Covert and Reliable Short-Packet Communications against A Proactive Warder

Manlin Wang[‡,†], Yao Yao[†], Bin Xia[†], Zhiyong Chen[†] and Jiangzhou Wang[‡]
[†]Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai, China
[‡]School of Engineering, University of Kent, Canterbury, U.K.
Email: [†]{wangmanlin, sandyyao, bxia, zhiyongchen}@sjtu.edu.cn; [‡]j.z.wang@kent.ac.uk

Abstract—Wireless short-packet communications pose challenges to the security and reliability of the transmission. Besides, the proactive warder compounds these challenges, who detects and interferes with the potential transmission. An extra jamming channel is introduced by the proactive warder compared with the passive one, resulting in the inapplicability of analytical methods and results in existing works. Thus, effective system design schemes are required for short-packet communications against the proactive warder. To address this issue, we consider the analysis and design of covert and reliable transmission for above systems. Specifically, to investigate the reliable and covert performance of the system, detection error probability at the warder and decoding error probability at the receiver are derived, which is affected by both the transmit power and the jamming power. Furthermore, to maximize the effective throughput, an optimization framework is proposed under reliability and covertness constraints. Numerical results verify the accuracy of analytical results and the feasibility of the optimization framework. It is shown that the tradeoff between covertness and reliability is changed by the proactive warder compared with the passive one. Besides, it is shown that longer block length is always beneficial to improve the throughput for systems with optimized transmission rates. But when transmission rates are fixed, the block length should be carefully designed since the maximum one is not optimal in this case.

Index Terms—covert and reliable transmission, short-packet communications, proactive warder, effective throughput

I. INTRODUCTION

Time-sensitive and mission-critical Internet of Things (IoT) applications have aroused great attention in the fifth-generation mobile communications systems [?]. The use of short packets meets the stringent low latency requirements, but a severe loss in coding gain exits with short packets, posing challenges to transmission reliability. Besides, massive confident messages are transmitted in wireless channels in IoT scenarios, which poses unprecedented challenges to transmission security. The exposure of transmission behaviors may bring unpredictable risks and losses in these scenarios. Notably, covert communication offers an solution for this issue, which prevents the transmission behaviors from being detected [?].

The fundamental work for covert communications [?] demonstrated that $\mathcal{O}(\sqrt{n})$ bits of information can be transmitted reliably and deniably over n channel use. In addition, [illegible] covert communication with short packets, [?] investigated the effective throughput of the system in additive white Gaussian noise (AWGN) channels. Similarly, [?] considered the achievability bounds [illegible] of given block length and error probability over AWGN channels. In addition, [?], [?] investigated the throughput over quasi-static fading channels, revealing the fundamental difference [illegible] between the case of quasi-static fading channel and that of AWGN channel. More complex scenarios with multiple warders, multi-antenna, [illegible] were considered in [?], [?]. [illegible]

This is because the proactive warder can not only detect the [illegible] transmission behaviours [?], [illegible] Different from the passive warder, the proactive warder behaves more dangerously. [illegible] to interfere with the potential transmission simultaneously [?]. However, [illegible] the proactive warder was considered [illegible] [?], [?], [?] [illegible] In addition, [?] investigated [illegible] power control [illegible] covert communication [?], [?], [?], [?] [illegible] the proactive warder, since [illegible] However, the analysis and the corresponding system design [illegible] were based on [illegible] infinite block length assumption, which is no longer suitable for short-packet transmission. Besides, the results [illegible] with passive warders [?], [?], [?], [?], [?] can not be directly applied to the system with the proactive warder [illegible]

To address this issue, in this paper, we consider the analysis and design of reliable and covert transmission against a proactive warder. [illegible] This compounds the challenges of both reliability and covertness in short-packet communications. [illegible] effective short-packet transmission design schemes to provide both reliability and covertness [illegible] which is tighter than the widely used Kullback–Leibler (KL) [illegible] To address this issue, in this paper, we consider the [illegible]

This work is supported in part by the National Natural Science Foundation of China under Grant 62271300, and Shanghai Municipal Science and Technology Major Project under grant 2021SHZDZX0102.

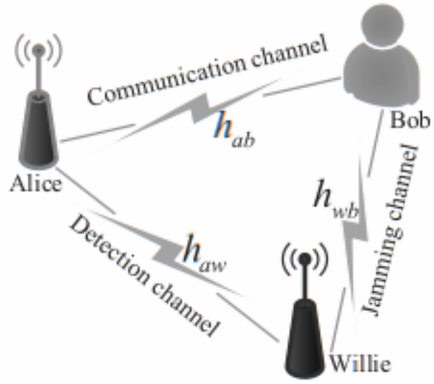

Fig. 1. Covert and reliable communication system against a proactive warder.

analysis and design of reliable and covert transmissions against a proactive warder. Specifically, to guarantee the system covertness requirement, the average detection error probability at the warder is derived. Besides, to facilitate system analysis and optimization, concise approximation expression is also proposed, which is tighter than the widely used Kullback-Leibler (KL) divergence approximation in existing works on covert communications. To guarantee the reliability requirement, the average decoding error probability at the receiver is derived. Furthermore, an optimization problem is formulated and an optimization framework is proposed to maximize the effective throughput of the system with reliability and covertness constraints by jointly designing the transmit power, transmission rate and blocklength. Numerical simulations verify the tightness of the proposed approximations and the feasibility of the proposed optimization framework for the system.

