically, we can draw the real and complex parts of the matrices's elements from a Rademacher distribution [9], with probability mass function given by

$$f_Q(q) = \frac{1}{2}\delta(q+1) + \frac{1}{2}\delta(q-1). \tag{8}$$

In the following, we also suppose that the elements of the channels, noise, and pilot randomization matrices are independent. It is worth noting that the precoding matrix $\mathbf{F}$ in the preceding section is not unique. Thus, we can consider the worst-case scenario that a sophisticated adversary can design its precoding so that $\mathbf{W}$ follows a zero-mean complex normal distribution with variance σ_W^2, i.e., $\mathbf{W} \sim \mathcal{CN}(\mathbf{0}, \sigma_W^2\mathbf{I})$.

B. Secret Key Rate in Multiple Input Single Output (MISO) system

It is worth noting that the majority of existing studies such as [3], [5] consider single antenna setting for legitimate users. Surprisingly, the secret key rate under the injection attack with this simple setting is still missing. For this special case, we can derive the secret key rate as follows.

Lemma 2. Utilizing the pilot randomization method, the secret key rate of the MISO system under an injection attack is given by

$$I_K = \log\left(1 + \frac{1}{2\sigma_{IN}^2 + \sigma_{IN}^4}\right), \tag{9}$$

where $\sigma_{IN}^2 = \sigma_W^2 + \sigma_Z^2$ is the interference-plus-noise power.

Interested readers can refer to Appendix B for the details of the lemma. The formulation for the general case, i.e., MIMO is derived in the following.

C. Secret Key Rate in MIMO System

In order to derive the achievable secret key rate, we need to investigate the common properties of the channel matrices. More specifically, we first consider the left product of the first terms of both equivalent channel matrices $\tilde{\mathbf{Y}}^A$ and $\tilde{\mathbf{Y}}^B$, which can be described as

$$\tilde{\mathbf{H}} = \mathbf{HQ}^A = \begin{bmatrix} \tilde{h}_{11} & \tilde{h}_{12} & \cdots & \tilde{h}_{1N} \\ \tilde{h}_{21} & \tilde{h}_{22} & \cdots & \tilde{h}_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{h}_{N1} & \tilde{h}_{N2} & \cdots & \tilde{h}_{NN} \end{bmatrix}, \tag{10}$$

whose individual elements can be expressed as

$$\tilde{h}_{jk} = h_{j1}q_{1k}^A + h_{j1}q_{2k}^A + \ldots + h_{jN}q_{Nk}^A \tag{11}$$

$$= \sum_{l=1}^{N} h_{jl}q_{lk}^A, \quad j = 1\ldots N, k = 1\ldots N. \tag{12}$$

Building on this remark, we obtain the following lemma.

Lemma 3. Each element of the sum (12) is a zero-mean Gaussian variable.

The proof is given in Appendix C. Based on Lemma 3, we can derive the following.

Proposition 1. *A vector* $\tilde{\mathbf{h}} = \text{vec}(\tilde{\mathbf{H}})$ *forms a Gaussian vector.*

The proof of Proposition 1 can be found in Appendix D. Similarly, we can prove that a vector of $(\mathbf{Q}^B)^T\mathbf{H}\mathbf{Q}^A$ can also form a Gaussian vector.

By exploiting the results of zero-mean Gaussian vectors [10], we can express the secret key rate as

$$I_K = \log \frac{|\mathbf{K}_a||\mathbf{K}_b|}{|\mathbf{K}_{AB}|}, \tag{13}$$

where

$$\mathbf{K}_{AB} = \begin{bmatrix} \mathbf{K}_a & \mathbf{K}_{ab} \\ \mathbf{K}_{ba} & \mathbf{K}_b \end{bmatrix}, \tag{14}$$

and $\mathbf{K}_a$ and $\mathbf{K}_b$ are the covariance matrices of the corresponding channels. Note that $K_a = \mathbb{E}(\mathbf{a}, \mathbf{a})$, $K_b = \mathbb{E}(\mathbf{b}, \mathbf{b})$, $K_{ab} = \mathbb{E}(\mathbf{a}, \mathbf{b})$ where $\mathbf{a}$ and $\mathbf{b}$ are the vectors of interest. In our case, $\mathbf{a}$ and $\mathbf{b}$ are vectors of length N^2 obtained by stacking the corresponding channel matrices of size $N \times N$ at Alice and Bob, respectively. As a result, we can express the secret key rate as follows.

