Leakage Rate Analysis for Artificial Noise Assisted Massive MIMO With Non-Coherent Passive Eavesdropper in Block-Fading

Changick Song, Senior Member, IEEE

Abstract—Massive MIMO is one of the salient techniques for achieving high spectral efficiency in the next-generation wireless networks. Recently, a combined strategy of the massive MIMO and the artificial noise (AN), namely, AN assisted massive (ANAM) MIMO has been investigated for security enhancement. However, most of the recent studies have been built upon the assumption of perfect channel state information (CSI) at the eavesdropper (ED), the results of which may be too pessimistic in terms of security because there exists no ED-friendly downlink training from the base station in the practical ANAM systems. In this paper, we provide more sophisticated analysis on the secrecy performance of the ANAM systems assuming that the CSI of the ED channels are unknown to both the base station and the ED or partially known to the ED. We measure the secrecy in terms of the leakage rate to the ED (or the secrecy rate to the legitimate users) and characterize their upper and lower bounds in the high signal-to-noise ratio regime. Our analytical result is represented by a single compact expression as a function of the number of ED antennas, the dimensionality of signal space, and the channel coherence time, and thus offers useful insights that help us fully exploit the secrecy potential of the ANAM systems. Finally, several useful observations are made from the numerical examples.

Index Terms—Massive MIMO, artificial noise, non-coherent, wiretap channels, leakage rate, secrecy rate.

I. INTRODUCTION

Along with the growing prevalence of wireless radio technologies, the security has become a major social challenge for both personal and professional sphere. Unlike the wired communications, however, the wireless security is in general a challenging task, since the radio transmission has no physical boundary, and thus any receivers nearby can listen to the transmitted signals. As the wireless link is also unreliable and severely constrained by energy and bandwidth, more sophisticated physical layer designs are required. An important objective of the physical layer wireless security is to protect such radio interface by negligibly low probability of interception (LPI) without relying on (but can be integrated with) the upper-layer crypto system.

Several classical solutions have already existed to achieve the LPI. For example, a waveform can be designed at the base-station (BS) using furtive frequency/time hopped or spread spectrum signals [1]. However, it may not be secure enough in wideband systems due to reduced space for the spreading gain. If the channel state information (CSI) of the eavesdropper (ED) is allowed at the BS, we may apply the directional beamforming scheme so that the data signals do not appear at the ED. However, achieving this is practically difficult because the EDs are typically passive so as to hide their existence. To address such issues, there are several of LPI designs that do not rely on the ED’s CSI at the BS. One popular way is to broadcast artificial noise (AN) signals isotropically within the null space of the legitimate user (LU) channels [2]–[6]. It is also possible to utilize a large excess of BS antennas as in [7], in which the channel hardening effect automatically arises rendering stable and predictable channel condition to the LUs, while nearly nullified channel gains to the ED. In addition, a combined strategy of the massive MIMO and the AN, namely, AN assisted massive MIMO (ANAM) has recently been actively investigated for security enhancement [8], [9].

For CSI acquisition in the massive MIMO systems, the time division duplexing (TDD) scheme is in general a better choice than the frequency division duplexing scheme because the channel reciprocity enables the BS to obtain the downlink CSI by utilizing the uplink channel training from the LUs without resorting to the downlink training that may incur prohibitively high overhead [10]. Our work is motivated by the fact that such no downlink training gives rise to a security advantage due to channel uncertainty at the ED. More specifically, unlike the LUs that can enjoy the deterministic channel condition via beamforming from the BS, the best strategy that can be taken by the ED without the downlink training is the blind estimation (or detection) schemes [11], [12] whose performance is heavily dependent on the coherence time and the dimensionality of the signal space. Thus, it is expected that as the coherence time of ED channels becomes smaller or the number of BS antennas increases, decoding ability of the ED will deteriorate. Nevertheless, most of previous research on the
ANAM systems has been built upon the full CSI assumption at the ED.

In this paper, we provide more sophisticated investigation on the performance of the TDD-based ANAM systems with non-coherent passive EDs in block fading. We make a practical assumption that the instantaneous CSI of the ED channels are unknown to anybody including the BS and the ED. Then, the goal of the paper is to quantify the amount of leakage information to the non-coherent ED in a simple mathematical expression. Note that once the leakage rate is identified, the secrecy rate can be easily computed, which represents the maximum secure data rate to the LUs [13], [14].

Similar security concepts based on no downlink training have been developed in [15]–[18], but the analysis is still based on the full CSI assumption at the ED. Our analysis also differs from the previous works for non-coherent capacity in [19]–[21] because the AN signal transmission over random propagation makes the analysis more challenging. The contribution of the paper can be summarized as follows.

- First, in Section IV, we analyze the leakage rate to the non-coherent ED, namely, ‘non-coherent leakage rate’ based on the high signal-to-noise ratio (SNR) approximation. Our analysis not only unveils the degrees of freedom (DoF) of the leakage rate, but also identifies upper and lower bounds of the constant terms that are irrespective of the ED’s SNR. The bounds become tight when the coherence time exceeds the number of BS antennas. One interesting observation from our analysis is that we can nullify the leakage rate DoF to the non-coherent ED by scaling up the dimensionality of the signal space, no matter how many antennas the ED has and even without the aid of wiretap codes and ED’s CSI at the BS. Obviously, this is not the case for the coherent ED because, in this case, a large number of receive antennas enable the ED to eliminate the AN via null-space projection scheme, which typically leads to a non-zero DoF in the leakage rate [3].

- Next, in Section V, we examine ‘partially coherent leakage rate’ considering a situation where the BS uses a downlink training over a few of beamforming vectors like the beamformed CSI-RS in the long term evolution advanced (LTE-A) [22]. The situation may occur when the CSI at the BS is imperfect or outdated, and the LUs need to estimate their effective channels more accurately. In this case, we should accept that the EDs also listen to the training signals and can perform the coherent detection using them. Nevertheless, the channel information that is associated with the AN is still unknown, which means that some security gain remains. In this section, the leakage rate is investigated in two folds, the universal upperbound that holds for all SNRs, and tighter upper and lower bounds that hold for high SNR. It turns out that the universal bound is tight in the low SNR regime.

- In Section VI, we provide in-depth discussion on the designs of the ANAM systems based on our analytical results. We examine the required number of BS antennas for achieving the zero-DoF leakage rate according to the coherence time of the ED channels. Also, we suggest a new artificial fast fading (AFF) design that deliberately shortens the ED’s channel coherence time. We also investigate achievable secrecy-rate to the LUs considering the wiretap codes at the BS. Finally, in VII, numerical results confirm the accuracy of our analysis and demonstrate the security potential of the ANAM systems.

**Notations:** In this paper, unless specified otherwise, boldface lowercase and uppercase letters indicate random vectors and matrices, respectively. The superscripts $(\cdot)^T$, $(\cdot)^*$, and $(\cdot)^H$ stand for the transpose, conjugate, and conjugate transpose operations, and $\mathbb{C}$ and $\mathbb{R}_+$ denote sets of complex and real positive numbers, respectively. We denote by $1_N$ an $N \times N$ identity matrix and by $E[\cdot]$ the expectation. We write $\text{Tr}(A)$ and $\text{det}(A)$ for the trace and determinant of a matrix $A$, respectively. We use $(\cdot)^+$ and $\mathcal{I}(X,Y)$ to denote $\max\{0,0\}$ and the mutual information between two random matrices $X$ and $Y$, respectively. We write $h(\cdot)$ as differential entropy to the base 2. The equivalence $d \overset{d}{=} d'$ means the same distribution.

**II. SYSTEM MODEL**

In this paper, we consider a multiuser downlink system where a BS having a large $M$ number of antennas supports $K$ LUs with a single antenna in the presence of an ED with $N_E$ antennas. Note that the multi-antenna ED can be interpreted as multiple cooperative EDs with a total of $N_E$ antennas. Similarly, the proposed analytical results are generally applicable to the case of a single LU with $K$ antennas. We adopt a Rayleigh flat fading model in which the baseband channels from the BS to the $k$-th LU and ED are respectively expressed by $h_k^H \in \mathbb{C}^{1 \times M}$ and $G \in \mathbb{C}^{N_E \times M}$ whose propagation coefficients are independent and identically distributed (i.i.d.) complex Gaussian with zero mean and unit variance, i.e., $CN(0,1)$. We consider the block-fading where all channel coefficients remain constant for $T$ symbol periods, and change to new independent values in the next time period.

The BS operates in a TDD mode and has perfect knowledge of $H = [h_1, \ldots, h_K]^H$ utilizing the uplink training from the LUs and the channel reciprocity. Meanwhile, the instantaneous CSIs of the ED channel $G$ is completely unknown to the BS due to the passive eavesdropping of the ED. The BS basically transmits no downlink training in order to prevent the ED from coherent detection, which means that the ED’s channel information is unavailable to anybody including the BS and the ED. As in the conventional massive MIMO, the system is designed such that $M$ is much greater than $K$. In contrast, $N_E$ could be arbitrarily large to be compared with $M$, given the situation where several multi-antenna EDs can cooperate. For simplicity, we consider homogeneous users, i.e., each LU experiences the same received signal power on average and ignore the uplink phase duration in the coherence time $T$.

