Covert Communications with a Full-Duplex Receiver in Non-Coherent Rayleigh Fading

Mengfan Zheng, Cong Ling and Alexander Hamilton

Abstract

In a majority of existing works on covert communications, knowledge about instantaneous channel state information (CSI) of the main channel and/or the warden channel is assumed to be known or partially known. However, due to the covertness requirement, a covert user may not afford to perform channel estimation in practice, and acquiring the warden channel's CSI is even impossible. In this paper, we investigate the problem of covert communications over non-coherent Rayleigh fading channels, in both the i.i.d. fast fading case and slow fading case. Considering that the purpose of covert communication in many scenarios is to hide the existence of the sender, not the receiver, we allow the receiver to work in a full-duplex mode such that it transmits some artificial noise (AN) while receiving radio signals. We analyse the achievable covert rates with fixed and varying AN power and show that in both the fast and slow fading cases, it is possible to achieve a positive covert rate with carefully designed AN strategy.

Index Terms

Covert communication, fading channels, full-duplex, artificial noise, physical layer security

I. Introduction

Security has always been a paramount concern in communication system design. Traditional security methods based on cryptography can not solve all security and privacy problems. Even if an encryption algorithm is unbreakable, the pattern of network traffic can reveal some sensitive information. For example, the fact that a terminal is transmitting signal may already raise suspicion. Also, the transmission activity may expose a terminal's position. If users hope to protect

M. Zheng and C. Ling are with the Department of Electrical and Electronic Engineering at Imperial College London, United Kingdom. Emails: {m.zheng,c.ling}@imperial.ac.uk.

A. Hamilton is with the Defence Science and Technology Laboratory (Dstl). Email: ajhamilton@dstl.gov.uk.

their activity or location privacy, they also need to prevent the adversary from detecting their transmission attempts. Previously, such situations are more appealing to military applications. However, as the Internet of Things (IoT) expands rapidly, they have also attracted interests from more general applications, especially for those ultra-low power terminals as complex encrypting processes could consume a large amount of power.

Covert communication, or low probability of detection (LPD) communication, aims at hiding communications from a watchful adversary. The seminal work of Bash et al. [1] established the fundamental limits of covert communications over additive white Gaussian noise (AWGN) channels. It is shown that using a pre-shared key between the covert sender Alice and the receiver Bob, it is possible to transmit $O(\sqrt{N})$ bits reliably and covertly to Bob over N channel uses such that the warden Willie cannot detect the existence of communication. This is known as the *square root law* (SLR) for covert communication. Later, information-theoretic limits of covert communication are further investigated for both discrete memoryless channels (DMC) and AWGN channels in [2]–[4], and the optimal signalling schemes are studied in [5], [6] for AWGN channels and in [7] for DMCs. The study of covert communication has also been extended to multi-user channels such as broadcast channels [8], multiple-access channels (MAC) [9] and relay channels [10]. In all these types of channels, the SRL holds, implying a discouraging fact that the asymptotic covert communication rate is 0 as N goes to infinity. There are also works attempting to identify scenarios in which the SRL may be broken. For example, when the channel statistics are imperfectly known [11]–[13], or when the warden has uncertainty about the time of communication [14]–[16], or when artificial noise (AN) is injected into the channel [17], [18], positive covert rates may be achieved.

Most of the existing works on covert communication have assumed some kind of known channel state information (CSI) (of either the main channel or the warden channel or both) by the covert sender. However, in many practical scenarios, it is impossible to obtain the CSI of the warden channel, and performing channel estimation of its own channel may already expose the covert user. The best that one can do is to estimate the channel distribution information (CDI) based on some empirical formulas according to the communication environment. For this reason, studying covert communications over non-coherent fading channels has practical significance. The covert capacity of non-coherent i.i.d. Rayleigh fading channel has been characterized in a recent paper of [19], which shows that the covert capacity is achieved with an amplitude-constrained input distribution that consists of a finite number of mass points including one at zero. Still, covert communication in this channel is governed by the SRL. Reference [20] considers covert communication in the case when only CDI is known to the warden and shows that with the help of a full-duplex receiver that transmits random AN, covert communication rate can be improved. Nevertheless, CSI of the main channel is still assumed to be known by the legitimate users.

In this paper, we consider covert communications over non-coherent Rayleigh fading channels with a full-duplex receiver, who emits AN with fixed or varying power while receiving signals simultaneously. Our work differs from existing ones on covert communications with a full-duplex receiver (e.g., [20]–[23]) in that both the legitimate users and the warden only have access to CDI of main and warden channels. Two different cases, namely i.i.d. fast fading and slow fading, are investigated under fixed and varying AN power assumptions. AN has been shown to be an effective way to boost covert communication throughputs or even achieve positive covert communication rates. An additional benefit of using a full-duplex receiver is that the receiver can perform self-interference cancellation so as to further improve covert communication. Note that in many cases the goal of covert communication is to hide the existence of the sender, not the receiver. As a matter of fact, the existence of the receiver (such as a communication tower) may already be known by the warden. Thus, the considered scenarios are very practical.

The main contributions of this paper are summarized as follows.

- For the i.i.d. fading and fixed AN power case, we analyse the covert capacity and show that AN improves covert throughput only when it harms the warden more than it does to the receiver. Thus, the receiver's ability of self-interference cancellation is a key to enhancing covert communication.
- For the i.i.d. fading and varying AN power case, we analyse the false alarm rate and missed detection rate of Willie and show that for any covertness requirement, it is possible to achieve a positive covert communication rate.
- For the slow fading case, we derive the optimal detecting threshold of Willie in both the fixed and varying AN power cases, and find the maximum transmitting power Alice can use. Our result shows that in both cases, a positive covert communication rate is achievable.

The rest of this paper is organized as follows. In Section II, we introduce the channel model for covert communication over non-coherent Rayleigh fading channels with a full-duplex receiver. In section III and IV we study the i.i.d. fast fading case and the slow fading case, respectively. Two types of AN strategies, i.e., fixed AN power and varying AN power, are investigated. Section V

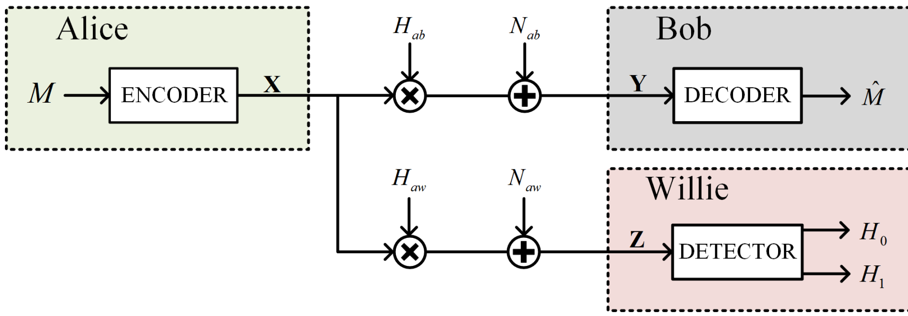

Fig. 1: The covert Rayleigh fading channel.

concludes this paper with some discussions.

Notation: Random variables are denoted by uppercase letters, and their realizations are denoted by the corresponding lowercase ones. Scalar variables are denoted by italic symbols, and vectors are denoted by boldface ones. Given a complex vector $\mathbf{X}$, $\mathbf{X}^\dagger$ denotes the conjugate transpose, and $|\mathbf{X}|$ its modulus. $[N]$ is the abbreviation for index set $\{1, 2, ..., N\}$.