Theorem 1. *The secret key rate of the considered system can be expressed as*

$$I_K = N^2 \log\left(1 + \frac{N^2}{2N\sigma_{IN}^2 + \sigma_{IN}^4}\right), \tag{15}$$

where $\sigma_{IN}^2 = \sigma_W^2 + \sigma_Z^2$ *is the interference-plus-noise power.*

The detailed proof can be found in Appendix E. We note that we obtain the term of N^2 outside the log as a result of using the channel matrix $N \times N$ in the SKG [11].

Remark 2. Having equipped with multiple antennas, we can retain a gain of $g = \log\left(\frac{(\sigma_{IN}^2+N)^{2N^2}(2\sigma_{IN}^2+\sigma_{IN}^4)}{(2N\sigma_{IN}^2+\sigma_{IN}^4)^{N^2}(\sigma_{IN}^2+1)^2}\right)$ in comparison with single antenna configuration. In case of single antenna at legitimate users, i.e., $N = 1$, the expression in Theorem 1 reduces to that of Lemma 2.

Remark 3. Denote $\xi = \frac{N}{\sigma_{IN}^2}$ the signal-to-interference-plus-noise (SINR), the secret key rate can be written in an equivalent form as

$$I_K = N^2 \log\left(1 + \frac{\xi^2}{1 + 2\xi}\right). \tag{16}$$

In the following simulations, we can observe the behavior of the secret key rate following the derived expression. In the first experiment, we study the advantages of the multiple antenna technology. In particular, the number of antennas is varied from 1 to 32, whereas the interference-plus-noise power changes from -30 to 30 dBW. As expected, MIMO setting helps to enhance the secret key rate significantly as shown in Fig. 2. However, when the interference-plus-noise power increases, the useful signal becomes less dominant and thus the secret key rate will decrease drastically.

In the second example, we study the impact of the signal-to-interference-plus-noise, i.e., ξ, on the system. In particular, SINR ranges from -5 to 20 dB and we also vary the number of antennas. As can be seen from Fig. 3, for a fixed injection signal power, increasing the number of antennas results in an increase of the secret key rate. If we fix the number of antennas, we also notice the degradation of the secret key rate when either interference or noise power becomes significant and therefore results in higher interference-plus-noise power as observed in the preceding simulation. Note that both axes of Fig. 3 are logarithm base 10 and thus the secret key rate increases linearly with the number of antenna as determined in Eq. (16).

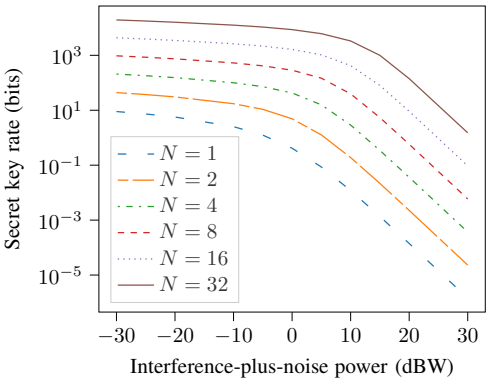

Fig. 2: Secret key rate under varying interference-plus-noise power.

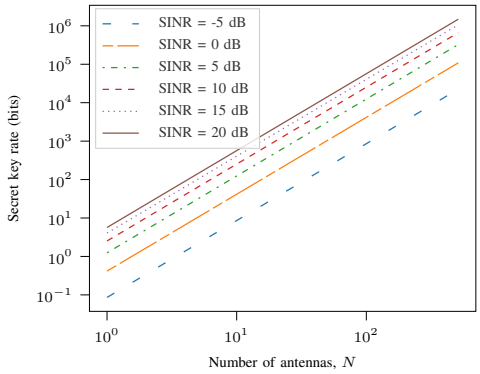

Fig. 3: Secret key rate under different SINR and number of antennas setting.