As we have assumed that $M \gg K$, the extra antenna dimension at the BS can be exploited for AN transmission to

\[\text{\textsuperscript{1}}\text{For mathematical convenience, we assume that at least the statistical information of the ED channel is known to the BS. This assumption is practically valid, since the BS can estimate the geometrical or stochastic channel information of a passive ED via detecting inevitable power leakage of the ED’s local oscillator [23] or utilizing the torch aided methods in [24].\]
interfere with the ED. Specifically, during each coherence interval, the BS generates and transmits an $M$ dimensional signal matrix $X \in \mathbb{C}^{M \times T}$ as

$$X = \alpha H^H S + \beta V_{H_2, N_J} N,$$

(1)

where $S \triangleq [s_1, \ldots, s_K]^H \in \mathbb{C}^{K \times T}$ denotes the $K$ dimensional message signals whose $k$-th row $s_k^H \in \mathbb{C}^{1 \times T}$ carries a message to the $k$-th LU, $H^H \triangleq \sqrt{M} H^H (H H^H)^{-1}$ refers to the pseudo-inverse of $H$, and $N \in \mathbb{C}^{N_J \times T}$ contains the $N_J$ dimensional artificial jamming noise with $N_J \leq M - K$. Here, $V_{H_2, N_J} \in \mathbb{C}^{M \times N_J}$ represents an orthonormal matrix with $V_{H_2, N_J}^H V_{H_2, N_J} = I_{N_J}$, and $HV_{H_2, N_J} = 0$, and is attainable from the following singular value decomposition (SVD)

$$H = U_H \Lambda H \big| V_H^H,$$

(2)

where $U_H \in \mathbb{C}^{K \times K}$ and $V_H \triangleq [V_{H_1}, V_{H_2}] \in \mathbb{C}^{M \times M}$ denote unitary matrices with $V_{H_1} \in \mathbb{C}^{M \times K}$ and $V_{H_2} \in \mathbb{C}^{M \times (M-K)}$, and $\Lambda H \in \mathbb{C}^{K \times K}$ refers to a diagonal matrix having non-zero singular values. Then, $V_{H_2, N_J}$ in (1) is attained by the first $N_J$ columns of $V_{H_2}$. In each symbol time, the BS can transmit at most $M - K$ dimensional AN signals isotropically in the null subspace of $H$.

It is assumed that $S$ and $N$ have i.i.d. Gaussian entries $s_{ij} \sim \mathcal{CN}(0,1)$ and $n_{ij} \sim \mathcal{CN}(0,1)$, respectively to maximize the information rate to the LUs while minimizing the leakage rate to the ED for a given $S$. The scaling factors $\alpha \in \mathbb{R}_+$ and $\beta \in \mathbb{R}_+$ distribute the power to the message and AN signals, subject to a power constraint $\text{Tr}(E[XX^H]) = MT$ where

$$\text{Tr}(E[XX^H]) = TM \alpha^2 \text{Tr}(H(HH^H)^{-1}) + \beta^2 N_J T,$$

$$= \frac{TMK}{M-K} \alpha^2 + \beta^2 N_J T.$$

(3)

The last equality follows from $E[(HH^H)^{-1}] = \frac{1}{M-K} I_K$ [25]. The values of $\alpha$ and $\beta$ are fixed over a codeword duration, and assumed to be known to all nodes.

After the signal in (1) is transmitted, corresponding received signals at the $k$-th LU and the ED can be respectively expressed by

$$y_k^H = \alpha \sqrt{M} s_k^H + w_k^H,$$

(4)

and

$$Y_E = G_1 S + G_2 N + Z,$$

(5)

where $w_k^H \in \mathbb{C}^{1 \times T}$ and $Z \in \mathbb{C}^{N_E \times T}$ denote the Gaussian thermal noise having i.i.d. entries $w_{kj} \sim \mathcal{CN}(0, \sigma_z^2)$ and $z_{kj} \sim \mathcal{CN}(0, \sigma_z^2)$, respectively. Also, we define $G_1 \triangleq \alpha GH^H \in \mathbb{C}^{N_E \times K}$ and $G_2 \triangleq \beta GV_{H_2, N_J} \in \mathbb{C}^{N_E \times N_J}$ as the effective channels of the ED. We refer to SNR$_L = \frac{\min(M, N_E)}{\sigma_z^2}$ and SNR$_E = \frac{\min(N_E, N_J)}{\sigma_z^2}$ as the expected SNRs at each receive antenna of the LU and the ED, respectively.

In our system, the channel coding can be applied over multiple $n$ fading blocks. We do not consider any wiretap codes until we discuss the secrecy rate in Section VI-C. The encoding function is thus a one-to-one mapping between the messages and the codewords. Let us define $\mathcal{M}_k = \{1, \ldots, 2^{nT_Hk}\}$ as a message set for the $k$-th LU with a transmission rate $R_k$. Then, the secrecy can be measured in terms of mutual information

$$I(M_1, \ldots, M_K; Y_E^{(n)} | Y_E^{(n-1)}) = I(S^{(n)}; Y_E^{(n)})$$

which measures the amount of information leaked to the ED.

Because of the independence among different coherence intervals, it is sufficient to study one coherence interval. Therefore, equivalently we can measure the secrecy via leakage rate that is defined by

$$\mathcal{L}_{\text{noncoh}}(K, N_E, N_J, T) \triangleq \frac{1}{T} I(Y_E; S).$$

(6)

In this paper, we call (6) non-coherent leakage rate, because neither the BS nor the ED knows the CSIs of both $G_1$ and $G_2$.

The exact values of the leakage information can be formulated as

$$I(Y_E; S) = h(Y_E) - h(Y_E | S) = E \left[ \log \frac{P(Y_E | S)}{P(Y_E)} \right]$$

$$= \int dSP(S) \int dY_E P(Y_E | S) \times \log \frac{P(Y_E | S)}{dSP(S) P(Y_E | S)}$$

$$= \int dSP(S) \int dY_E \int dN \hat{P}(N) P(Y_E | S, N) \times \log \left( \frac{\int dN \hat{P}(N) P(Y_E | S, N)}{\int dS P(S) \int dN \hat{P}(N) P(Y_E | S, N)} \right).$$

(7)

A direct evaluation of (7) is prohibitive even if the numerical methods are utilized, because the exact expression of $P(Y_E | S, N)$ is unknown and a myriad of integrals over multi-dimensional complex space are computationally intractable. The goal of the paper is to find an analytic expression of (7), thereby helping readers easily evaluate the leakage rate for many cases of interest in the ANAM systems.

III. PRELIMINARIES

In this section, we introduce some useful results and approximations that will be used for our mathematical derivations.

A. Ergodic Leakage Rate

Supposing that the effective ED channels $G_1$ and $G_2$ are perfectly known at the ED (but not at the BS), the ergodic secrecy capacity was investigated in [26]. In the following lemma, we reinterpret some of the results in terms of leakage rate considering the i.i.d. Gaussian input signals.

**Lemma 1 (Ergodic Leakage Rate):** For the given knowledge on $G \triangleq [G_1 G_2] \in \mathbb{C}^{N_E \times M}$ with $M \triangleq N_J + K$, the ergodic leakage rate is approximated at high SNR$_E$ as

$$\mathcal{L}_{\text{ergodic}}(K, N_E, N_J) \triangleq \frac{1}{T} I(Y_E; S | G)$$

$$= \sigma_z^2 \min((N_E - N_J)^+, K) \log \text{SNR}_E$$

$$+ c_K, N_E, N_J, \text{ergodic} + o(1)$$

where $o(1)$ represents the vanishing terms as $\sigma_z^2 \to 0$ and $c_K, N_E, N_J, \text{ergodic}$ denote a constant value irrespective of $\sigma_z^2$, which is defined by

$$c_K, N_E, N_J, \text{ergodic} = E \left[ \sum_{i=1}^{\min(M, N_E)} \log \lambda_{G_1,i}^2 - \sum_{i=1}^{\min(N_E, N_J)} \log \lambda_{G_2,i}^2 \right]$$
with $\lambda_{G,i}$ for $i = 1, \ldots, \min(M, N_L)$ and $\lambda_{G,j}$ for $j = 1, \ldots, \min(N_E, N_J)$ being the non-zero singular values of $G$ and $G_2$, respectively.

**Proof:** For given $G$, $Y_E$ is Gaussian. Thus, we can express the mutual information as

$$T(Y_E; S|G) = TE \left[ \log \det \left( GG^H + \sigma_z^2 I_{N_E} \right) \right] - \log \det \left( G_2 G_2^H + \sigma_z^2 I_{N_E} \right).$$

By applying the SVD, the above equation can be reformulated as

$$TE \left[ \sum_{i=1}^{N_E} \log(\lambda_{G,i}^2 + \sigma_z^2) - \sum_{i=1}^{N_E} \log(\lambda_{G,i}^2 + \sigma_z^2) \right]$$

$$= -T \min((N_E - N_J)^+, K) \log \sigma_z^2 + TE \left[ \sum_{i=1}^{\min(M, N_E)} \log(\lambda_{G,i}^2 + \sigma_z^2) \right] - \sum_{i=1}^{\min(N_E, N_J)} \log(\lambda_{G,i}^2 + \sigma_z^2).$$

Finally, we obtain the lemma for $\sigma_z^2 \to 0$.