II. System Model and Problem Statement

A. Channel Model

We consider covert communications over the Rayleigh fading channel as shown in Fig. 1, where Alice and Bob are the legitimate users, while Willie is a warden who tries to detect whether Alice is communicating with Bob or not. Alice encodes its covert message M into a length-N codeword $\mathbf{X} = (X^{(1)}, ..., X^{(N)})$ and sends it into the channel. Bob receives $\mathbf{Y} = (Y^{(1)}, ..., Y^{(N)})$ and recovers the message as $\hat{M}$. Willie observes $\mathbf{Z} = (Z^{(1)}, ..., Z^{(N)})$ and performs a statistical hypothesis test to determine whether Alice is communicating or not, where the null hypothesis $\mathcal{H}_0$ indicates that Alice is not communicating and the alternate hypothesis $\mathcal{H}_1$ the converse. In the i.i.d. fading case, channel gains change symbol by symbol. The input-output relationships at each time instant i are given by

$$Y^{(i)} = H_{ab}^{(i)} X^{(i)} + N_{ab}^{(i)}, \tag{1}$$

$$Z^{(i)} = H_{aw}^{(i)} X^{(i)} + N_{aw}^{(i)}, \tag{2}$$

where fading coefficients $H_{ab}^{(i)}$ and $H_{aw}^{(i)}$ are independent complex circular Gaussian random variables with zero mean and variances θ_{ab}^2 and θ_{aw}^2, respectively, and noises $N_{ab}^{(i)}$ and $N_{aw}^{(i)}$ are independent complex circular Gaussian random variables with zero mean and variances σ_b^2 and

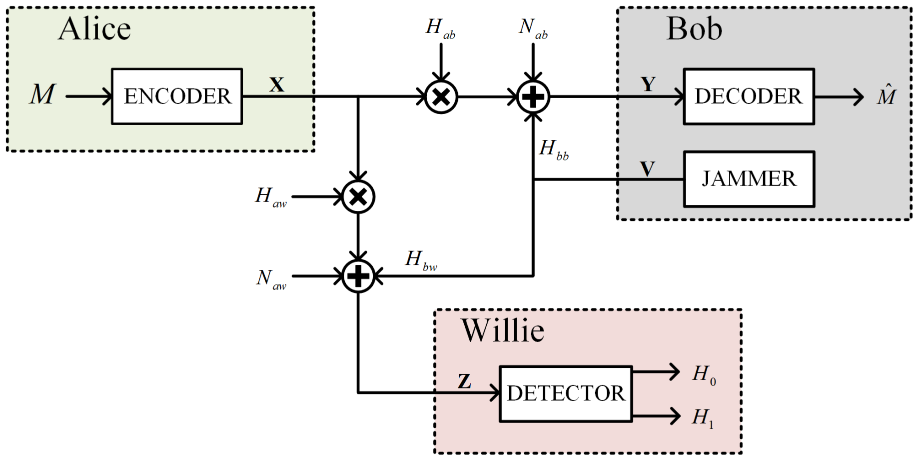

Fig. 2: The covert Rayleigh fading channel with a full-duplex receiver.

σ_w^2, respectively. In the slow fading case, channel gains remain constant during each transmission block of length N. The input-output relationships are given by

$$\mathbf{Y} = H_{ab}\mathbf{X} + \mathbf{N}_{ab}, \tag{3}$$

$$\mathbf{Z} = H_{aw}\mathbf{X} + \mathbf{N}_{aw}. \tag{4}$$

In the non-coherent setting, the fading coefficients are unknown to all parties, and only the statistical distributions of them are available. Since the phases of the fading coefficients are uniform, only the magnitude of X can carry useful information. Thus, $|Y|^2$ and $|Z|^2$ are sufficient statistics for detection.

In this paper, we have assumed that Bob works in a full-duplex mode such that it sends AN signal $\mathbf{V} = (V^{(1)}, ..., V^{(N)})$ while receiving signals from Alice, as shown in Fig. 2. The input-output relationships in the i.i.d. and slow fading cases are respectively given by

$$Y^{(i)} = H_{ab}^{(i)} X^{(i)} + \sqrt{\phi} H_{bb}^{(i)} V^{(i)} + N_{ab}^{(i)}, \tag{5}$$

$$Z^{(i)} = H_{aw}^{(i)} X^{(i)} + H_{bw} V^{(i)} + N_{aw}^{(i)}, \tag{6}$$

and

$$\mathbf{Y} = H_{ab}\mathbf{X} + \sqrt{\phi} H_{bb}\mathbf{V} + \mathbf{N}_{ab}, \tag{7}$$

$$\mathbf{Z} = H_{aw}\mathbf{X} + H_{bw}\mathbf{V} + \mathbf{N}_{aw}, \tag{8}$$

where the fading coefficient H_{bb} and H_{bw} are independent complex circular Gaussian random variables with zero mean and variances θ_{bb}^2 and θ_{bw}^2, respectively, and $\phi \in [0, 1]$ is the self-interference cancellation coefficient, with $\phi = 0$ corresponding to perfect self-interference cancellation and $\phi = 1$ no self-interference cancellation. The transmit power is denoted by $P_a = \mathbb{E}[X^{(i)}X^{(i)\dagger}]$ and the AN power by $P_b = \mathbb{E}[V^{(i)}V^{(i)\dagger}]$.

B. Willie's Detection Performance

Willie's detector may have two type of errors, namely the rejection of $\mathcal{H}_0$ when it is true, known as *false alarm*, the probability of which is denoted by α, and the acceptance of $\mathcal{H}_0$ when it is false, known as *missed detection*, the probability of which is denoted by β. The lower bound on the sum $\alpha + \beta$ characterizes the necessary tradeoff between the false alarms and the missed detections in the design of a hypothesis test. Let P_0 and P_1 be the probability distributions of Willie's channel observations when Alice is inactive and active, respectively. Willie can construct an optimal statistical hypothesis test (such as the Neyman–Pearson test) that minimizes the sum of error probabilities $\alpha + \beta$. The following holds for such an optimal test [1]:

$$\alpha + \beta = 1 - \mathbb{V}(P_0, P_1), \tag{9}$$

where $\mathbb{V}(\cdot, \cdot)$ is the *variational distance* between two distributions P_X and P_Y, defined as

$$\mathbb{V}(P_X, P_Y) \triangleq \frac{1}{2}\int_{\mathbb{R}^n} |P_X(\mathbf{x}) - P_Y(\mathbf{x})| d\mathbf{x}.$$

To limit Willie's detector performance, Alice should design her transmission scheme carefully such that the variational distance is negligible. This assures that no hypothesis tests will work for Willie: he either suffers from a high probability of false alarm, or from a high probability of missed detection. We can also use the Kullback-Leibler (KL) divergence (aka relative entropy) $\mathbb{D}(P\|Q)$ to measure covertness. The KL divergence of distributions P_X and P_Y is defined as [24]

$$\mathbb{D}(P_X\|P_Y) = \int_{\mathbb{R}^n} P_X(\mathbf{x}) \log \frac{P_X(\mathbf{x})}{P_Y(\mathbf{x})} d\mathbf{x}.$$

Pinsker's inequality (see [25, pp.58-59]) asserts that

$$\mathbb{V}(P, Q) \le \sqrt{\frac{1}{2}\mathbb{D}(P\|Q)}$$

for any distributions P and Q.

C. Covert Capacity of Non-Coherent I.I.D. Rayleigh Fading Channel

In the i.i.d. fading case, by relabelling $|X|^2$, $|Y|^2$ and $|Z|^2$ as X, Y and Z respectively, we obtain an effective channel with transition probabilities

$$W_{Y|X}(y|x) = \frac{1}{\theta_{ab}^2 x + \sigma_b^2} \exp\Big(- \frac{y}{\theta_{ab}^2 x + \sigma_b^2}\Big), \tag{10}$$

$$W_{Z|X}(z|x) = \frac{1}{\theta_{aw}^2 x + \sigma_w^2} \exp\Big(- \frac{z}{\theta_{aw}^2 x + \sigma_w^2}\Big). \tag{11}$$

For brevity, we can assume that $\sigma_b = \sigma_w = 1$ by properly normalizing Y and Z, and $\theta_{aw} = 1$ by properly normalizing X. Then we obtain a simplified channel parameterized by a single parameter θ'_{ab}:

$$W_{Y|X}(y|x) = \frac{1}{\theta_{ab}'^2 x + 1} \exp\Big(- \frac{y}{\theta_{ab}'^2 x + 1}\Big), \tag{12}$$

$$W_{Z|X}(z|x) = \frac{1}{x + 1} \exp\Big(- \frac{z}{x + 1}\Big), \tag{13}$$

where

$$\theta'_{ab} = \frac{\theta_{ab}^2}{\theta_{aw}^2}. \tag{14}$$

It is known that the SRL exists in many covert communication scenarios, meaning that the covert communication rate vanishes as the code length goes large. Thus, the traditional definition of channel capacity is unsuitable for covert communications. Instead, the *covert throughput* is defined as

$$T_{covert} = \frac{\log M_s}{\sqrt{N\delta}},$$

where M_s is the size of the message set, N is the code length, and δ is the target KL divergence between the innocent channel output distribution $Q_0^{\otimes N}$ and the induced distribution $P_{\mathbf{Z}}$ by the coding scheme. For any $\delta > 0$, a covert throughput T_{covert} is said to be achievable if there exist (M_s, N) codes of increasing block length N such that

$$\lim_{n\to\infty} \frac{\log M_s}{\sqrt{N\delta}} \geq T_{covert},$$

$$\lim_{n\to\infty} P(M \neq \hat{M}) = 0,$$

$$\lim_{n\to\infty} \mathbb{D}(P_{\mathbf{Z}} || Q_0^{\otimes N}) \leq \delta.$$

The supremum of all achievable covert throughputs is called the covert capacity. For the non-coherent i.i.d. Rayleigh fading channel, the covert capacity has been characterised in [19] in the following theorem.