IV. Conclusions

We have modelled an injection attack for MIMO secret key generation systems and have proven that such attacks are feasible as long as the attacker has one extra antenna compared to legitimate users. Interestingly, by using a pilot randomization scheme to counteract the attack, we are able to take advantage of the special structure of the formulation to derive a closed-form expression for the secret key rate, which was missing from literature. Finally, we have evaluated numerically the impact of the injected signal on the performance of a SKG system using the pilot randomization method in different MIMO settings.

Appendix A
Proof of Lemma 1

Denote

$$\mathbf{H}^A - \mathbf{H}^B = [\bar{\mathbf{H}} \quad \hat{\mathbf{H}}], \tag{17}$$

where $\bar{\mathbf{H}} \in \mathbb{C}^{N\times N}$, and $\hat{\mathbf{H}}^M \in \mathbb{C}^{N\times(M-N)}$. Similarly, we can denote $\mathbf{F} = [\bar{\mathbf{F}}^T \quad \hat{\mathbf{F}}^T]$ where $\bar{\mathbf{F}} \in \mathbb{C}^{N\times N}$, and $\hat{\mathbf{F}} \in \mathbb{C}^{(M-N)\times N}$. Therefore, we can rewrite the condition (5) as

$$[\bar{\mathbf{H}} \quad \hat{\mathbf{H}}] \begin{bmatrix} \bar{\mathbf{F}} \\ \hat{\mathbf{F}} \end{bmatrix} = 0, \tag{18}$$

or equivalently

$$\bar{\mathbf{H}}\bar{\mathbf{F}} = -\hat{\mathbf{H}}\hat{\mathbf{F}}. \tag{19}$$

By fixing the value of $\hat{\mathbf{F}}$ we have

$$\bar{\mathbf{F}} = -(\bar{\mathbf{H}})^{-1}\hat{\mathbf{H}}\hat{\mathbf{F}}. \tag{20}$$

Due to the randomness of the channel, an invertible channel $\bar{\mathbf{H}}$ can be chosen and thus completes the proof. Note that if Alice and Bob have single antenna while Mallory has two, the equation above reduces to [5, Eq. 1].

Appendix B
Proof of Lemma 2

Considering the single-antenna case at legitimate users, we can rewrite the estimated channels at Alice and Bob as follows:

$$y^B = hq^A + w + z^B, \tag{21}$$

$$y^A = q^B h + w + z^A. \tag{22}$$

Utilizing randomization method, we can rewrite the aforementioned equations as

$$\tilde{y}^B = q^B y^B = q^B h q^A + q^B w + q^B z^B, \tag{23}$$

$$\tilde{y}^A = y^A q^A = q^B h q^A + w q^A + z^A q^A. \tag{24}$$

It is worth noting that for two complex i.i.d random variables X, Y with correlation coefficient ρ, the mutual information is given by

$$I_K = -\log(1-\rho^2), \tag{25}$$

where the correlation coefficient $\rho = \frac{\mathbb{E}(X,Y)}{\sqrt{\mathbb{E}(Y,Y)\mathbb{E}(X,X)}}$.

In our case, we can compute

$$\mathbb{E}(\tilde{y}^B, \tilde{y}^A) = 1, \tag{26}$$

$$\mathbb{E}(\tilde{y}^B, \tilde{y}^B) = 1 + \sigma_W^2 + \sigma_Z^2, \tag{27}$$

$$\mathbb{E}(\tilde{y}^A, \tilde{y}^A) = \mathbb{E}(\tilde{y}^B, \tilde{y}^B). \tag{28}$$

Combining these results with (25), we obtain

$$I_K = -\log(1 - \frac{1}{(1+\sigma_W^2+\sigma_Z^2)^2}), \tag{29}$$

$$= \log\left(\frac{(1+\sigma_W^2+\sigma_Z^2)^2}{2(\sigma_W^2+\sigma_Z^2)+(\sigma_W^2+\sigma_Z^2)^2}\right), \tag{30}$$

$$= \log\left(1 + \frac{1}{2(\sigma_W^2+\sigma_Z^2)+(\sigma_W^2+\sigma_Z^2)^2}\right). \tag{31}$$

Define $\sigma_{IN}^2 = \sigma_W^2 + \sigma_Z^2$, the secret key rate for the scalar case can be written as

$$I_K = \log\left(1 + \frac{1}{2\sigma_{IN}^2 + \sigma_{IN}^4}\right), \tag{32}$$

which completes the proof.