We first find from Lemma 1 that the ergodic leakage rate has a DoF $\min((N_E - N_J)^+, K)$, which means that the amount of information leaked to the coherent ED increases by $\min((N_E - N_J)^+, K)$ bps/Hz for each 3dB SNR increase. If $N_E \leq N_J$, the DoF converges to zero, since all spatial dimension of the ED is corrupted by the AN. In this case, thus, the leakage rate will be saturated at high SNR. In contrast, when the ED has a sufficiently large number of antennas such that $N_E > M$, the DoF increases up to $K$. Therefore, in a system with coherent ED, a number of ED antennas may be a security threat despite the assistance of the AN, which we call a large antenna-array attack.

**B. Probability Distributions and Gaussian Approximation**

The goal of this section is to study the distribution of the effective ED channels $G_1$ and $G_2$ in (5). We summarize the results in the following three properties.

**Property 1:** The right singular matrix $V_H$ of the LU channel $H$ in (2) is statistically independent of $A_H$ and $U_H$.

**Proof:** Recall that $H$ is isotropically distributed (i.d.), since we have $H \overset{d}{=} HQ$ for any i.d. square unitary matrix $Q$ that is independent of $H$. It is also true that $V_H^H Q \overset{d}{=} V_H^H$ [20]. The two equalities in distribution imply that the rotational variation of $V_H^H$ will not change the distribution of $A_H$ and $U_H$, which means that $A_H$ and $U_H$ are statistically independent of $V_H$. Interested readers may refer to Appendix A for more rigorous proof.

**Property 2:** The effective ED channel $G_2 = \beta GV_{H_2, N_J}$ has i.i.d. $CN(0, \beta^2)$ entries being independent of $G_1 = \alpha GV_{H_1}^H$.

**Proof:** The original ED channel matrix $G$ is i.d., which means that $GV_{H_1}$ also contains i.i.d. Gaussian entries. Thus, $G_2$ consists of i.i.d. $CN(0, \beta^2)$ elements being independent of $GV_{H_1}$. By invoking (2), $G_1 = \alpha GV_{H_1} A_H^{-1} U_H^H$. As it is true that $G_2$ is independent of $GV_{H_1}$, the remaining problem is to verify the independence of $G_2$ and $A_H^{-1} U_H^H$. Obviously, $G$ is independent of $A_H^{-1} U_H^H$ and so is $V_{H_2, N_J}$ due to Property 1. Thus, the proof is concluded.

**Property 3:** The entries of $G_1$ can be approximated to i.i.d. $CN(0, \alpha^2)$. As $\frac{M}{K} \to \infty$.

**Proof:** By the law of large numbers, we have $\lim_{\frac{M}{K} \to \infty} \frac{1}{M} HH^H = I_K$, which implies that $\frac{1}{\sqrt{M}} A_H \overset{d}{\to} I_K$ and $G_1 \overset{d}{\to} \alpha GV_{H_1} U_H^H \overset{d}{=} \alpha GV_{H_1}$. Thus, we obtain Property 3.

Figure 1 demonstrates our statement in Property 3. We see that as $\frac{M}{K}$ increases, each element of $G_1$ gets close to the corresponding Gaussian distribution. It is observed that only a small value of $\frac{M}{K}$ e.g., 4 or more, is sufficient for the approximation.

**C. Mathematical Definitions**

In what follows, we briefly introduce some mathematical definitions that will be used throughout the paper.

**Definition 1 (Stiefel Manifold [27]):** The Stiefel manifold $S(T, M)$ for $T \geq M$ is defined as the set of all unitary $M \times T$ matrices $S(T, M) = \{ U \in \mathbb{C}^{M \times T} | U U^H = I_M \}$ and the total volume of the Stiefel manifold is computed as

$$|S(T, M)| = \prod_{i=T-M+1}^{T} \frac{2\pi^{i}}{(i-1)!}.$$ 

**Definition 2 (Grassmann Manifold [27]):** The Grassmann manifold $G(T, M)$ for $T \geq M$ is defined as the quotient space of $S(T, M)$, which represents the set of all $M$-dimensional subspaces of $\mathbb{C}^T$. The volume of the Grassmann manifold...
equals
\[ |G(T, M)| = \frac{|S(T, M)|}{|S(M, M)|} = \frac{\prod_{i=T-M+1}^{T} 2^{\pi^2 i^2 / (i-1)}}{\prod_{i=1}^{M} 2^{\pi^2 i^2 / (i-1)}}. \]

**Definition 3 (Digamma Function [28]):** In mathematics, the digamma function \( \psi(x) \) is defined as the logarithmic derivative of the gamma function \( \Gamma(x) \).

Theorem 1: For a random Gaussian matrix \( S \in \mathbb{C}^{M \times T} \) whose entries are i.i.d. \( \mathcal{CN}(0, 1) \) for \( T \geq M \), the following equivalence holds:
\[ \ln \det(SS^H) = \ln \det(G) + \ln \det(T) + \ln \det(I) \]
where \( G \) is a unitary matrix with \( T \) rows, \( \Omega \) is an i.d. random unitary matrix that is independent of \( G \), and \( \Delta \) is a diagonal matrix containing non-zero singular values of \( G \).

**Proof:** See Appendix B.

In fact, the study on the differential entropy of \( \mathbf{Y}_E \) in (10) is not new [19, 20]. Proposition 1 is however tailored for our purposes, because the input distribution is fixed to the i.i.d. complex Gaussian rather than the unitary matrix as is the case for the non-secure communication systems [20]. In this proposition, we are also able to find a compact expression of non-coherent differential entropy, which is applicable to all antenna configurations with \( T \geq M \).

**Lemma 2:** For \( \mathbf{Y}_E \) in Proposition 1, we can establish the upper and lower bounds of \( \frac{1}{T} h(\mathbf{Y}_E) \). To get insight, let us investigate the gap between the bounds in (11) for a simple case of \( T = M \) and \( \alpha^2 = \beta^2 = 1 \), which is approximately expressed by
\[ \text{bound gap} = \frac{N_E}{T} (M \log T - \sum_{i=1}^{M} \psi(T - i + 1) \log e) \]
where (12) holds from \( \lim_{a \to \infty} \psi(a) = \ln a \). We see that the bound gap becomes smaller as the coherence time \( T \) exceeds \( M \). Thus, the upper and lower bounds in (11) are asymptotically tight for \( T \gg M \).

Now, to compute \( h(\mathbf{Y}_E|\mathbf{S}) \), we can rewrite \( \mathbf{Y}_E \) in (5) as
\[ \mathbf{Y}_E^T = \mathbf{N}^T \mathbf{G}_2^T + \mathbf{S}^T \mathbf{G}_1^T + \mathbf{Z}^T. \]

Then, the system can be interpreted as the fictitious case of the BS sending a virtual signal matrix \( \mathbf{G}_2^T \) through a random propagation matrix \( \mathbf{N}^T \) with effective noise \( \mathbf{S}^T \mathbf{G}_1^T + \mathbf{Z}^T \). For a given \( \mathbf{S} \), each column of the effective noise has an equal covariance matrix \( \alpha^2 \mathbf{S}^T \mathbf{S} + \sigma_0^2 \mathbf{I}_T \). Thus, we can perform the noise whitening as
\[ \mathbf{Y}_E^T \triangleq \mathbf{S}_E (\alpha^2 \mathbf{S}^T \mathbf{S} + \sigma_0^2 \mathbf{I}_T)^{-\frac{1}{2}} \mathbf{Y}_E^T \]
where (13) holds from \( \lim_{a \to \infty} \psi(a) = \ln a \). We see that the bound gap becomes smaller as the coherence time \( T \) exceeds \( M \). Thus, the upper and lower bounds in (11) are asymptotically tight for \( T \gg M \).

Now, to compute \( h(\mathbf{Y}_E|\mathbf{S}) \), we can rewrite \( \mathbf{Y}_E \) in (5) as
\[ \mathbf{Y}_E^T = \mathbf{N}^T \mathbf{G}_2^T + \mathbf{S}^T \mathbf{G}_1^T + \mathbf{Z}^T. \]

Then, the system can be interpreted as the fictitious case of the BS sending a virtual signal matrix \( \mathbf{G}_2^T \) through a random propagation matrix \( \mathbf{N}^T \) with effective noise \( \mathbf{S}^T \mathbf{G}_1^T + \mathbf{Z}^T \). For a given \( \mathbf{S} \), each column of the effective noise has an equal covariance matrix \( \alpha^2 \mathbf{S}^T \mathbf{S} + \sigma_0^2 \mathbf{I}_T \). Thus, we can perform the noise whitening as
\[ \mathbf{Y}_E^T \triangleq \mathbf{S}_E (\alpha^2 \mathbf{S}^T \mathbf{S} + \sigma_0^2 \mathbf{I}_T)^{-\frac{1}{2}} \mathbf{Y}_E^T \]
where (13) holds from \( \lim_{a \to \infty} \psi(a) = \ln a \). We see that the bound gap becomes smaller as the coherence time \( T \) exceeds \( M \). Thus, the upper and lower bounds in (11) are asymptotically tight for \( T \gg M \).
Here, we observe that the resulting sum of \((a)\) has rank \(N_J\), while the rank of \(V_{S_2}^*V_{S_2}^H N^*G_2^T\) equals \(\min(T - K, N_J)\). Therefore as long as \(T - K \geq N_J\) (i.e., \(T \geq M\)), the first term of \((14)\) becomes negligible as \(\sigma_0 \rightarrow 0\).