Theorem 1 ([19]). *Let $\Omega_a(t)$ be the set of all discrete probability measures μ over $\mathcal{X}$ such that (i) μ has a finite number of mass points; (ii) $\mu(\{0\}) > 0$; (iii) $\sup(\text{support}(\mu)) \leq a$; (iv) $0 < \mathbb{D}(W_{Z|X} \circ \mu || Q_0) \leq t$[1]. Define*

$$\psi_a(t) \triangleq \sup_{\mu \in \Omega_a(t)} \frac{I(\mu, W_{Y|X})}{\sqrt{\mathbb{D}(W_{Z|X} \circ \mu || Q_0)}}. \tag{15}$$

For any $\zeta > 0$, the covert capacity of the non-coherent Rayleigh fading channel is

$$\begin{aligned} C_{no-CSI} &= \inf_{t>0} \psi_{1+\zeta}(t) \\ &= \sup_{\{\mu_N\}_{N\geq 1}: \forall N \geq 1, \mu_N \in \Omega_{1+\zeta}(\frac{\delta}{N})} \liminf_{N\to\infty} \frac{I(\mu_N, W_{Y|X})}{\sqrt{\mathbb{D}(W_{Z|X} \circ \mu_N || Q_0)}}. \end{aligned} \tag{16}$$

In addition, the following simple bounds hold:

$$\max_{\tilde{x} \in [0,1]} \sqrt{2(1-\tilde{x}^2)} \big(\theta_{ab}^{\prime 2} \tilde{x} - \log(1 + \theta_{ab}^{\prime 2} \tilde{x})\big) \leq C_{no-CSI} \leq \sqrt{2} \theta_{ab}^{\prime 2}. \tag{17}$$

III. I.I.D. Fading Channel

In this section we discuss AN strategies in the i.i.d. Rayleigh fading channel. We start by considering a simple case of fixed-power (more specifically, fixed-amplitude) AN. In this case, due to the Gaussian nature of the channel gain, the AN creates another Gaussian noise in addition to the background noise, and we can analyse the covert capacity based on the result of [19]. Then we consider the varying AN power case. By analysing Willie's false alarm probability (FAP) and missed detection probability (MDP), we derive the maximum transmit power of Alice under a given covertness requirement and discuss the effect of AN on covert communication rates.

A. Fixed AN Power: Is AN Always Helpful?

We first consider a simple case that AN has fixed amplitude and phase, i.e., $V = \sqrt{P_b} e^{\rho j}$, where $\rho \in [0, 2\pi]$ can be arbitrary. In this case, the AN item seen by Willie is still complex Gaussian due to the channel coefficient.

[1] $W_{Z|X} \circ \mu$ denotes the induced marginal distribution of Z by a probability measure μ on $\mathcal{X}$.

1) Covert Capacity Analysis: In the non-coherent setting, since the phases of H_{ab} and H_{aw} are uniform, only the magnitude of X can carry useful information. Thus, $|Y|^2$ and $|Z|^2$ are sufficient statistics for detection. By relabelling $|X|^2$, $|Y|^2$ and $|Z|^2$ by X, Y and Z respectively, we obtain an effective channel with transition probability

$$W_{Y|X}(y|x) = \frac{1}{\theta_{ab}^2 x + \phi\theta_{bb}^2 P_b + \sigma_b^2} \exp\Big(-\frac{y}{\theta_{ab}^2 x + \phi\theta_{bb}^2 P_b + \sigma_b^2}\Big), \tag{18}$$

$$W_{Z|X}(z|x) = \frac{1}{\theta_{aw}^2 x + \theta_{bw}^2 P_b + \sigma_w^2} \exp\Big(-\frac{z}{\theta_{aw}^2 x + \theta_{bw}^2 P_b + \sigma_w^2}\Big). \tag{19}$$

Similar to the discussion in the previous section, by properly normalizing X, Y and Z, we can assume $\sigma_b = \sigma_w = 1$ and obtain a simplified channel parameterized by a single parameter θ'_{ab}:

$$W_{Y|X}(y|x') = \frac{1}{\theta'^2_{ab} x' + 1} \exp\Big(-\frac{y}{\theta'^2_{ab} x' + 1}\Big), \tag{20}$$

$$W_{Z|X}(z|x') = \frac{1}{x+1} \exp\Big(-\frac{z}{x+1}\Big), \tag{21}$$

where

$$\theta'_{ab} = \frac{\theta_{ab}^2(\theta_{bw}^2 P_b + 1)}{\theta_{aw}^2(\phi\theta_{bb}^2 P_b + 1)} \tag{22}$$

and the equivalent transmit power is $P'_a = \frac{\theta_{aw}^2}{P_b+1} P_a$.

Comparing (14) and (22), we can find that if the AN power is fixed, the use of AN does not always improve covert throughput as the equivalent channel coefficient θ'_{ab} may not be increased. Only when Bob's residual self-interference is below a certain level, i.e.,

$$\phi\theta_{bb}^2 < \theta_{bw}^2, \tag{23}$$

can the covert throughput be enhanced. Note that we can also separate the jammer from Bob, which corresponds to the external jammer scenario. In this case Bob cannot do self-interference cancellation, i.e., $\phi = 1$, and therefore, AN can help improve covert throughput if and only if the jammer-Willie channel is better than the jammer-Bob channel.

2) Numerical Results: As an example, assume that $\theta_{ab}^2/\theta_{aw}^2 = 0.5$, $\theta_{bw}^2 = 2$ and $\theta_{bb}^2 = 1$. First, we fix $P_b = 3$, increase ϕ from 0 to 1 and compute the lower bound in (17). The result is shown in Fig. 3. We can see that in this case (the jammer-Willie channel is better than the jammer-Bob channel), AN always helps boost covert throughput. However, the covert throughput with perfect self-interference cancellation is over 8 times larger than that with no self-interference cancellation.

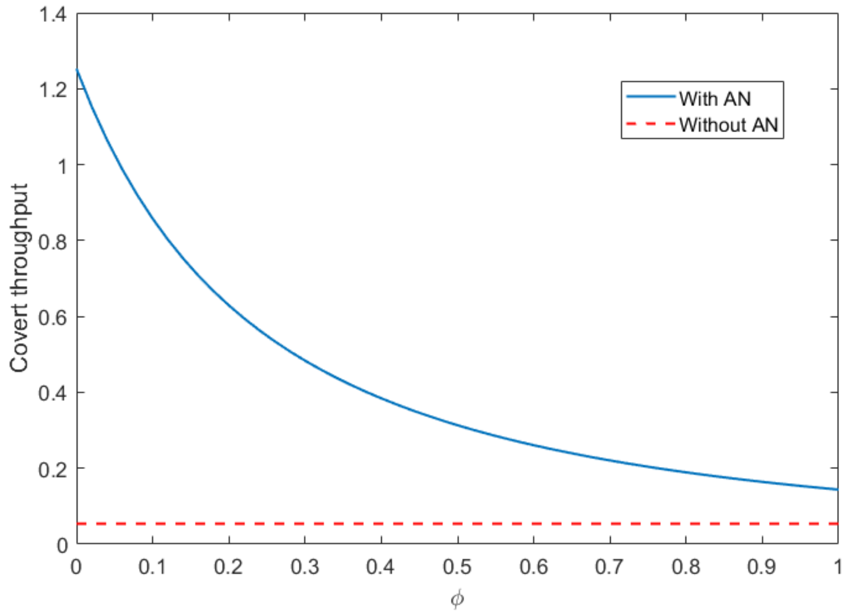

Fig. 3: Covert throughput vs. self-interference cancellation coefficient ϕ.