Appendix C
Proof of Lemma 3

Note that we use conventional notation that h is a realization and H is a random variable. In the following, we will prove that each component in the sum of (12) can be derived from zero-mean independent Gaussian variables. Since H is i.i.d. and Q^A has Rademacher distribution for both real and complex parts, we obtain the product $H_{jl}Q_{lk}^A \sim \mathcal{CN}(0,1)$ which is a direct result of [5].

We can further study the cross correlation between components in the sum (12). Note that for fixed j and k, we can obtain arbitrary pair of random variables $H_l = H_{jl}Q_{lk}^A$ and $H_v = H_{jv}Q_{vk}^A$, $l \neq v$, which results in

$$K_{h_l,h_v} = \mathbb{E}(H_l, H_v) \tag{33}$$

$$= \mathbb{E}(H_l H_v) - \mathbb{E}(H_l)\mathbb{E}(H_v)$$
$$= \mathbb{E}(H_{jl}Q_{lk}^A H_{jv}Q_{vk}^A) - \mathbb{E}(H_{jl}Q_{lk}^A)\mathbb{E}(H_{jv}Q_{vk}^A)$$
$$= \mathbb{E}(H_{jl}H_{jv})E(Q_{lk}^A Q_{vk}^A) - \mathbb{E}(H_{jl})\mathbb{E}(Q_{lk}^A)\mathbb{E}(H_{jv})\mathbb{E}(Q_{vk}^A)$$
$$= 0. \tag{34}$$

We now prove that a vector of aforementioned variables $\mathbf{X} = [H_1, H_2, \cdots, H_N]^T$ can follow the multivariate normal distribution. First, consider the covariance

$$\mathbf{K}_x = \begin{bmatrix} \mathbb{E}(H_1^2) & \mathbb{E}(H_1,H_2) & \cdots & \mathbb{E}(H_1,H_N) \\ \mathbb{E}(H_2,H_1) & \mathbb{E}(H_2^2) & \cdots & \mathbb{E}(H_2,H_N) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{E}(H_N,H_1) & \mathbb{E}(H_N,H_2) & \cdots & \mathbb{E}(H_N^2) \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = \mathbf{I}. \tag{35}$$

Note that $\mathbb{E}(\mathbf{X}) = \mathbf{0}$, thus a multivariate Gaussian probability density function can be written as

$$f_{\mathbf{X}}(x) = \frac{1}{(\pi)^{\frac{d}{2}}|\mathbf{K}_x|^{\frac{1}{2}}} x^{-(x-\mathbb{E}(X))^H K_x^{-1}(x-\mathbb{E}(X))}$$
$$= \frac{1}{(\pi)^{\frac{N}{2}}} e^{-(|h|_1^2+|h|_2^2+\cdots+|h|_N^2)}$$
$$= \frac{1}{\sqrt{\pi}} e^{-|h|_1^2} \frac{1}{\sqrt{\pi}} e^{-|h|_2^2} \cdots \frac{1}{\sqrt{\pi}} e^{-|h|_N^2}$$
$$= f_X(h_1) f_X(h_2) \cdots f_X(h_N). \tag{36}$$

As a consequence, we can always assume that components of the considered sum originate from independent Gaussian variables.

Appendix D
Proof of Proposition 1

In this appendix, we investigate the independence properties of the matrix in (10). In particular, considering a pair, e.g., $\tilde{H}_{jk}$ and $\tilde{H}_{ju}, k \neq u$, we can obtain the covariance matrix as

$$K_{\tilde{h}_{jk},\tilde{h}_{ju}} = \mathbb{E}(\tilde{H}_{jk}, \tilde{H}_{ju}) \tag{37}$$