To simplify the expressions, we invoke the equivalence

\[
h(Y_E^T|S) = h(V_S^H Y_E),
\]

where

\[
V_S^H Y_E = \left[\begin{array}{c} Z_1^* \cr N_S^* G_2^T \end{array}\right] + Z_2^*
\]

with \(Z_1 = ZV_{S_1}\) and \(Z_2 = ZV_{S_2}\) having i.i.d entries, and the components of \(N_2 = NV_{S_2}\) being i.i.d. \(\mathcal{CN}(0,1)\), then the conditional entropy \(h(Y_E|S) = h(Y_E^T|S)\) is now computed by

\[
h(Y_E^T|S) = h(V_S^H Y_E) + NE \log \det \left(\alpha^2 S^T S^* + \sigma_2^2 I_T\right)
\]

\[
- N_E T \log \sigma_2^2
\]

\[
= h(Z_1) + h(G_2 N_2 + Z_2) + NE \log \det \left(\alpha^2 S^T S^* + \sigma_2^2 I_T\right)
\]

\[
- N_E T \log \sigma_2^2
\]

\[
= h(G_2 N_2 + Z_2) + KN_E \log \pi e \sigma_2^2 + NE \sum_{i=1}^{K} \log(\alpha^2 \lambda_{S,i}^2) + N_E T \log \sigma_2^2
\]

\[
+ N_E (T - K) \log \sigma_2^2 - N_E T \log \sigma_2^2
\]

\[
= h(G_2 N_2 + Z_2) + N_E \sum_{i=1}^{K} \log(\alpha^2 \lambda_{S,i}^2) + N_E T \log \sigma_2^2,
\]

Finally, by combining Propositions 1 and 2, we can evaluate the high SNR approximate to the non-coherent leakage rate as summarized in the following theorem.

**Theorem 1 (Non-Coherent Leakage Rate):** When the effective ED channels of both \(G_1\) and \(G_2\) are unknown to the ED with \(T \geq M\) and \(\sigma_2^2 \rightarrow 0\), the non-coherent leakage rate in \((6)\) is computed as

\[
L_{\text{non-coh}}(K, N_E, N_J, T)
\]

\[
= \min((N_E - N_J)\dagger, K) \left(1 - \frac{M}{T}\right)
\]

\[
\times \log \text{SNR} + c_{K,N_E,N_J,T,\text{non-coh}} + o(1),
\]

where \(c_{K,N_E,N_J,T,\text{non-coh}}\) denotes the constant that are irrespective of \(\sigma_2^2\), which is defined by

\[
c_{K,N_E,N_J,T,\text{non-coh}} \triangleq \frac{1}{T} h(G \Lambda_Q X) - K \frac{h(G_2 \Lambda_{N_2} Q)}{T} + d_{K,N_E,N_J,T,\text{non-coh}},
\]

and bounded by

\[
c_{K,N_E,N_J,T,\text{non-coh}} \leq \frac{N_E}{T} \left[ K \log \pi e \dfrac{MT}{K} + N_J \log \dfrac{MT}{K} \beta^2 \right]
\]

\[
- \sum_{i=K+1}^{2K} \varphi(T - i + 1) \log e
\]

\[
+ d_{K,N_E,N_J,T,\text{non-coh}},
\]

and

\[
c_{K,N_E,N_J,T,\text{non-coh}} \geq \frac{N_E}{T} \left[ K \log \pi e \sigma_2^2 + N_J \log \dfrac{\beta^2}{K - T} \right]
\]

\[
+ \sum_{i=1}^{K} \varphi(T - i + 1) \log e
\]

\[
+ d_{K,N_E,N_J,T,\text{non-coh}},
\]

with \(d_{K,N_E,N_J,T,\text{non-coh}}\) being defined in \((20)\), shown at the top of the next page. Note that the bounds \((18)\) and \((19)\) in Theorem 1 are based on the results in \((11)\) and \((17)\).

It is interesting to contrast the ergodic leakage rate in Lemma 1 and the non-coherent leakage rate in Theorem 1. As mentioned previously, the DoF in the ergodic leakage rate is independent of the coherence time \(T\), because the ED can exploit its full spatial DoF by using the CSIs. Therefore, a large number of ED antennas may be a security threat.

On the contrary, the result in Theorem 1 reveals that the channel uncertainty reduces the ED’s spatial DoF by the factor of \((1 - \frac{M}{T})\). This is because some of the space-time dimension of the ED channel is consumed for channel uncertainty resolution. Interestingly, if the BS increases the AN signal space dimension \(N_J\) such that \(M = T\) or more, we achieve the zero DoF leakage rate regardless of the number of ED antennas \(N_E\). Therefore, the proposed ANAM system having a large number...
of transmit antennas with no downlink training is inherently robust to the passive eavesdropping especially in a mobile environment where the coherence time $T$ is not too long.

Now, we investigate the leakage rate in Theorem 1 in terms of the constant value. Let us consider an example of $M = M = T$ and $\alpha^2 = \beta^2 = 1$. In this case, the DoF equals zero, and the leakage rate upperbound is saturated at high $\text{SNR}_E$ as

$$L_{\text{non-coh}}(K, N_E, N_J, T) \leq N_E \log T - N_E \log e \left\{ \sum_{i=1}^{M} \varphi(T - i + 1) \log e \right\}$$

where $\gamma \simeq 0.57721566$ denotes the Euler’s constant. We see that the constant value of the leakage rate is non-zero and linearly increases with the number of ED antennas. This implies that we may not be able to make the leakage rate completely (or asymptotically) zero even if $M$ exceeds the coherence time. Nevertheless, the leakage rate growth is marginal considering the LUs’ increasing data rate at high $\text{SNR}_E$. Also, we notice that the increase of LUs does not affect the total amount of leakage information to the ED. An interesting point here is that we do not assume any wiretap codes or the ED’s channel information at the BS, both of which are practically difficult to realize. Therefore, the ANAM system is of practical significance and has the potential of physical layer security.

V. PARTIALLY COHERENT LEAKAGE RATE

In this section, we consider a practical situation where the BS transmits the downlink training precoded by $H$ (the same precoder that is applied to the message signal $S$) to the LUs. The scheme is useful when the knowledge on the downlink CSI, i.e., $H$, at the BS is imperfect or outdated, because in this case the LUs need to estimate their effective channels more accurately. However, the downlink training may also give the ED a chance of estimating its own effective channel $G_1$. Therefore, it is also practically of interest to investigate the partially coherent leakage rate supposing that the effective ED channel $G_1$ is known to the ED whereas the AN channel $G_2$ is still unknown.

In order to estimate the fading coefficients of the $K$ different LUs, a training phase of duration should be no smaller than $K$ [29]. This represents the minimum cost for using the downlink training to the LUs, which reduces the effective coherence time for the data transmission from $T$ to $T' \leq T - K$. Thus, with the downlink training, the leakage rate can be reformulated by

$$L_{\text{partial-coh}}(K, N_E, N_J, T') = \frac{1}{T'} I(Y'_E; S'|G_1) \quad (21)$$

where $Y'_E = G_1 S' + G_2 N' + Z' \in \mathbb{C}^{N_E \times T'}$ denotes the received signals at the ED over the remaining coherence time $T'$ with input signals $S' \in \mathbb{C}^{K \times T'}$ (message) and $N' \in \mathbb{C}^{N_E \times T'}$ (AN) both having i.i.d. $\mathcal{CN}(0, 1)$ entries, and additive noise $Z' \in \mathbb{C}^{N_E \times T'}$ having i.i.d. $\mathcal{CN}(0, \sigma^2_Z)$ entries. As a worst case, we consider no channel estimation error at the ED. Nevertheless, a direct evaluation of (21) is difficult due to the hidden variables in $G_2$. To tackle the problem, we first investigate a universal upperbound that generally holds without any assumption. Then, we obtain tighter upper and lower bounds under the assumption that $N_E \geq M$, $T' \gg N_J$, and $\sigma^2_Z \to 0$.

A. Universal Upperbound

The following theorem evaluates the universal upperbound of the leakage rate in (21). We note that the universal bound is tight at the low $\text{SNR}_E$ regime.