Next, we fix $\phi = 0.5$ and increase P_b from 0 to 20. It can be seen that the covert throughput gain converges as P_b goes large. This is due to the fact that

$$\lim_{P_b \to \infty} \frac{\theta_{bw}^2 P_b + 1}{\phi \theta_{bb}^2 P_b + 1} = \frac{\theta_{bw}^2}{\phi \theta_{bb}^2}.$$

To conclude, in the non-coherent i.i.d. Rayleigh fading channel, even fixed-amplitude AN can sometimes improve covert capacity, but increasing the AN power may not always has a positive effect, depending on the channel condition and Bob's self-interference cancellation ability.

B. Varying AN Power: Achieving Positive Covert Rates

Now we consider a varying AN power scenario. Assume that P_b is uniformly distributed over $[0, P_{max}]$. In this case, Willie has some uncertainty about its received signal noise variance under both hypotheses, and we will show that a positive covert rate is possible.

1) Willie's Hypothesis Testing: Assume that P_b is constant during a transmission block of length N, but changes to another value randomly in the next block. Then the signal received by

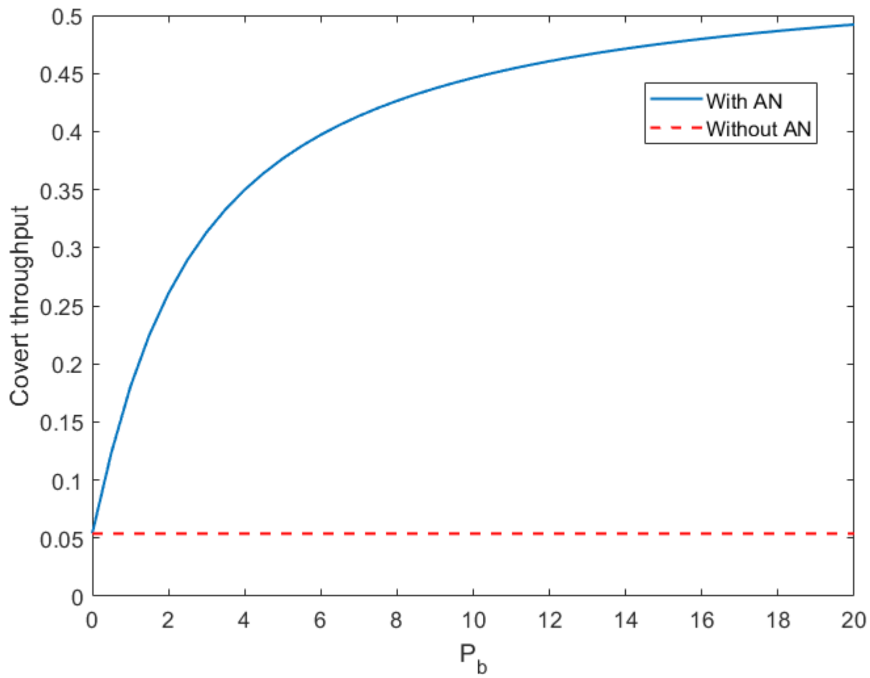

Fig. 4: Covert throughput vs. AN power P_b.

Willie during a block is given by:

$$\mathbf{z}[i] = \begin{cases} \mathbf{h}_{bw}[i]\mathbf{v}[i] + \mathbf{n}_{aw}[i] & \mathcal{H}_0, \\ \mathbf{h}_{aw}[i]\mathbf{x}[i] + \mathbf{h}_{bw}[i]\mathbf{v}[i] + \mathbf{n}_{aw}[i] & \mathcal{H}_1, \end{cases} \tag{24}$$

where $i \in [N]$. Let $\mathbf{v}[i] = \sqrt{P_b} e^{\rho j}$ for some $\rho \in [0, 2\pi]$. Then Willie's received signal $\mathbf{z}[i]$ is complex Gaussian distributed. When Alice is active, Willie observes both Alice's signal and Bob's jamming signal in addition to background noise. In the absence of Alice's transmission, $\mathbf{z}[i]$ is just the combination of the jamming signal and background noise. Let μ denote the variance of $\mathbf{z}[i]$. The distribution of μ under hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ can be written as

$$f_{\mu|\mathcal{H}_h}(\mu|\mathcal{H}_h) = \begin{cases} \frac{1}{P_{max}\theta_{bw}^2}, \quad \sigma_w^2 \leq \mu \leq P_{max}\theta_{bw}^2 + \sigma_w^2, & h = 0 \\ \frac{1}{P_{max}\theta_{bw}^2}, \quad P_a\theta_{aw}^2 + \sigma_w^2 \leq \mu \leq P_{max}\theta_{bw}^2 + P_a\theta_{aw}^2 + \sigma_w^2, & h = 1 \\ 0, \quad \text{otherwise} \end{cases} \tag{25}$$

We assume that Willie has full knowledge of the statistical model: Alice's signal power P_a, the distribution of AN power P_b, the variances of h_{aw} and h_{bw}, and the noise variance σ_w^2. Then by the Neyman-Pearson criterion, the optimal test for Willie is to use the likelihood ratio test (LRT) [26, Ch. 3.3]. Define the test statistic as

$$T(\mathbf{z}) = \sum_{i=1}^{N} |\mathbf{z}[i]|^2. \tag{26}$$

Let $\mathcal{D}_0$ and $\mathcal{D}_1$ denote the decisions that Alice is inactive and active, respectively, and define

$$\Lambda(\mathbf{z}) = \frac{f_{T(\mathbf{z})|\mathcal{H}_1}(T(\mathbf{z})|\mathcal{H}_1)}{f_{T(\mathbf{z})|\mathcal{H}_0}(T(\mathbf{z})|\mathcal{H}_0)}. \tag{27}$$

The decision rule of Willie is given by

$$\begin{cases} \mathcal{D}_0: & \text{if } \Lambda(\mathbf{z}) < \gamma, \\ \mathcal{D}_1: & \text{if } \Lambda(\mathbf{z}) > \gamma, \end{cases} \tag{28}$$

where $\gamma = \frac{P(\mathcal{H}_0)}{P(\mathcal{H}_1)}$, and $f_{T(\mathbf{z})|\mathcal{H}_1}(\cdot|\mathcal{H}_1)$ and $f_{T(\mathbf{z})|\mathcal{H}_0}(\cdot|\mathcal{H}_0)$ are the PDFs of $T(\mathbf{z})$ when Alice is active and inactive, respectively. In this paper, we assume that Willie has no knowledge about Alice's active probability, which indicates $\gamma = 1$. Therefore, Willie can alternatively find a proper detecting threshold τ such $f_{T(\mathbf{z})|\mathcal{H}_1}(\tau|\mathcal{H}_1) = f_{T(\mathbf{z})|\mathcal{H}_0}(\tau|\mathcal{H}_0)$ and make the following decision:

$$\begin{cases} \mathcal{D}_0: & \text{if } \frac{T(\mathbf{z})}{N} < \tau, \\ \mathcal{D}_1: & \text{if } \frac{T(\mathbf{z})}{N} > \tau. \end{cases} \tag{29}$$

The FAP and the MDP are given by

$$P_{fa} = P(\mathcal{D}_1|\mathcal{H}_0), \tag{30}$$

$$P_{md} = P(\mathcal{D}_0|\mathcal{H}_1), \tag{31}$$

respectively. To ensure covertness, Alice need to design a scheme such that for any covertness requirement $\epsilon > 0$, $P_{fa} + P_{md} > 1 - \epsilon$ for sufficiently large N.