$$= \mathbb{E}(\tilde{H}_{jk}\tilde{H}_{ju}) - \mathbb{E}(\tilde{H}_{jk})\mathbb{E}(\tilde{H}_{ju})$$
$$= \mathbb{E}\left(\left(\sum_{l=1}^{N} H_{jl}Q_{lk}^A\right)\left(\sum_{u=1}^{N} H_{ju}Q_{uk}^A\right)\right)$$
$$= \mathbb{E}\left(\sum_{l=1}^{N}\left(\sum_{u=1}^{N} H_{jl}Q_{lk}^A H_{ju}Q_{uk}^A\right)\right)$$
$$= \sum_{l=1}^{N}\left(\sum_{u=1}^{N} \mathbb{E}(H_{jl}Q_{lk}^A H_{ju}Q_{uk}^A)\right)$$
$$= 0. \tag{38}$$

We can continue in this fashion obtaining the same results for arbitrary pair of elements in the matrix of (10). We can then apply the same arguments in the proof of Lemma 3 (c.f. Eqs. (35)-(36)) and thus conclude the proof.

APPENDIX E
PROOF OF THEOREM 1

To derive the secret key rate, we can first vectorize the obtained channel matrices as follows $\tilde{\mathbf{y}}^A = \text{vec}(\tilde{\mathbf{Y}}^A), \tilde{\mathbf{y}}^B = \text{vec}(\tilde{\mathbf{Y}}^B)$. By definition of the covariance matrix, we thus get

$$\begin{aligned}\mathbf{K}_a &= \mathbb{E}(\tilde{\mathbf{y}}^A(\tilde{\mathbf{y}}^A)^H) \quad (39)\\ &= \mathbb{E}\Big(\text{vec}((\mathbf{Q}^B)^T\mathbf{H}\mathbf{Q}^A)\text{vec}^H((\mathbf{Q}^B)^T\mathbf{H}\mathbf{Q}^A)\Big)\\ &+ \mathbb{E}\Big(\text{vec}((\mathbf{Q}^B)^T\mathbf{H}\mathbf{Q}^A)\text{vec}^H((\mathbf{W}\mathbf{Q}^A))\Big)\\ &+ \mathbb{E}\Big(\text{vec}((\mathbf{Q}^B)^T\mathbf{H}\mathbf{Q}^A)\text{vec}^H(\mathbf{Z}^A\mathbf{Q}^A)\Big)\\ &+ \mathbb{E}\Big(\text{vec}(\mathbf{W}\mathbf{Q}^A)\text{vec}^H((\mathbf{Q}^B)^T\mathbf{H}\mathbf{Q}^A)\Big)\\ &+ \mathbb{E}\Big(\text{vec}(\mathbf{W}\mathbf{Q}^A)\text{vec}^H((\mathbf{W}\mathbf{Q}^A))\Big)\\ &+ \mathbb{E}\Big(\text{vec}(\mathbf{W}\mathbf{Q}^A)\text{vec}^H((\mathbf{Z}^A\mathbf{Q}^A)\Big)\\ &+ \mathbb{E}\Big(\text{vec}(\mathbf{Z}^A\mathbf{Q}^A)\text{vec}^H((\mathbf{Q}^B)^T\mathbf{H}\mathbf{Q}^A)\Big)\\ &+ \mathbb{E}\Big(\text{vec}(\mathbf{Z}^A\mathbf{Q}^A)\text{vec}^H((\mathbf{W}\mathbf{Q}^A)\Big)\\ &+ \mathbb{E}\Big(\text{vec}(\mathbf{Z}^A\mathbf{Q}^A)\text{vec}^H((\mathbf{Z}^A\mathbf{Q}^A))\Big). \quad (40)\end{aligned}$$