Theorem 2 (Universal Upperbound): Supposing that the ED perfectly knows $G_1$ (not $G_2$), the partially coherent leakage rate to the ED is universally upperbounded by

$$L_{\text{partial-coh}}(K, N_E, N_J, T') \leq \min(N_E, K) \left( 1 - \frac{N_J}{T'} \right) + \min(N_E, K) \left[ \frac{1}{T'} \sum_{i=1}^{\min(T', N_J)} \sum_{j=1}^{\min(N_E, K)} \log \left( 1 + \frac{\lambda_{G_{1,i,j}}^2}{\beta^2 \lambda_{N'_{i,j}}^2 + \sigma^2_Z} \right) \right]$$

$$+ \min(N_E, K) \left[ \frac{1}{T'} \sum_{j=1}^{\min(N_E, K)} \log \left( \frac{\lambda_{G_{1,j}}^2 + \sigma^2_Z}{\lambda_{N'_{j}}^2} \right) \right]$$

where the equality holds at low $\text{SNR}_E$. Here, $c_{\text{univ}}(K, N_E, N_J, T', \text{partial-coh})$ represents the constant terms that are irrespective of $\sigma^2_Z$ as

$$c_{\text{univ}}(K, N_E, N_J, T', \text{partial-coh}) = \frac{1}{T'} \left[ \sum_{i=1}^{\min(T', N_J)} \sum_{j=1}^{\min(N_E, K)} \log \left( 1 + \frac{\lambda_{G_{1,i,j}}^2}{\beta^2 \lambda_{N'_{i,j}}^2 + \sigma^2_Z} \right) \right]$$

$$+ \left[ \frac{1}{T'} \sum_{j=1}^{\min(N_E, K)} \log \left( \frac{\lambda_{G_{1,j}}^2 + \sigma^2_Z}{\lambda_{N'_{j}}^2} \right) \right].$$

with $\lambda_{N'_{i,j}}$ for $i = 1, \ldots, \min(T', N_J)$ and $\lambda_{G_{1,j}}$ for $j = 1, \ldots, \min(N_E, K)$ denoting the non-zero singular values of $N'$ and $G_1$, respectively.

Proof: See Appendix D.
Corollary 1: In the low SNR regime, the partially coherent leakage rate is asymptotically equal to
\[
\mathcal{L}_{\text{partial-coh}}(K, N_E, N_J, T') \xrightarrow{\sigma^2 \to \infty} E \left[ \log (\mathbf{I}_K + \text{SNR}_E \mathbf{G}_1^\dagger \mathbf{G}_1) \right].
\]

Proof: Assuming that \(\sigma^2 \to \infty\), we have
\[
\mathcal{L}_{\text{partial-coh}}^{\text{univ}}(K, N_E, N_J, T', \text{partial-coh}) = \min(N_E, K) \left( 1 - \frac{N_J}{T'} \right) + \log \sigma^2_z
\]
\[
+ \left\{ \min(T', N_J) + \left( 1 - \frac{N_J}{T'} \right) + \right\} \sum_{j=1}^{\min(N_E, K)} \log \left( 1 + \frac{\lambda_2^2 \mathbf{G}_1 \mathbf{g}_j}{\sigma^2_z} \right)
\]
which leads to the fact that \(\mathcal{L}_{\text{partial-coh}}(K, N_E, N_J, T') \sigma^2 \to \infty\)
\[
= \mathcal{E} \left( \sum_{j=1}^{\min(N_E, K)} \log \left( 1 + \frac{\lambda_2^2 \mathbf{G}_1 \mathbf{g}_j}{\sigma^2_z} \right) \right).
\]
Since the universal bound in Theorem 2 is tight at the low SNR regime, we obtain the corollary.

The results in Theorem 2 and Corollary 1 provide useful insight into the system. First, observe that at low SNR (\(\sigma^2 \to \infty\)), the partially coherent leakage rate converges to the ergodic capacity of the ED channel with no AN, which implies that there is no security gain that stems from the AN and its channel uncertainty. In contrast, from the moderate to high SNR regime, we can reduce the leakage rate by increasing \(\beta^2\) or decreasing \(T\), which means that the AN and its channel uncertainty are still effective for emasculating the ED’s decoding ability even if \(\mathbf{G}_1\) is known to the ED. The universal bound in Theorem 2 also serves as an upperbound of the non-coherent leakage rate in Theorem 1, since the known CSI \(\mathbf{G}_1\) gives a benefit to the ED compared to its non-coherent counterpart.

B. Tight Upper and Lower Bounds With High SNR Approximation

In this subsection, we investigate tighter upper and lower bounds for the partially coherent leakage rate based on the high SNR approximation. By definition, we have
\[
\mathcal{L}_{\text{partial-coh}}(K, N_E, N_J, T') = \frac{1}{T'} h(\mathbf{Y}_E | \mathbf{G}_1) - \frac{1}{T'} h(\mathbf{Y}_E' | \mathbf{G}_1, \mathbf{S}')
\]
\[
= \frac{1}{T'} h(\mathbf{Y}_E | \mathbf{G}_1) - \frac{1}{T'} h(\mathbf{G}_2 \mathbf{N}'+ \mathbf{Z}')
\]
(23)
Then, the high SNR approximate to the first and second terms of (23) can be attained following the arguments in Theorem 1 and 2, respectively. The result is summarized below. Detailed proof is omitted here for brevity.

*Theorem 3 (Partially Coherent Leakage Rate):* Define \(\mathbf{G}_{2,2} \in \mathcal{C}^{N(E-K) \times N_J}\) as a random matrix having i.i.d \(\mathcal{C}^{N(0, \beta^2)}\) entries. Then, for \(N_E \geq M, T' \geq N_J\), and \(\sigma^2 \to 0\), the partially coherent leakage rate can be evaluated by
\[
\mathcal{L}_{\text{partial-coh}}(K, N_E, N_J, T') = K \left( 1 - \frac{N_J}{T'} \right) + \log \text{SNR}_E
\]
+ \(c_{K, N_E, N_J, T', \text{partial-coh}} + o(1)\)
where the constant term
\[
c_{K, N_E, N_J, T', \text{partial-coh}} \triangleq \frac{1}{T'} h(\mathbf{G}_{2,2} \mathbf{U}_N | \mathbf{N}, \mathbf{Q}) - \frac{1}{T'} h(\mathbf{G}_2 \mathbf{U}_N | \mathbf{N}, \mathbf{Q})
\]
+ \(d_{K, N_E, N_J, T', \text{partial-coh}}\)
is bounded by
\[
c_{K, N_E, N_J, T', \text{partial-coh}} \leq \frac{1}{T'} \left( N_J N_E \log T' - K N_J \log \pi e T' \beta^2
\]
\[
- N_E \sum_{i=1}^{N_J} \phi(T' - i + 1) \log e
\]
+ \(d_{K, N_E, N_J, T', \text{partial-coh}}\);
and
\[
c_{K, N_E, N_J, T', \text{partial-coh}} \geq \frac{1}{T'} \left( (N_E - K) \sum_{i=1}^{N_J} \phi(T' - i + 1) \log e
\]
\[
- N_E N_J \log T' - K N_J \log \pi e \beta^2
\]
+ \(d_{K, N_E, N_J, T', \text{partial-coh}}\);
with
\[
d_{K, N_E, N_J, T', \text{partial-coh}} \triangleq \left( 1 - \frac{N_J}{T'} \right) \left( \sum_{i=1}^{N_J} E \left[ \log \frac{\lambda_2^2 \mathbf{G}_{2,2} \mathbf{N}' \mathbf{g}_j}{\mathbf{G}_2 \mathbf{N}' \mathbf{g}_j} \right] - K \log \pi e \right)
\]
\[
+ \sum_{i=1}^{K} \phi(N_E - i + 1) \log e + K \log \pi e \beta^2.
\]

The result in Theorem 3 shows the exact DoF of the partially coherent leakage rate. From the result, we recognize that the universal bound in Theorem 2 is tight in terms of the DoF for the case of \(N_E \geq M\), but otherwise loose. A zero-DoF condition \(N_J \geq T'\) arises from Theorem 3. Interestingly, for \(T' = T - K\), it is equivalently \(M \geq T\) which is the same condition for the non-coherent ED. The bounds in Theorem 3 are tight for \(T' \gg M\). However, as we have noted in Section IV, when \(T'\) is compatible with \(M\), the bounds in Theorem 3 may be loose. In this case, one can find a tighter upperbound by combining the results in Theorems 2 and 3.

VI. DISCUSSION

In this section, we introduce useful design methods for the ANAM systems by leveraging the proposed analysis on the leakage rates.

A. How Many Antennas Do We Need for Secrecy?

In practice, a popular rule of thumb for determining the coherence time \(T\) in modern digital communications is \(T \approx \frac{234}{f_m}\) (sec) where \(f_m\) denotes the maximum Doppler frequency shift. In LTE-A, for example, one OFDM symbol duration is about 72.4\(\mu s\), which implies that the coherence symbol length \(T\) of the ED channel is approximately
\[
T = \frac{T_c}{72.4\mu s} = \frac{5842.5}{f_m},
\]
(24)
where we have $f_m = \frac{v}{c}$ for carrier frequency $f_c(\text{Hz})$, speed of a moving object $v(\text{m/s})$, and speed of light $c = 3 \times 10^8(\text{m/s})$.