2) Alice's Transmit Power: Let χ_l^2 denote a chi-squared random variable with l degrees of freedom. Then $T(\mathbf{z}) \sim \mu\chi_{2N}^2$. By the weak law of large numbers, χ_{2N}^2/N converges to 1 as N goes large. Therefore, under $\mathcal{H}_0$ and given $P_b = \xi P_{max}$ for some $\xi \in [0, 1]$, for any $\delta > 0$ there exists an N_0 such that for any $N > N_0$,

$$P\bigg(\frac{T(\mathbf{z})}{N} \in \Big((\xi P_{max}\theta_{bw}^2 + \sigma_w^2)(1 - \frac{\delta}{P_{max}\theta_{bw}^2 + \sigma_w^2}), (\xi P_{max}\theta_{bw}^2 + \sigma_w^2)(1 + \frac{\delta}{P_{max}\theta_{bw}^2 + \sigma_w^2})\Big)\bigg) > 1 - \frac{\epsilon}{2}. \tag{32}$$

Since $\xi \in [0, 1]$, we have $\xi P_{max}\theta_{bw}^2 + \sigma_w^2 \leq P_{max}\theta_{bw}^2 + \sigma_w^2$. Therefore,

$$P\bigg(\frac{T(\mathbf{z})}{N} \in \Big(\xi P_{max}\theta_{bw}^2 + \sigma_w^2 - \delta, \xi P_{max}\theta_{bw}^2 + \sigma_w^2 + \delta\Big)\bigg) > 1 - \frac{\epsilon}{2}. \tag{33}$$

Then we know that for any $\tau < \xi P_{max}\theta_{bw}^2 + \sigma_w^2 - \delta$

$$P_{fa}(\xi) > 1 - \frac{\epsilon}{2}.$$

Similarly we can obtain that for any $\tau > \xi P_{max}\theta_{bw}^2 + P_a\theta_{aw}^2 + \sigma_w^2 + \delta$

$$P_{md}(\xi) > 1 - \frac{\epsilon}{2}.$$

Define the following set

$$\mathcal{B} = \{\xi : \ \xi P_{max}\theta_{bw}^2 + \sigma_w^2 - \delta < \tau < P_{max}\theta_{bw}^2 + P_a\theta_{aw}^2 + \sigma_w^2 + \delta\}. \tag{34}$$

The probability that $\xi \in \mathcal{B}$ can be bounded by

$$\begin{aligned} P(\mathcal{B}) &= P(\frac{\tau - P_a\theta_{aw}^2 - \sigma_w^2 - \delta}{P_{max}\theta_{bw}^2} \leq \xi \leq \frac{\tau - \sigma_w^2 + \delta}{P_{max}\theta_{bw}^2}) \\ &\leq \frac{P_a\theta_{aw}^2 + 2\delta}{P_{max}\theta_{bw}^2} \end{aligned} \tag{35}$$

By choosing

$$\delta = P_{max}\theta_{bw}^2\epsilon/8,$$
$$P_a = P_{max}\theta_{bw}^2\epsilon/4\theta_{aw}^2,$$

we can guarantee that

$$P(\mathcal{B}^c) \geq 1 - \frac{\epsilon}{2}. \tag{36}$$

Therefore, the sum of Willie's missed detection rate and false alarm rate can be lower bounded by

$$\begin{aligned} P_{fa} + P_{md} &= \mathbb{E}_\xi\big[P_{fa}(\xi) + P_{md}(\xi)\big] \\ &\geq \mathbb{E}_\xi\big[P_{fa}(\xi) + P_{md}(\xi)|\mathcal{B}^c\big]P(\mathcal{B}^c) \\ &\geq 1 - \epsilon. \end{aligned} \tag{37}$$

This shows that at a given covertness requirement ϵ, Alice can transmit with non-vanishing power $P_{max}\theta_{bw}^2\epsilon/4\theta_{aw}^2$. The average signal-to-interference-plus-noise ratio (SINR) at Bob when Alice is transmitting is

$$\begin{aligned} \mathrm{SINR}_b &= \frac{P_{max}\theta_{bw}^2\theta_{ab}^2\epsilon}{4\theta_{aw}^2(\phi P_b\theta_{bb}^2 + \sigma_b^2)} \\ &\geq \frac{\theta_{bw}^2\theta_{ab}^2\epsilon}{4\theta_{aw}^2(\phi\theta_{bb}^2 + \sigma_b^2/P_{max})}. \end{aligned}$$

Now we discuss two cases. If self-interference can be perfectly cancelled by Bob, i.e., $\phi = 0$, then we have $\mathrm{SINR}_b \geq \frac{P_{max}\theta_{bw}^2\theta_{ab}^2\epsilon}{4\theta_{aw}^2\sigma_b^2}$. Then for any $\epsilon > 0$, we can always guarantee a positive SINR by choosing P_{max} sufficiently large, and a positive covert communication is achievable. However, if $\phi \neq 0$, then SINR_b goes to 0 as $\epsilon \to 0$, which means the covert communication rate is still asymptotically 0.

IV. Slow Fading Channel

In the previous section, we have considered covert communications over non-coherent i.i.d. Rayleigh fading channels with AN. Now we turn to the slow-fading case. We still assume that only the CDI is available to all parties, but the CSI remains constant during a whole transmission block of length N.

A. Willie's Hypothesis Testing

We assume that Willie uses a radiometer to detect the activity of Alice. The signal received by Willie during a block of length N is given by

$$\mathbf{z}[i] = \begin{cases} h_{bw}\mathbf{v}[i] + \mathbf{n}_{aw}[i] & \mathcal{H}_0, \\ h_{aw}\mathbf{x}[i] + h_{bw}\mathbf{v}[i] + \mathbf{n}_{aw}[i] & \mathcal{H}_1, \end{cases} \tag{38}$$

where $i \in [N]$. Define the test statistic as

$$T(\mathbf{z}) = \frac{1}{N}\sum_{i=1}^{N} |\mathbf{z}[i]|^2. \tag{39}$$

We assume that Willie has full knowledge of the statistical model: the distribution of AN power P_b, θ_{aw}^2 and θ_{bw}^2 (the variances of h_{aw} and h_{bw}), and the noise variance σ_w^2. Then by the Neyman-Pearson criterion, the optimal test for Willie is to use the likelihood ratio test (LRT) [26, Ch. 3.3]. Let $\mathcal{D}_0$ and $\mathcal{D}_1$ denote the decisions that Alice is inactive and active, respectively, and define

$$\Lambda(\mathbf{z}) = \frac{f_{T(\mathbf{z})|\mathcal{H}_1}(T(\mathbf{z})|\mathcal{H}_1)}{f_{T(\mathbf{z})|\mathcal{H}_0}(T(\mathbf{z})|\mathcal{H}_0)}. \tag{40}$$

The decision rule of Willie can be similarly given by (29), and the FAP and MDP can also be given by (30) and (31), respectively.

B. Fixed AN Power

1) Optimal Detecting Threshold: We first consider the fixed AN power case. When P_b is fixed and known to all parties and $\mathbf{v}[i] = \sqrt{P_b}e^{\rho j}$ for some $\rho \in [0, 2\pi]$, Willie's received signal $\mathbf{z}[i]$ is complex Gaussian distributed with zero mean and variance $P_b|h_{bw}|^2 + \sigma_w^2$ in $\mathcal{H}_0$ and $P_a|h_{aw}|^2 + P_b|h_{bw}|^2 + \sigma_w^2$ in $\mathcal{H}_1$, respectively. To determine the optimal detecting threshold for Willie, we have the following lemma.

Lemma 1. *The FAP and MDP of Willie are given by*

$$P_{fa} = \begin{cases} e^{-\frac{\tau-\sigma_w^2}{P_b\theta_{bw}^2}}, & \tau \geq \sigma_w^2 \\ 1, & \textit{otherwise} \end{cases} \tag{41}$$

and

$$P_{md} = \begin{cases} 1 - e^{-\frac{\tau-\sigma_w^2}{P_b\theta_{bw}^2}} - \frac{P_a\theta_{aw}^2}{P_a\theta_{aw}^2 - P_b\theta_{bw}^2}\left(e^{-\frac{\tau-\sigma_w^2}{P_a\theta_{aw}^2}} - e^{-\frac{\tau-\sigma_w^2}{P_b\theta_{bw}^2}}\right), & \tau \geq \sigma_w^2 \\ 0, & \textit{otherwise} \end{cases}, \tag{42}$$

respectively.