Notation: $|\cdot|$ denotes the absolute value operator. $\mathcal{CN}(0,\sigma^2)$ denotes the complex Gaussian distribution with zero mean and variance σ^2. $\Pr(\cdot)$ denotes the probability of an event. $Q(x)=\int_x^\infty \frac{1}{\sqrt{2\pi}}\exp(-t^2/2)\,dt$ denotes the Q-function. $\Gamma(n)=(n-1)!$ denotes the Gamma function, and $\gamma(n,x)=\int_0^x e^{-t}t^{n-1}\,dt$ denotes the lower incomplete Gamma function. $\psi(x)=\frac{d\ln(\Gamma(x))}{dx}$ denotes the digamma function while $\psi^{(n)}(x)$ denotes its n-th derivative. $E_1(x)=\int_x^\infty \frac{e^{-t}}{t}\,dt$ denotes the exponential integral function.

II. System Model

A. Signal and Channel Models

As shown in Fig. 1, a covert wireless communication scenario is considered, where the transmitter (Alice) desires to deliver messages to the receiver (Bob) while keeping a full-duplex warder (Willie) unaware of the transmission. Willie operates in full-duplex receiving signals from Alice and transmitting jamming signals to Bob simultaneously. Alice and Bob are assumed to be equipped with a single antenna, while Willie is assumed to be equipped with two antennas to support full-duplex functionality (detecting and jamming) [?].

In one transmission round, Alice transmits n covert signals $x_a[i], i\in\{1,\cdots,n\}$ to Bob, while Willie sends n jamming signals $x_w[i], i\in\{1,\cdots,n\}$ to Bob. Besides, Willie collects received signals to determine whether or not Alice has transmitted signals. The transmit power of Alice is denoted as P_a, $x_a[i]\sim\mathcal{CN}(0,P_a)$. Similarly, the jamming power of Willie is denoted as P_w, $x_w[i]\sim\mathcal{CN}(0,P_w)$. We denote the AWGN at Bob and Willie as $n_b[i]\sim\mathcal{CN}(0,\sigma_b^2)$ and $n_w[i]\sim\mathcal{CN}(0,\sigma_w^2)$, where σ_b^2 and σ_w^2 are the noise variances at Bob and Willie, respectively.

The wireless channels from Alice to Bob (communication channel, h_{ab}), Alice to Willie (detection channel, h_{aw}) and Willie to Bob (jamming channel, h_{wb}) are subject to the quasi-static Rayleigh fading [?]. Specifically, $h_{ab}\sim\mathcal{CN}(0,\lambda_{ab})$, $h_{aw}\sim\mathcal{CN}(0,\lambda_{aw})$ and $h_{wb}\sim\mathcal{CN}(0,\lambda_{wb})$. The channel coefficients remain constant during one transmission round, and are independently and identically distributed (i.i.d.) among different rounds. The instantaneous channel state information (CSI) h_{aw} is unavailable for Alice since Willie does not cooperate with Alice as an adversarial node while the statistical CSI is able to be estimated through the jamming signal [?]. Besides, the instantaneous CSI is available for Willie from a worst case perspective for covert communication.

B. Binary Hypothesis Testing at Willie

In order to detect the presence of covert communications, Willie must distinguish between the following two hypotheses in each transmission round

$$y_w[i]=\begin{cases}\sqrt{\varphi}x_w[i]+n_w[i], & \mathcal{H}_0\\ h_{aw}x_a[i]+\sqrt{\varphi}x_w[i]+n_w[i], & \mathcal{H}_1\end{cases} \tag{1}$$

where $\mathcal{H}_0$ denotes the null hypothesis where Alice has not transmitted, $\mathcal{H}_1$ denotes the alternative hypothesis where Alice has transmitted. $y_w[i]$ is the received signal at Willie, and $\varphi\in[0,1]$ is the self-interference cancellation coefficient [?], [?].

With a radiometer [?], Willie makes a binary decision as

$$T=\frac{1}{n}\sum_{i=1}^{n}|y_w[i]|^2 \underset{\mathcal{D}_0}{\overset{\mathcal{D}_1}{\gtrless}} \tau, \tag{2}$$

where T is the average power of each received signal at Willie, τ denotes the detection threshold, $\mathcal{D}_0$ and $\mathcal{D}_1$ denote the binary decisions that infer whether Alice transmits or not.