Therefore, the BS can determine its number of antennas such that $M = \left\lceil \frac{\text{SNR}_{\text{ED}}}{\gamma} \right\rceil$ to achieve the zero-DoF leakage rate. Obviously as the ED’s mobility $v$ or the carrier frequency $f_c$ increases, the required number of BS antennas will decrease. For example, for $[f_c = 5 \text{ GHz}, v = 0.8 \text{ m/s}]$ and $[f_c = 10 \text{ GHz}, v = 1.3 \text{ m/s}]$, the BS may need $M = 351$ and $135$ number of antennas, respectively, both of which are reasonable numbers from the massive MIMO point of view.\(^3\)

**B. Artificial Fast Fading (AFF) Design**

When the coherence time $T$ is measured undesirably long, one can deliberately randomize the artificial noise channel $G_2$ by reformulating the transmit signal in (1) as

$$X = \alpha H S + \beta V h_2 N_j A(t) N, \quad (25)$$

where $A(t)$ denotes the AFF precoder that randomly changes in each time instant $t$ within the coherence time. Then, the received signal at Eve can be expressed by

$$Y_E = G_1 S + G_2(t) N + Z, \quad (26)$$

where $G_2(t) \triangleq G_2 A(t)$. Observe that as the changing period of $A(t)$ becomes shorter, the effective coherence time of $G_2(t)$ reduces, and therefore we obtain security gain.

A major difference point of (25) from the existing AFF schemes in [15]–[18] is that we apply the AFF precoder to the artificial noise term $N$, not to the message signal $S$. This is because applying the AFF precoder to $S$ is cost ineffective since an additional effort is needed to make the LU channels remain deterministic. In contrast, the proposed AFF precoder in (25) is free from such a condition because the AN will not appear in the LU channels, and thus the AFF precoder can be fully random. From the analysis in Section V, we know that reducing the coherence time of $G_2(t)$ is also an effective method for degrading the ED’s decoding ability even if $G_1$ is perfectly known to the ED. Unfortunately, however, a rigorous analysis for its leakage-rate is unavailable yet because independence between two fading channel matrices $G_2(t)$ and $G_2(t')$ for $t \neq t'$ is not guaranteed. We leave this topic as future works.

**C. Achievable Secrecy Rate**

With a knowledge on the ED’s channel statistics at the BS, it is also of interest to investigate the secrecy rate that specifies the secure data rate to the LUs with the aid of the wiretap codes at the BS.\(^4\) In general, the achievable secrecy rate is computed by the difference of the information rate to the LUs and the leakage rate to the ED for a given input distribution [13]. In what follows, we consider two scenarios, a single LU having $K$ antennas and $K$ LUs each having a single antenna. Note that the achievable secrecy rates are obtained based on the leakage rates upperbounds in Theorems 1 and 3 rather than the lowerbounds.

1) **Single-User Secrecy Rate:** For the case of a single LU with $K$ antennas, the wiretap code can be applied across the $K$ different data streams, because we have only one message to encode. Therefore, the achievable secrecy rate with non-coherent ED is simply given by

$$R_{\text{sec,non-coh}} = [K \log (1 + M \alpha^2 \text{SNR}_L)] - L_{\text{non-coh}}(K, N_E, N_J, T) +. \quad (27)$$

As for the partially coherent ED, the data transmission takes place over $T'$ symbol times during the coherent time $T$ due to the downlink training phase. Therefore, the secrecy rate is computed as

$$R_{\text{sec,partial-coh}} = \frac{T'}{T}[K \log (1 + M \alpha^2 \text{SNR}_L)] - L_{\text{partial-coh}}(K, N_E, N_J, T'). \quad (28)$$

2) **Multi-User Secrecy Rate:** For the case of multiple $K$ LUs with a single antenna, the wiretap code must be applied independently to each data stream, because otherwise the encoded data may not be decodable at each LU. A reasonable approach in this case is to consider the worst case scenario where the ED can intercept all other messages $M_1, \forall i \neq k$ when we encode the message for the $k$-th user. Thus, the achievable non-coherent and partially coherent secrecy-rates to the $K$ LUs are respectively computed as

$$R_{\text{sec,non-coh}} = K [\log (1 + M \alpha^2 \text{SNR}_L)] - L_{\text{non-coh}}(1, N_E, N_J, T)] +$$

and

$$R_{\text{sec,partial-coh}} = \frac{KT'}{T}[\log (1 + M \alpha^2 \text{SNR}_L)] - L_{\text{partial-coh}}(1, N_E, N_J, T'). \quad (29)$$

**VII. Numerical Results**

In this section, we investigate the leakage and secrecy rate performance of the TDD based ANAM systems through some numerical examples. Unless stated otherwise, we focus on a system with $M = 64, K = 16, \alpha = \beta = 1, \text{ and } T' = T - K$, but observations made in this section are generally applicable to other scenarios.

First, Figure 2 compares the leakage rate performance for the ANAM systems with $\text{SNR}_E = 30 \text{ dB}$, and $T = \gamma M$ for $\gamma \geq 1$. Here, “Ergodic Leakage” indicates the ergodic leakage rate in Lemma 1, which amounts to the worst case scenario in terms of security, and “UB” and “LB” denote the proposed upper and lower bounds of the leakage rates. It is seen that as the coherence time $T$ increases, the non-coherent leakage and the partially coherent leakage, both approach their full CSI counter part because a long channel coherence.
time gives enough time for the ED to perform the blind detection [11], [12]. The figure also confirms our statement in equation (12) that as the coherence time increases, the gap between the leakage rate bounds reduces. The bounds seems tight when $T > 5M$, at which we can estimate the exact amount of leakage to the ED. The partially coherent leakage is slightly higher than the non-coherent one due to $G_1$ that is known to the ED.

Figure 3 shows the leakage rate performance according to various $N_E$ and $N_J$ with $T = 5M$ and $SNR_E = 30$ dB. For the case of no AN ($N_J = 0$), we set $\alpha = \sqrt{M/K}$ according to (8). We observe that when $N_E \ll M$, the CSI at the ED may be no importance. However, as $N_E$ grows, the blindness of the ED has a critical impact on protecting the messages. The impact is maximized when the transmitted signals are corrupted by the AN, because the AN signals coming through random propagation amplifies the channel uncertainty at the non-coherent ED, and thus cannot be simply removed even if the ED utilizes a large number of receive antennas. This verifies the robustness of the ANAM systems to the large antenna array attack of the ED.

Figure 4 exhibits the leakage-rate bounds to the partially coherent ED according to various channel coherence time $T$ with $N_E = 64$. For the case of zero DoF leakage rate with $T = M$, we take the minimum between the bounds in
Theorem 2 and 3 to exhibit a tighter upperbound. Otherwise, the results are based on Theorem 3. The performance bounds in this figure are thus accurate in the high SNR region. We see that the AN channel uncertainty degrades the leakage rate DoF by a factor of $(1 - N_f/T')^+$ compared to its ergodic counter part, which leads to significant security advantage as SNR grows high. It is also interesting to observe that as in the case of non-coherent ED, the upper and lower bounds for the partially coherent ED are tight for $T > 5M$, while being looser when $T$ decreases toward $M$.

Figure 5 shows the performance trend of the universal upperbound in Theorem 2 with $T = M$, which is valid over all SNR range. First, the figure verifies our statement in Corollary 1 that when SNR is sufficiently small, the universal bound is tight and behaves like the ergodic leakage rate. In addition, we confirm from the figure that the partially coherent ED scenario still attains robustness to the ED’s large antenna array attack compared to the ergodic case.

Figure 6 illustrates the achievable secrecy rate to a single LU with $K$ antennas in the ANAM systems with $T = 2M$ and SNR$_E = $ SNR$_L = $ SNR. As expected, the LU with a non-coherent ED enjoys highest secrecy rate for all $N_E$ compared to other scenarios. One interesting observation is that the partially or non-coherent secrecy rates continuously grow with SNR regardless of $N_E$, whereas the ergodic secrecy-rate vanishes as $N_E$ increases from 64 to 128.

In Figure 7, we investigate the tendency of the secrecy-rate change according to the coherence time $T = \gamma M$ with SNR$_E = $ SNR$_L = $ 30 dB. We observe that the smaller the ED’s channel coherence time is, the higher the achievable secrecy-rate becomes. For example, the coherent time $T = 2M$ amounts to the ED’s mobility 1.3 km/h in the carrier frequency of $f_c = 10$ GHz as illustrated in Section VI-A, which implies that a small movement of the ED or environmental change around the ED may produce a high security gain in the ANAM systems.

Figure 8 compares the performance of single-user and multi-user secrecy rates with $T = 7M$ and SNR$_E = $ SNR$_L$. As explained in Section VI-C.2, the single-user secrecy-rate outperforms the multi-user secrecy-rate, because each encoding function for multiple LUs considers a genie aided ED that is more capable than the one for the single-user case.