Proof. For $\tau < \sigma_w^2$, it is clear that $P_{fa} = 1$ and $P_{md} = 0$. Since $|h_{bw}|^2$ is exponentially distributed with density

$$\frac{1}{\theta_{bw}^2}\exp\left\{\frac{-x}{\theta_{bw}^2}\right\}, \qquad x > 0, \tag{43}$$

For $\tau \geq \sigma_w^2$, we have

$$\begin{aligned} P_{fa} &= \mathbb{P}\Big(P_b|h_{bw}|^2 + \sigma_w^2 > \tau|\mathcal{H}_0\Big) \\ &= \int_{(\tau-\sigma_w^2)/P_b}^{+\infty} \frac{1}{\theta_{bw}^2} e^{-\frac{1}{\theta_{bw}^2}x}\mathrm{d}x \\ &= e^{-\frac{\tau-\sigma_w^2}{P_b\theta_{bw}^2}}, \end{aligned} \tag{44}$$

and

$$\begin{aligned}
P_{md} &= \mathbb{P}\Big(P_a|h_{aw}|^2 + P_b|h_{bw}|^2 + \sigma_w^2 < \tau|\mathcal{H}_1)\Big) \qquad (45)\\
&= \int_0^{(\tau-\sigma_w^2)/P_b} \frac{1}{\theta_{bw}^2} e^{-\frac{1}{\theta_{bw}^2}y} \int_0^{(\tau-\sigma_w^2-P_b y)/P_a} \frac{1}{\theta_{aw}^2} e^{-\frac{1}{\theta_{aw}^2}x} \mathrm{d}x\mathrm{d}y\\
&= \int_0^{(\tau-\sigma_w^2)/P_b} \frac{1}{\theta_{bw}^2} e^{-\frac{1}{\theta_{bw}^2}y} \Big(1 - e^{-\frac{\tau-\sigma_w^2-P_b y}{P_a\theta_{aw}^2}}\Big)\mathrm{d}y\\
&= 1 - e^{-\frac{\tau-\sigma_w^2}{P_b\theta_{bw}^2}} - \int_0^{(\tau-\sigma_w^2)/P_b} \frac{1}{\theta_{bw}^2} e^{-\frac{\tau-\sigma_w^2}{P_a\theta_{aw}^2}} e^{-\Big(\frac{1}{\theta_{bw}^2}-\frac{P_b}{P_a\theta_{aw}^2}\Big)y}\mathrm{d}y\\
&= 1 - e^{-\frac{\tau-\sigma_w^2}{P_b\theta_{bw}^2}} - \frac{P_a\theta_{aw}^2}{P_a\theta_{aw}^2 - P_b\theta_{bw}^2} e^{-\frac{\tau-\sigma_w^2}{P_a\theta_{aw}^2}} \Big(1 - e^{-\frac{(P_a\theta_{aw}^2-P_b\theta_{bw}^2)(\tau-\sigma_w^2)}{P_a\theta_{aw}^2 P_b\theta_{bw}^2}}\Big)\\
&= 1 - e^{-\frac{\tau-\sigma_w^2}{P_b\theta_{bw}^2}} - \frac{P_a\theta_{aw}^2}{P_a\theta_{aw}^2 - P_b\theta_{bw}^2} \Big(e^{-\frac{\tau-\sigma_w^2}{P_a\theta_{aw}^2}} - e^{-\frac{\tau-\sigma_w^2}{P_b\theta_{bw}^2}}\Big). \qquad (46)
\end{aligned}$$

□

Theorem 2. *The optimal detecting threshold of Willie is given by*

$$\tau = \begin{cases} \sigma_w^2 + P_a P_b \theta_{aw}^2 \theta_{bw}^2 \frac{\log(P_a\theta_{aw}^2) - \log(P_b\theta_{bw}^2)}{P_a\theta_{aw}^2 - P_b\theta_{bw}^2}, & P_a\theta_{aw}^2 \neq P_b\theta_{bw}^2 \\ \sigma_w^2 + P_a\theta_{aw}^2, & P_a\theta_{aw}^2 = P_b\theta_{bw}^2 \end{cases}. \qquad (47)$$

Proof. Willie's detecting failure probability is given by

$$\begin{aligned}
P_f(\tau) &= P_{fa} + P_{md}\\
&= 1 - \frac{P_a\theta_{aw}^2}{P_a\theta_{aw}^2 - P_b\theta_{bw}^2} \Big(e^{-\frac{\tau-\sigma_w^2}{P_a\theta_{aw}^2}} - e^{-\frac{\tau-\sigma_w^2}{P_b\theta_{bw}^2}}\Big). \qquad (48)
\end{aligned}$$

Then we have

$$\frac{\mathrm{d}P_f(\tau)}{\mathrm{d}\tau} = \frac{P_a\theta_{aw}^2}{P_a\theta_{aw}^2 - P_b\theta_{bw}^2} \Big(\frac{1}{P_a\theta_{aw}^2} e^{-\frac{\tau-\sigma_w^2}{P_a\theta_{aw}^2}} - \frac{1}{P_b\theta_{bw}^2} e^{-\frac{\tau-\sigma_w^2}{P_b\theta_{bw}^2}}.\Big)$$

Letting $\frac{\mathrm{d}P_f(\tau)}{\mathrm{d}\tau} = 0$, we have

$$\tau = \sigma_w^2 + P_a P_b \theta_{aw}^2 \theta_{bw}^2 \frac{\log(P_a\theta_{aw}^2) - \log(P_b\theta_{bw}^2)}{P_a\theta_{aw}^2 - P_b\theta_{bw}^2}.$$

It is easy to verify that such a τ minimizes $P_f(\tau)$. Thus, (47) is the optimal detecting threshold for Willie. □

2) Covert Rate: From Theorem 2 and (48) we can obtain that the minimal detecting error probability of Willie is

$$\min P_f(\tau) = 1 - \frac{P_a\theta_{aw}^2}{P_a\theta_{aw}^2 - P_b\theta_{bw}^2}\left[\left(\frac{P_a\theta_{aw}^2}{P_b\theta_{bw}^2}\right)^{\frac{P_b\theta_{bw}^2}{P_b\theta_{bw}^2 - P_a\theta_{aw}^2}} - \left(\frac{P_a\theta_{aw}^2}{P_b\theta_{bw}^2}\right)^{\frac{P_a\theta_{aw}^2}{P_b\theta_{bw}^2 - P_a\theta_{aw}^2}}\right]. \tag{49}$$

Therefore, for fixed transmit power P_a, to guarantee that Willie has a detecting failure probability at least $1-\epsilon$, we need to determine the minimum P_b such that

$$\frac{P_a\theta_{aw}^2}{P_a\theta_{aw}^2 - P_b\theta_{bw}^2}\left[\left(\frac{P_a\theta_{aw}^2}{P_b\theta_{bw}^2}\right)^{\frac{P_b\theta_{bw}^2}{P_b\theta_{bw}^2 - P_a\theta_{aw}^2}} - \left(\frac{P_a\theta_{aw}^2}{P_b\theta_{bw}^2}\right)^{\frac{P_a\theta_{aw}^2}{P_b\theta_{bw}^2 - P_a\theta_{aw}^2}}\right] \leq \epsilon. \tag{50}$$

Let $\eta = \frac{P_b\theta_{bw}^2}{P_a\theta_{aw}^2}$ denote the relative AN power with respect to the transmit power. Then the above inequality can be rewritten as

$$f(\eta) = \frac{1}{1-\eta}\left(\eta^{\frac{\eta}{1-\eta}} - \eta^{\frac{1}{1-\eta}}\right) \leq \epsilon. \tag{51}$$

It can be shown that $f(\eta)$ is strictly decreasing as η increases and $\lim_{\eta\to\infty} f(\eta) = 0$ (see Fig. 5). Thus, for any $0 < \epsilon < 1$ and fixed transmit power P_a, without any constraint on the maximum AN power, it is always possible to achieve covert communication with any covertness requirement of ϵ. Thus, a positive outage[2] covert communication rate is achievable.

C. Varying AN Power

Now we assume that P_b is uniformly distributed over $[0, P_{max}]$ in different transmission blocks (but remains constant during a block).

1) Optimal Detecting Threshold:

Lemma 2. *The FAP and MDP of Willie in the varying AN power case are given by*

$$P_{fa} = \begin{cases} e^{\frac{\tau-\sigma_w^2}{P_{max}\theta_{bw}^2}} - \frac{\tau-\sigma_w^2}{P_{max}\theta_{bw}^2}\mathrm{Ei}\left(\frac{\tau-\sigma_w^2}{P_{max}\theta_{bw}^2}\right), & \tau \geq \sigma_w^2 \\ 1, & \textit{otherwise} \end{cases} \tag{52}$$

and

$$P_{md} = \begin{cases} P_1, & \tau \geq \sigma_w^2 \\ 0, & \textit{otherwise} \end{cases}, \tag{53}$$

respectively, where $\mathrm{Ei}(x) = \int_{-\infty}^{x} \frac{e^y}{y}\mathrm{d}y$ *is the exponential integral function.*

[2]Recall that for slow fading channels without CSI, it is more convenient to define channel capacity with an outage probability [27, Chapter 5.4.1].