Since the channel, randomization matrices and noise are independent, after some manipulations, we arrive at

$$\begin{aligned}\mathbf{K}_a &= \mathbb{E}\Big(((\mathbf{Q}^A)^T\otimes(\mathbf{Q}^B)^T)((\mathbf{Q}^A)^T\otimes(\mathbf{Q}^B)^T)^H\Big)\\ &+ (\sigma_W^2+\sigma_Z^2)\mathbb{E}\Big(((\mathbf{Q}^A)^T\otimes\mathbf{I})((\mathbf{Q}^A)^T\otimes\mathbf{I})^H\Big), \quad (41)\end{aligned}$$

where $\sigma_{IN}^2 = \sigma_W^2 + \sigma_Z^2$. Due to the properties of mixed products, i.e., $(\mathbf{A}\otimes\mathbf{B})(\mathbf{C}\otimes\mathbf{D}) = (\mathbf{AC})\otimes(\mathbf{BD})$, we can easily achieve

$$\begin{aligned}\mathbf{K}_a &= \mathbb{E}\Big((\mathbf{Q}^A(\mathbf{Q}^A)^H\otimes\mathbf{Q}^B(\mathbf{Q}^B)^H)^T\Big)\\ &+ \sigma_{IN}^2\mathbb{E}\Big((((\mathbf{Q}^A)^H\mathbf{Q}^A)^T\otimes\mathbf{I})\Big). \quad (42)\end{aligned}$$

Repeating derivation to that of Bob enables us to write

$$\begin{aligned}\mathbf{K}_b &= \mathbb{E}\Big((\mathbf{Q}^A(\mathbf{Q}^A)^H\otimes\mathbf{Q}^B(\mathbf{Q}^B)^H)^T\Big)\\ &+ \sigma_{IN}^2\mathbb{E}\Big((\mathbf{I}\otimes((\mathbf{Q}^B)^H\mathbf{Q}^B)^T)\Big), \quad (43)\end{aligned}$$

$$\mathbf{K}_{ab} = \mathbb{E}\Big((\mathbf{Q}^A(\mathbf{Q}^A)^H\otimes\mathbf{Q}^B(\mathbf{Q}^B)^H)^T\Big). \quad (44)$$

The same is applicable to $\mathbf{K}_{ba}$ and it is easy to check that $\mathbf{K}_a = \mathbf{K}_b, \mathbf{K}_{ab} = \mathbf{K}_{ba}$. Applying the property of determinants and the derivations above to (13) yield

$$\begin{aligned}I_K &= \log\frac{|\mathbf{K}_a|}{|\mathbf{K}_a - \mathbf{K}_{ab}\mathbf{K}_a^{-1}\mathbf{K}_{ab}|} \quad (45)\\ &= \text{logdet}(\mathbf{K}_a) - \text{logdet}(\mathbf{K}_a - \mathbf{K}_{ab}\mathbf{K}_a^{-1}\mathbf{K}_{ab}) \quad (46)\\ &= I_1 - I_2. \quad (47)\end{aligned}$$

An element of the product $\tilde{\mathbf{Q}} = \mathbf{Q}^A(\mathbf{Q}^A)^H$ is given by

$$\tilde{q}_{jk} = \sum_{l=1}^{N} q_{jl}^A(q_{lk}^A)^*, \quad j = 1\dots N, k = 1\dots N. \quad (48)$$

On the diagonal, i.e., $j = k$, we obtain $\tilde{q}_{jj} = \sum_{l=1}^{N}|q_{jl}^A|^2$. Thus

$$\mathbb{E}(\tilde{q}_{jj}) = N. \quad (49)$$

We can apply similar arguments to the product $(\mathbf{Q}^B)^H\mathbf{Q}^B$ and therefore obtain

$$I_1 = \sum_{l=1}^{N^2}\log((N+\sigma_{IN}^2)N) = N^2\log((N+\sigma_{IN}^2)N). \quad (50)$$

We note that another indirect result of this derivation is the computation of SINR which is given by

$$\begin{aligned}\xi &= \frac{\text{tr}\Big(\mathbb{E}(\tilde{\mathbf{y}}_s\tilde{\mathbf{y}}_s^H)\Big)}{\text{tr}\Big(\mathbb{E}(\tilde{\mathbf{y}}_w\tilde{\mathbf{y}}_w^H)\Big) + \text{tr}\Big(\mathbb{E}(\tilde{\mathbf{y}}_n\tilde{\mathbf{y}}_n^H)\Big)} \quad (51)\\ &= \frac{\text{tr}\Big(\mathbb{E}\Big((\mathbf{Q}^A(\mathbf{Q}^A)^H\otimes\mathbf{Q}^B(\mathbf{Q}^B)^H)^T\Big)\Big)}{\sigma_{IN}^2\,\text{tr}\Big(\mathbb{E}\Big((((\mathbf{Q}^A)^H\mathbf{Q}^A)^T\otimes\mathbf{I})\Big)\Big)} = \frac{N}{\sigma_{IN}^2} \quad (52)\end{aligned}$$