VIII. Conclusion

In this paper, we have investigated the leakage-rate performance of the ANAM systems with a non-coherent passive ED in block fading, where CSI of the ED channels are unknown to both the BS and the ED. First, we analyzed tight upper and lower bounds of the non-coherent leakage rate with the high SNR approximation. Then, we derived the leakage rate to the coherent ED with partial CSI considering a situation in which the BS transmits the precoded downlink training. We also computed single- and multi-user secrecy rates that are attainable via appropriate wiretap codes at the BS. From our analysis, it was found that the conventional assumption on the full CSI ED may be too pessimistic to provide meaningful information on the security, because the channel uncertainty at the ED may significantly degrade the ED’s message interception ability. Finally, we provided several interesting observations from the numerical results.

Appendix

A. Proof of Property 1

Define $A \triangleq U_H[A_H \ 0]$ and $B \triangleq V_H^H$ for notational simplicity. Also, define $f_{B}(B)$ as the probability density function of a random matrix $B$. Then, $H \triangleq HQ$ if and only if $f_{A,B}(A,B) = f_{\bar{A},B}(A,B)$ with $B \triangleq BQ$ [31], which is equivalently

$$f_{A|B}(A|B = B) = f_{\bar{A}|B}(A|B = B)\frac{f_B(B)}{f_{\bar{B}}(B)}$$

$$= f_{A|B,Q}(A|B = BQ^{-1}, Q = Q)$$

$$= \bar{f}_{A|B}(A|B = BQ^{-1}),$$

(30)

where the normal letters $A$, $B$, and $Q$ represent deterministic realizations of the corresponding random matrices. Here, the second equality follows due to $B \triangleq \bar{B}$ and the last equality stems from the independence between $A$ and $Q$. We see that
the equation (30) holds for any values of $Q$, which implies that $A$ is independent of $B$, and thus the proof is concluded.

B. Proof of Proposition 1

Note that in this proof, we basically assume that $M \leq T$. Let us define SVD of $Y_E = U_E A_E V_E^H$ where $U_E \in \mathbb{C}^{N_E \times T}$ and $V_E \in \mathbb{C}^{T \times \Omega}$ are unitary matrices and $A_E \in \mathbb{C}^{\Omega \times \Omega}$ denotes a diagonal matrix having $\Omega \triangleq \min(N_E, T)$ non-zero singular values of $Y_E$ ordered in descending order, i.e., $\{\sigma_1, \ldots, \sigma_{\Omega}\}$ on its main diagonal. Also, observe that $Y_E$ is i.i.d., since its distribution is invariant over both left and right unitary transformations, which means that the singular vectors of $Y_E$ are i.i.d and independent of the singular values. Thus, by the SVD coordinate change [32], we write

$$h(Y_E) = h(U_E) + h(V_E) + h(\sigma_1, \ldots, \sigma_{\Omega})$$

$$+ \mathbb{E}[\log|J_{Y',\Omega}(\sigma_1, \ldots, \sigma_{\Omega})|]$$

$$= \log|S(N_E, \Omega)| + \log|S(T, \Omega)| + h(\sigma_1, \ldots, \sigma_{\Omega})$$

$$+ \mathbb{E}[\log|J_{Y',\Omega}(\sigma_1, \ldots, \sigma_{\Omega})|],$$

where $\Omega' \triangleq \max(T, N_E)$ and $J_{Y',\Omega}(\cdot)$ denotes the Jacobian that is induced by the change of SVD coordinate, which is defined by

$$J_{Y',\Omega}(\sigma_1, \ldots, \sigma_{\Omega}) \triangleq \prod_{i<j}(\sigma_i^2 - \sigma_j^2)^{\frac{1}{2}} \prod_{i=1}^{\Omega} \sigma_i^{2(\Omega'-\Omega)+1}.$$  

Now, the remaining problem is to find the differential entropy of the singular values of $Y_E$. To this end, we first introduce the following lemma.

**Lemma 3:** For the singular values of $Y_E$ with $G\bar{X}$ having rank $\Xi \triangleq \min(N_E, M)$, the following property holds as $\sigma_2^2 \to 0$,

$$\{\sigma_1, \ldots, \sigma_{\Xi}\} \xrightarrow{d} \{\mu_1, \ldots, \mu_{\Xi}\}$$

and

$$\{\sigma_{\Xi+1}, \ldots, \sigma_{\Omega}\} \xrightarrow{d} \{\mu_{\Xi+1}, \ldots, \mu_{\Omega}\}$$

where $\{\mu_1, \ldots, \mu_{\Xi}\}$ and $\{\mu_{\Xi+1}, \ldots, \mu_{\Omega}\}$ are the singular values of $G\bar{X}$ and an independent $(N_E-\Xi \times (T-\Xi))$ matrix $\tilde{Z}$ having i.i.d. $\mathcal{CN}(0, \sigma_2^2)$ entries, respectively.

**Proof:** By the circular symmetry of the noise matrix $Z$, the singular values of $Y_E$ has the same distribution with those of

$$Y_0 = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix},$$

where $D \in \mathbb{R}^{T \times T}$ denotes the non-zero singular value matrix of $G\bar{X}$, and $Z_{11} \in \mathbb{C}^{(\Xi \times (T-\Xi))}$, $Z_{12} \in \mathbb{C}^{(\Xi \times (T-\Xi))}$, $Z_{21} \in \mathbb{C}^{(N_E-\Xi \times (T-\Xi))}$, and $Z_{22} \in \mathbb{C}^{(N_E-\Xi \times (T-\Xi))}$ represents random matrices having i.i.d. $\mathcal{CN}(0, \sigma_2^2)$ entries. Now, consider an equation $f(\lambda) = \det(\lambda I_{N_E} - Y_0 Y_0^H)$ with the roots $(\sigma_1^2, \ldots, \sigma_{N_E}^2)$. Then, it can be simplified at high SNR $E$ as

$$f(\lambda) \sigma^2 = \det \begin{bmatrix} \lambda(-D - \bar{Z}^H_{21} D) & -DZ_{21}^H \\ -Z_{21}^H D & \lambda(-\bar{Z}^H_{11} - Z_{21}^H) \end{bmatrix} - \det(\lambda I_{N_E} - \bar{Z}^H_{22}) \det \begin{bmatrix} \lambda I_{N_E} - \bar{Z}^H_{22} & -Z_{22}^H \ 0 & \lambda I_{N_E} - Z_{22}^H \end{bmatrix}$$

$$+ Z_{21} D (D^2 - \lambda E)^{-1} DZ_{21}^H, \quad (31)$$

where the second equality follows from the Schur’s identity for determinant of a block matrix. To find the roots of $f(\lambda) = 0$, we observe that the first $\Xi$ roots are the entries in $D^2$. The remaining $N_E - \Xi$ eigenvalues asymptotically vanishes as $\sigma_2^2 \to 0$, which means that they are much smaller than the entries of $D$. Thus, we can approximate $(D^2 - \lambda E)$ as $D^2$ and the second determinant of (31) becomes $\det(\lambda E - \bar{Z}^H_{22} Z_{22}^H)$. Therefore, the remaining $N_E - \Xi$ eigenvalues of $Y_0 Y_0^H$ are approximately the eigenvalues of $Z_{22}^H Z_{22}^H$.

**Lemma 3** states that the two sets of singular values $\{\sigma_1, \ldots, \sigma_{\Xi}\}$ and $\{\sigma_{\Xi+1}, \ldots, \sigma_{\Omega}\}$ of $Y_E$ are asymptotically independent of each other. Thus, we have

$$h(Y_E) = \log|S(N_E, \Omega)| + \log|S(T, \Omega)| + h(\sigma_1, \ldots, \sigma_{\Xi})$$

$$+ h(\sigma_{\Xi+1}, \ldots, \sigma_{\Omega})$$

$$+ \mathbb{E}[\log|J_{Y',\Xi}(\sigma_1, \ldots, \sigma_{\Xi})|].$$

Note that the singular values of $\tilde{G}\tilde{X}$ and $\tilde{G}A_X$ are the same. Also, a matrix $G\Lambda_X Q$ is i.i.d., since $Q$ is independent of $G\bar{X}$. Thus, we can consider the following differential entropy via the SVD coordinate change as

$$h(\tilde{G}A_X Q) = \log|S(N_E, \Xi)| + \log|S(M, \Xi)| + h(\sigma_1, \ldots, \sigma_{\Xi})$$

$$+ \mathbb{E}[\log|J_{Y',\Xi}(\sigma_1, \ldots, \sigma_{\Xi})|], \quad (33)$$

where $\Xi' \triangleq \max(N_E, M)$. Also, we write

$$h(Z) = (N_E - \Xi)(T - \Xi) \log \pi e \sigma_2^2$$

$$+ \log|S(\Omega - \Xi, \Omega - \Xi)| + \log|S(\Omega' - \Xi, \Omega - \Xi)|$$

$$+ h(\sigma_{\Xi+1}, \ldots, \sigma_{\Omega})$$

$$+ \mathbb{E}[\log|J_{Z',\Xi}(\sigma_{\Xi+1}, \ldots, \sigma_{\Omega})|].$$

where the second equality follows from the SVD coordinate change of $Z$. Then, combining the three equations from (32) to (34), we get

$$h(Y_E) = h(GA_X Q) + (N_E - \Xi)^+ (T - \Xi) \log \pi e \sigma_2^2$$

$$+ \mathbb{E}[\log|J_{T,N_E}(\sigma_1, \ldots, \sigma_{N_E})|]$$

$$- \mathbb{E}[\log|J_{T,N_E}(\sigma_1, \ldots, \sigma_{N_E})|]$$

$$+ \mathbb{E}[\log|J_{Z',\Xi}(\sigma_{\Xi+1}, \ldots, \sigma_{\Omega})|]$$

$$+ \log|S(N_E, \Xi)| + \log|S(T, \Omega)|$$

$$- \log|S(N_E, \Xi)| - \log|S(M, \Xi)|$$

$$- \log|S(\Omega - \Xi, \Omega - \Xi)|$$

$$- \log|S(\Omega' - \Xi, \Omega - \Xi)|.$$  

(35)