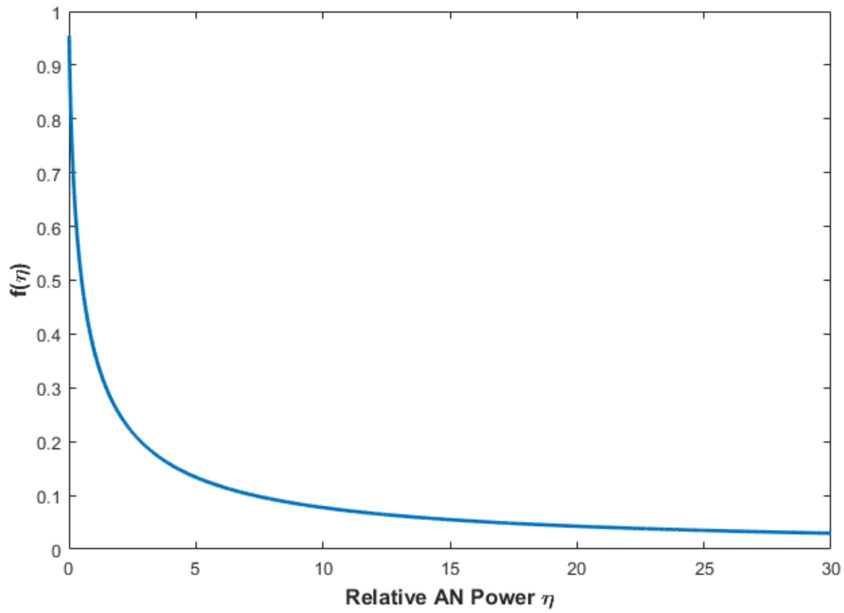

Fig. 5: $f(\eta)$ vs. relative AN power η.

Proof. For $\tau < \sigma_w^2$, it is clear that $P_{fa} = 1$ and $P_{md} = 0$. For $\tau \geq \sigma_w^2$, we have

$$\begin{aligned} P_{fa} &= \mathbb{P}\Big(P_b|h_{bw}|^2 + \sigma_w^2 > \tau|\mathcal{H}_0)\Big) \\ &= \int_0^{P_{max}} \frac{1}{P_{max}} \int_{(\tau-\sigma_w^2)/y}^{+\infty} \frac{1}{\theta_{bw}^2} e^{-\frac{1}{\theta_{bw}^2}x} \mathrm{d}x\mathrm{d}y \\ &= \int_0^{P_{max}} \frac{1}{P_{max}} e^{-\frac{(\tau-\sigma_w^2)}{\theta_{bw}^2 y}} \mathrm{d}y \\ &= e^{-\frac{\tau-\sigma_w^2}{P_{max}\theta_{bw}^2}} + \frac{\tau-\sigma_w^2}{P_{max}\theta_{bw}^2} \mathrm{Ei}\big(-\frac{\tau-\sigma_w^2}{P_{max}\theta_{bw}^2}\big), \end{aligned} \tag{54}$$

and

$$\begin{aligned} P_{md} &= \mathbb{P}\Big(P_a|h_{aw}|^2 + P_b|h_{bw}|^2 + \sigma_w^2 < \tau|\mathcal{H}_1)\Big) \qquad (55)\\ &= \int_0^{P_{max}} \frac{1}{P_{max}} \int_0^{(\tau-\sigma_w^2)/z} \frac{1}{\theta_{bw}^2} e^{-\frac{1}{\theta_{bw}^2}y} \int_0^{(\tau-\sigma_w^2-zy)/P_a} \frac{1}{\theta_{aw}^2} e^{-\frac{1}{\theta_{aw}^2}x} \mathrm{d}x\mathrm{d}y\mathrm{d}z \\ &= \int_0^{P_{max}} \frac{1}{P_{max}} \Big[1 - e^{-\frac{\tau-\sigma_w^2}{\theta_{bw}^2 z}} - \frac{P_a\theta_{aw}^2}{P_a\theta_{aw}^2 - \theta_{bw}^2 z}\Big(e^{-\frac{\tau-\sigma_w^2}{P_a\theta_{aw}^2}} - e^{-\frac{\tau-\sigma_w^2}{\theta_{bw}^2 z}}\Big)\Big]\mathrm{d}z \\ &= 1 - e^{-\frac{\tau-\sigma_w^2}{P_{max}\theta_{bw}^2}} - \frac{\tau-\sigma_w^2}{P_{max}\theta_{bw}^2}\mathrm{Ei}\big(-\frac{\tau-\sigma_w^2}{P_{max}\theta_{bw}^2}\big) + \frac{P_a\theta_{aw}^2}{P_{max}\theta_{bw}^2}\Big[e^{-\frac{\tau-\sigma_w^2}{P_a\theta_{aw}^2}} \ln\big|1 - \frac{P_{max}\theta_{bw}^2}{P_a\theta_{aw}^2}\big| \\ &\quad + \mathrm{Ei}\big(-\frac{\tau-\sigma_w^2}{P_{max}\theta_{bw}^2}\big) - e^{-\frac{\tau-\sigma_w^2}{P_a\theta_{aw}^2}}\mathrm{Ei}\big(\frac{\tau-\sigma_w^2}{P_a\theta_{aw}^2} - \frac{\tau-\sigma_w^2}{P_{max}\theta_{bw}^2}\big)\Big] \qquad (56) \end{aligned}$$

□

Theorem 3. *The optimal detecting threshold of Willie in the varying AN power case is given by*

$$\tau = \sigma_w^2 + P_a P_b \theta_{aw}^2 \theta_{bw}^2 \frac{\mathrm{Ei}^{-1}\big(\ln\big|1 - \frac{P_{max}\theta_{bw}^2}{P_a\theta_{aw}^2}\big|\big)}{P_{max}\theta_{bw}^2 - P_a\theta_{aw}^2}, \qquad (57)$$

where $y = \mathrm{Ei}^{-1}(\cdot)$ *is the inverse function of* $\mathrm{Ei}(\cdot)$ *with* y *being restricted in* $(0, \infty)$.

Proof. Willie's detecting failure probability is given by

$$\begin{aligned} P_f(\tau) &= P_{fa} + P_{md} \\ &= 1 + \frac{P_a\theta_{aw}^2}{P_{max}\theta_{bw}^2}\Big[e^{-\frac{\tau-\sigma_w^2}{P_a\theta_{aw}^2}} \ln\big|1 - \frac{P_{max}\theta_{bw}^2}{P_a\theta_{aw}^2}\big| + \mathrm{Ei}\big(-\frac{\tau-\sigma_w^2}{P_{max}\theta_{bw}^2}\big) \\ &\quad - e^{-\frac{\tau-\sigma_w^2}{P_a\theta_{aw}^2}}\mathrm{Ei}\big(\frac{\tau-\sigma_w^2}{P_a\theta_{aw}^2} - \frac{\tau-\sigma_w^2}{P_{max}\theta_{bw}^2}\big)\Big]. \qquad (58) \end{aligned}$$

Then we have

$$\frac{\mathrm{d}P_f(\tau)}{\mathrm{d}\tau} = \frac{e^{-\frac{\tau-\sigma_w^2}{P_a\theta_{aw}^2}}}{P_{max}\theta_{bw}^2}\Big[\mathrm{Ei}\big(\frac{\tau-\sigma_w^2}{P_a\theta_{aw}^2} - \frac{\tau-\sigma_w^2}{P_{max}\theta_{bw}^2}\big) - \ln\big|1 - \frac{P_{max}\theta_{bw}^2}{P_a\theta_{aw}^2}\big|\Big]$$

Letting $\frac{\mathrm{d}P_f(\tau)}{\mathrm{d}\tau} = 0$, we have

$$\tau = \sigma_w^2 + P_a P_{max} \theta_{aw}^2 \theta_{bw}^2 \frac{\mathrm{Ei}^{-1}\big(\ln\big|1 - \frac{P_{max}\theta_{bw}^2}{P_a\theta_{aw}^2}\big|\big)}{P_{max}\theta_{bw}^2 - P_a\theta_{aw}^2}.$$

It is easy to verify that such a τ minimizes $P_f(\tau)$. Thus, (57) is the optimal detecting threshold for Willie. □