where $\tilde{\mathbf{y}}_s = \text{vec}((\mathbf{Q}^B)^T\mathbf{H}\mathbf{Q}^A), \tilde{\mathbf{y}}_w = \text{vec}(\mathbf{W}\mathbf{Q}^A)$, and $\tilde{\mathbf{y}}_z = \text{vec}(\mathbf{Z}^A\mathbf{Q}^A)$. The second term of (47) can be handled in much the same way which results in

$$\begin{aligned}I_2 &= \sum_{l=1}^{N^2}\log\left((N+\sigma_{IN}^2)N - \frac{N^3}{(N+\sigma_{IN}^2)}\right), \quad (53)\\ &= N^2\log((2N\sigma_{IN}^2+\sigma_{IN}^4)N) - N^2\log(N+\sigma_{IN}^2).\end{aligned}$$

Combining these results, we can rewrite (47) as

$$\begin{aligned}I_K = I_1 - I_2 &= N^2\log\left(\frac{(N+\sigma_{IN}^2)^2}{2N\sigma_{IN}^2+\sigma_{IN}^4}\right), \quad (54)\\ &= N^2\log\left(1+\frac{N^2}{2N\sigma_{IN}^2+\sigma_{IN}^4}\right), \quad (55)\end{aligned}$$

which proves the theorem.

REFERENCES

[1] Y. Zou, J. Zhu, X. Wang, and L. Hanzo, "A survey on wireless security: Technical challenges, recent advances, and future trends," *Proc. IEEE*, vol. 104, no. 9, pp. 1727–1765, 2016.

[2] A. Chorti *et al.*, "Context-aware security for 6G wireless: The role of physical layer security," *IEEE Commun. Stand. Mag.*, vol. 6, no. 1, pp. 102–108, 2022.

[3] R. Jin and K. Zeng, "Physical layer key agreement under signal injection attacks," in *Proc. IEEE CNS*, 2015, pp. 254–262.

[4] A. Chorti, "A study of injection and jamming attacks in wireless secret sharing systems," *Int. Workshop on Comm. Secur., Springer Cham*, 2017, pp. 1–14.

[5] M. Mitev, A. Chorti, E. V. Belmega, and M. Reed, "Man-in-the-middle and denial of service attacks in wireless secret key generation," in *Proc. IEEE GLOBECOM*, 2019, pp. 1–6.

[6] S. Eberz, M. Strohmeier, M. Wilhelm, and I. Martinovic, "A practical man-in-the-middle attack on signal-based key generation protocols," *Proc. 17th Europ. Symp. Research Comput. Secur. (ESORICS)*, 2012, pp. 235–252.

[7] M. Mitev, A. Chorti, E. V. Belmega, and H. V. Poor, "Protecting physical layer secret key generation from active attacks," *Entropy*, vol. 23, no. 8, 2021.

[8] E. A. Jorswieck, A. Wolf, and S. Engelmann, "Secret key generation from reciprocal spatially correlated MIMO channels," in *Proc. IEEE GLOBECOM Workshops*, 2013, pp. 1245–1250.

[9] P. Hitczenko and S. Kwapień, "On the Rademacher series," in *Probability in Banach Spaces, 9*, J. Hoffmann-Jørgensen, J. Kuelbs, and M. B. Marcus, Eds. Boston, MA: Birkhäuser Boston, 1994, pp. 31–36.

[10] J. W. Wallace and R. K. Sharma, "Automatic secret keys from reciprocal MIMO wireless channels: Measurement and analysis," *IEEE Trans. Inf. Forensics Secur.*, vol. 5, no. 3, pp. 381–392, 2010.

[11] G. Pasolini and D. Dardari, "Secret information of wireless multi-dimensional Gaussian channels," *IEEE Trans. Wirel. Commun.* , vol. 14, no. 6, pp. 3429–3442, 2015.