First, for $N_E \leq M$, it follows that

$$h(Y_E) = h(GA_X Q) + \mathbb{E}[\log|J_{T,N_E}(\sigma_1, \ldots, \sigma_{N_E})|]$$

$$+ \log|S(T, N_E)| - \log|S(M, N_E)|$$

$$= h(GA_X Q) + \log \left[ \frac{|G(T, N_E)|}{|G(M, N_E)|} \right]$$

$$+ \sum_{i=1}^{N_E} \mathbb{E} \left[ \log \sigma_i^{2(T-N_E)+1} - \log \sigma_i^{2(M-N_E)+1} \right]$$

$$= h(GA_X Q) + \log \left[ \frac{|G(T, N_E)|}{|G(M, N_E)|} \right]$$

$$+ (T - M) \mathbb{E} \left[ \sum_{i=1}^{N_E} \log \sigma_i^2 \right].$$  

(36)
Similarly, for $N_E > M$, we can show that

$$h(Y_E) = h(GA_XQ) + (N_E - M)^+ (T - M) \log \pi e \sigma^2_z$$

$$+ \log |G(T, M)| + (T - M)E \left[ \sum_{i=1}^{M} \log \sigma^2_i \right].$$ (37)

Finally, combining (36) and (37), we obtain Proposition 1.

C. Proof of Lemma 2

First, we compute the sum power of the elements in $\hat{G}A_XQ$ as $E[\text{Tr}(GU_XA_XU_H)] = E[\text{Tr}(GXXH)] = TN_E(\alpha^2 + \beta^2 N_j)$ by Properties 2 and 3, and the power constraint in (8). Thus, the differential entropy $h(GA_XQ)$ is maximized by $N_E \times M$ matrix with i.i.d. $CN(0, \frac{M}{N_E})$ entries, which leads to an upper bound

$$h(GA_XQ) \leq MN_E \log \pi e \frac{M}{N_E}.$$

Now, the lower bound can be computed as

$$h(GA_XQ) \geq h(GU_XA_XQX, Q)$$

$$= h(G) + N_EE \left[ \log \det (XX^H) \right]$$

$$= N_E(K \log \pi e + N_j \log \pi e/\beta^2)$$

$$+ N_E \sum_{i=1}^{M} \varphi(T - i + 1) \log e$$

where the last equality follows from Definition 3, and the proof is completed.

D. Proof of Theorem 2

First, let us introduce the following lemma.

Lemma 4: For $Y_E$ in (21), the following bound holds as

$$I(Y_E'; N'|G_1) \leq I(Y_E', N'|G_1, S')$$

which is tight as $\sigma^2_z \to \infty$.

Proof: First, observe that $I(Y_E'; N'|G_1)$ amounts to a fictitious case of the BS sending the signal $N$ to the ED at the existence of additive noise $G_1S' + Z'$ where $G_1$ is known to the ED. Therefore, the leakage rate will further increase if the ED knows both $G_1$ and $S'$ because in this case the effective noise reduces to $Z$, which is the case of $I(Y_E'; N'|G_1, S')$. When $\sigma^2_z \to \infty$, $G_1S'$ is ignorable relative to $Z$, which implies that $I(Y_E'; N'|G_1) \simeq I(Y_E'; N'|G_1, S')$. Therefore, the bound in (38) is tight in the low SNR$_E$ regime.

Now, we can verify the following inequality as

$$I(Y_E'; S', G_2; G_1)$$

$$= I(Y_E'; S', G_2, N'|G_1) - I(Y_E'; G_2, N'|G_1, S')$$

$$= I(Y_E'; S', G_2, G_1, N') - I(Y_E'; G_2, N', G_1, S')$$

$$+ I(Y_E'; G_2, N'|G_1) + I(Y_E'; N'|G_1, S')$$

$$\leq I(Y_E'; S', G_2, G_1, N') - I(Y_E'; G_2, N', G_1, S'),$$ (39)

where the first two equalities follow from the chain rules and the last inequality stems from Lemma 4. Note that the bound in (39) is universal and tight at low SNR$_E$.

The universal bound in (39) can be further evaluated as in the following. First, we consider that

$$I(Y_E'; G_2|N', G_1, S')$$

$$= I(N'^T G_2^T + Z'^T; G_2^T |N')$$

$$= N_EE \left[ \log \det \left( \beta^2 N'^T N'^* + \sigma^2_z I_T \right) \right]$$

$$- N_E T' \log \sigma^2_z$$

$$= N_E \min(N_j, T') \log \text{SNR}_E$$

$$+ N_EE \left[ \sum_{i=1}^{\min(N_j, T')} \log \left( \beta^2 \lambda^2_{N,i} + \sigma^2_z \right) \right]$$ (41)

where the second equality is due to the fact that (40) represents the information rate in a virtual MIMO channel where a transmitter with $N_j$ antennas sends the i.i.d. $CN(0, \beta^2)$ signal matrix $G_2^T$ during $N_E$ symbol times through the random propagation $N'^T$ that is known to the ED.

Now, to evaluate $I(Y_E'; S', G_2, G_1, N')$ in (39), let us write $Y_E'$ in the vectorization form as

$$\text{vec}(Y_E') = (I_T \otimes G_1) \text{vec}(S')$$

$$+ (\alpha N'^T \otimes I_{N_E}) \text{vec}(G_2/\alpha) + \text{vec}(Z')$$

$$= [I_T \otimes G_1 \beta N'^T \otimes I_{N_E}] \begin{bmatrix} \text{vec}(S') \text{vec}(G_2/\beta) \end{bmatrix} + \text{vec}(Z')$$ (42)

where the components of $S'$ and $G_2/\beta$ are i.i.d. $CN(0, 1)$.

Property 4: As for the Kronecker product, the following properties hold $A \otimes (B \circ D) = AC \otimes BD$ and $(A \otimes B)^H = A^H \otimes B^H$. Property 5: For two square matrices $A \in \mathbb{C}^{p \times p}$ and $B \in \mathbb{C}^{q \times q}$, the Kronecker sum is defined as $A \oplus B \triangleq A \otimes I_q + I_p \otimes B$. Then, the $(ij)\text{-th}$ eigenvalue of $A \oplus B$ equals $a_i^2 + b_j^2$ where $a_i^2$ and $b_j^2$ denote the $i\text{-th}$ and $j\text{-th}$ eigenvalues of $A$ and $B$, respectively.

Then, utilizing the above two properties, we can verify the following equivalences

$$I(Y_E'; S', G_2, G_1, N')$$

$$= I(\text{vec}(Y_E'); \text{vec}(S'), \text{vec}(G_2/\beta) \otimes G_1, N')$$

$$= E \left[ \log \det \left( (I_T \otimes G_1) \text{vec}(G_2/\beta) \otimes I_{N_E} \right) + \sigma^2_z I_{N_E} \right]$$

$$- N_E T' \log \sigma^2_z$$

$$= E \left[ \log \det \left( \beta^2 N'^T N'^* + G_1 \text{vec}(G_2/\beta) \right) + \sigma^2_z I_{N_E} \right]$$

where the rank of a kronecker sum $\beta^2 N'^T N'^* + G_1 \text{vec}(G_2/\beta)$ equals $N_E (T' - N_j)^+]$, which denotes the number of zero eigenvalues. Thus, we have

$$I(Y_E'; S', G_2; G_1, N')$$

$$= (N_E T' - R) \log \text{SNR}_E$$

$$+ \left[ \min(T', N_j) \min(N_E, K) \right]$$

$$+ \left( N_E - K \right) E \left[ \sum_{i=1}^{\min(T', N_j)} \sum_{j=1}^{\min(N_E, K)} \log \left( \beta^2 \lambda^2_{N,i} + \sigma^2_z \right) \right]$$

$$+ (T' - N_j)^+ E \left[ \sum_{i=1}^{\min(N_E, K)} \log \left( \lambda^2_{G_1,i} + \sigma^2_z \right) \right].$$ (44)

Finally, combining the results in (41) and (44), we obtain the theorem.