2) Covert Rate: From Theorem 3 and (58) we have that the minimal detecting error probability of Willie is

$$\min P_f(\tau) = 1 + \frac{P_a\theta_{aw}^2}{P_{max}\theta_{bw}^2}\mathrm{Ei}\Big(- \frac{P_a\theta_{aw}^2 \mathrm{Ei}^{-1}\big(\ln\big|1 - \frac{P_{max}\theta_{bw}^2}{P_a\theta_{aw}^2}\big|\big)}{P_{max}\theta_{bw}^2 - P_a\theta_{aw}^2}\Big). \tag{59}$$

Therefore, for fixed transmit power P_a, to guarantee that Willie has a detecting failure probability at least $1-\epsilon$, we need to determine the minimum P_b such that

$$\frac{P_a\theta_{aw}^2}{P_{max}\theta_{bw}^2}\mathrm{Ei}\Big(- \frac{P_a\theta_{aw}^2 \mathrm{Ei}^{-1}\big(\ln\big|1 - \frac{P_{max}\theta_{bw}^2}{P_a\theta_{aw}^2}\big|\big)}{P_{max}\theta_{bw}^2 - P_a\theta_{aw}^2}\Big) \geq -\epsilon. \tag{60}$$

Let $\eta = \frac{P_{max}\theta_{bw}^2}{P_a\theta_{aw}^2}$. Then the above inequality can be rewritten as

$$f(\eta) = \frac{1}{\eta}\mathrm{Ei}\Big(- \frac{\mathrm{Ei}^{-1}\big(\ln\big|1-\eta\big|\big)}{\eta - 1}\Big) \geq -\epsilon. \tag{61}$$

We can verify that for $\eta > 7.77$, $\mathrm{Ei}^{-1}\big(\ln\big|1-\eta\big|\big) > 1$. Since $\mathrm{Ei}(x)$ is negative and strictly decreasing in $x \in (-\infty, 0)$, we then have

$$\frac{1}{\eta}\mathrm{Ei}\Big(- \frac{1}{\eta-1}\Big) < f(\eta) < 0.$$

It can be shown that $\lim_{\eta\to\infty} \frac{1}{\eta}\mathrm{Ei}\big(- \frac{1}{\eta-1}\big) = 0$. Thus, for any $0 < \epsilon < 1$ and fixed transmit power P_a, without any constraint on the maximum AN power, it is always possible to achieve covert communication with a given covertness requirement ϵ. Thus, a positive outage covert communication rate is achievable.

V. Conclusion and Future Work

In this work, we have investigated covert communications over non-coherent i.i.d. and slow Rayleigh fading channels with a full-duplex receiver that generates AN while receiving signals. We show that with the help of the full-duplex receiver, covert communication can be improved in the non-coherent fading environment, especially when the receiver has good ability of self-interference cancellation, and even positive covert communication rates are achievable.

Note that in our analysis, although AN has the simple form of fixed amplitude and phase during a whole transmission block, covert communication is already greatly benefited. This shows that while non-coherent fading is bad news for conventional communications, it can offer much opportunity for covert communications. In the future we may investigate whether more complicated AN forms can further enhance covert communication.

References

[1] B. A. Bash, D. Goeckel, and D. Towsley, "Limits of reliable communication with low probability of detection on awgn channels," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 1921–1930, 2013.

[2] L. Wang, G. W. Wornell, and L. Zheng, "Fundamental limits of communication with low probability of detection," *IEEE Transactions on Information Theory*, vol. 62, no. 6, pp. 3493–3503, 2016.

[3] M. R. Bloch, "Covert communication over noisy channels: A resolvability perspective," *IEEE Transactions on Information Theory*, vol. 62, no. 5, pp. 2334–2354, 2016.

[4] M. Tahmasbi and M. R. Bloch, "First- and second-order asymptotics in covert communication," *IEEE Transactions on Information Theory*, vol. 65, no. 4, pp. 2190–2212, 2019.

[5] S. Yan, Y. Cong, S. V. Hanly, and X. Zhou, "Gaussian signalling for covert communications," *IEEE Transactions on Wireless Communications*, vol. 18, no. 7, pp. 3542–3553, 2019.

[6] M. T. Ishaque Ashar Kadampot and M. R. Bloch, "Codes for covert communication over additive white Gaussian noise channels," in *2019 IEEE International Symposium on Information Theory (ISIT)*, 2019, pp. 977–981.

[7] M. R. Bloch and S. Guha, "Optimal covert communications using pulse-position modulation," in *2017 IEEE International Symposium on Information Theory (ISIT)*, 2017, pp. 2825–2829.

[8] K. S. Kumar Arumugam and M. R. Bloch, "Embedding covert information in broadcast communications," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 10, pp. 2787–2801, 2019.

[9] K. S. K. Arumugam and M. R. Bloch, "Covert communication over a k -user multiple-access channel," *IEEE Transactions on Information Theory*, vol. 65, no. 11, pp. 7020–7044, 2019.

[10] K. S. Kumar Arumugam, M. R. Bloch, and L. Wang, "Covert communication over a physically degraded relay channel with non-colluding wardens," in *2018 IEEE International Symposium on Information Theory (ISIT)*, 2018, pp. 766–770.

[11] S. Lee, R. J. Baxley, M. A. Weitnauer, and B. Walkenhorst, "Achieving undetectable communication," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 7, pp. 1195–1205, 2015.

[12] D. Goeckel, B. Bash, S. Guha, and D. Towsley, "Covert communications when the warden does not know the background noise power," *IEEE Communications Letters*, vol. 20, no. 2, pp. 236–239, 2016.

[13] B. He, S. Yan, X. Zhou, and V. K. N. Lau, "On covert communication with noise uncertainty," *IEEE Communications Letters*, vol. 21, no. 4, pp. 941–944, 2017.

[14] B. A. Bash, D. Goeckel, and D. Towsley, "LPD communication when the warden does not know when," in *2014 IEEE International Symposium on Information Theory*, 2014, pp. 606–610.

[15] K. S. K. Arumugam and M. R. Bloch, "Keyless asynchronous covert communication," in *2016 IEEE Information Theory Workshop (ITW)*, 2016, pp. 191–195.

[16] B. A. Bash, D. Goeckel, and D. Towsley, "Covert communication gains from adversary's ignorance of transmission time," *IEEE Transactions on Wireless Communications*, vol. 15, no. 12, pp. 8394–8405, 2016.

[17] T. V. Sobers, B. A. Bash, S. Guha, D. Towsley, and D. Goeckel, "Covert communication in the presence of an uninformed jammer," *IEEE Transactions on Wireless Communications*, vol. 16, no. 9, pp. 6193–6206, 2017.

[18] R. Soltani, D. Goeckel, D. Towsley, B. A. Bash, and S. Guha, "Covert wireless communication with artificial noise generation," *IEEE Transactions on Wireless Communications*, vol. 17, no. 11, pp. 7252–7267, 2018.

[19] M. Tahmasbi, A. Savard, and M. R. Bloch, "Covert capacity of non-coherent rayleigh-fading channels," *IEEE Transactions on Information Theory*, pp. 1–1, 2019.

[20] T. Xu, L. Xu, X. Liu, and Z. Lu, "Covert communication with a full-duplex receiver based on channel distribution information," in *2018 12th International Symposium on Antennas, Propagation and EM Theory (ISAPE)*, 2018, pp. 1–4.

[21] K. Shahzad, X. Zhou, S. Yan, J. Hu, F. Shu, and J. Li, "Achieving covert wireless communications using a full-duplex receiver," *IEEE Transactions on Wireless Communications*, vol. 17, no. 12, pp. 8517–8530, 2018.

[22] J. Hu, S. Yan, X. Zhou, F. Shu, and J. Li, "Covert wireless communications with channel inversion power control in rayleigh fading," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 12, pp. 12 135–12 149, 2019.

[23] F. Shu, T. Xu, J. Hu, and S. Yan, "Delay-constrained covert communications with a full-duplex receiver," *IEEE Wireless Communications Letters*, vol. 8, no. 3, pp. 813–816, 2019.

[24] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[25] I. Csiszar and J. Korner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Akademiai Kiado, December 1981.

[26] S. M. Kay, *Fundamentals of statistical signal processing*. Prentice Hall PTR, 1993.

[27] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.