Suppose there is no prior knowledge for Willie about when Alice will transmit, the priori probability of either hypothesis is equal. Mathematically, the detection error probability ξ at Willie is defined as follows [?], [?], [?], [?]

$$\begin{aligned}\xi(\tau)&=\Pr(\mathcal{D}_1|\mathcal{H}_0)+\Pr(\mathcal{D}_0|\mathcal{H}_1)\\&=\Pr(T>\tau|\mathcal{H}_0)+\Pr(T<\tau|\mathcal{H}_1),\end{aligned} \tag{3}$$

where $\Pr(\mathcal{D}_1|\mathcal{H}_0)$ denotes the false alarm probability and $\Pr(\mathcal{D}_0|\mathcal{H}_1)$ denotes the missed detection probability. In covert communications, Willie's ultimate goal is to detect the presence of Alice's transmission with the minimum detection error probability ξ^*, which is achieved by using the optimal detection threshold τ^* that minimizes (3).

C. Effective Throughput with Finite Blocklength

When Alice transmits, the received signal at Bob can be expressed as

$$y_b[i]=h_{ab}x_a[i]+h_{wb}x_w[i]+n_b[i]. \tag{4}$$

Based on the received signal (??), Bob can decode the messages. The decoding error cannot be ignored in short-packet communications, which is given by [?]

$$\varepsilon\approx Q\left(\frac{\ln 2\sqrt{n}(\log_2(1+\gamma_b)-R)}{\sqrt{1-(1+\gamma_b)^{-2}}}\right), \tag{5}$$

where $\Pr(\mathcal{D}_1|\mathcal{H}_0)$ denotes the false alarm probability and $\Pr(\mathcal{D}_0|\mathcal{H}_1)$ denotes the miss detection probability. In covert communications, Willie's ultimate goal is to detect the presence of Alice's transmission with the minimum detection error probability ξ^*, which is achieved by using the optimal detection threshold τ^* that minimizes ξ.

C. Effective Throughput with Finite Blocklength

When Alice transmits, the received signal at Bob can be expressed as

$$y_b[i] = h_{ab}x_a[i] + h_{wb}x_w[i] + n_b[i]. \quad (4)$$

Based on the received signal (??), Bob can decode the messages. The decoding error cannot be ignored in short-packet communications, which is given by [?]

$$\delta = \mathcal{Q}\left(\frac{\ln 2\sqrt{n}\left(\log_2(1+\gamma_b) - R\right)}{\sqrt{1-(1+\gamma_b)^{-2}}}\right), \quad (5)$$

where $\gamma_b = P_a|h_{ab}|^2 / \left(P_w|h_{wb}|^2 + \sigma_b^2\right)$ denotes the received signal-to-noise ratio (SNR) at Bob, and R is the transmission rate measured by bits per channel use (bpcu).

Since the decoding error probability (??) is affected by fading channels h_{ab} and h_{wb}, the average decoding error probability $\overline{\delta}$ is adopted to evaluate the reliability performance. And the effective throughput of the system is given by [?]

$$\eta = nR\left(1-\overline{\delta}\right), \quad (6)$$

which quantifies the expected number of information bits that can be reliably transmitted from Alice to Bob.

III. Covertness Performance Analysis

In this section, to analyze the covertness performance of the system, the average detection error probability is derived.

With detection threshold τ, the detection error probability is expressed as [?]

$$\xi(\tau) = 1 - \frac{\gamma\left(n, \frac{n\tau}{\sigma^2}\right)}{\Gamma(n)} + \frac{\gamma\left(n, \frac{n\tau}{\sigma^2 + P_a|h_{aw}|^2}\right)}{\Gamma(n)}, \quad (7)$$

where $\sigma^2 = \varphi P_w + \sigma_w^2$ for expression simplification.

Since Willie knows h_{aw} in each round, Willie can adjust the optimal threshold τ^* to minimize the detection error probability for each round, which is given by [?]

$$\tau^* = \frac{\sigma^2\left(\sigma^2 + P_a|h_{aw}|^2\right)}{P_a|h_{aw}|^2}\ln\left(\frac{\sigma^2 + P_a|h_{aw}|^2}{\sigma^2}\right). \quad (8)$$

Notably, from the perspective of Alice, only statistical CSI is available. Therefore, the average detection error probability is derived as the covertness metric [?].

Theorem 1. *The average detection error probability at Willie with optimal detection threshold under Rayleigh fading channels can be derived as*

$$\overline{\xi}(\tau^*) = 1 - \frac{\pi}{BP_a\lambda_{aw}\Gamma(n)}\sum_{i=1}^{B}\left[\gamma\left(n, \frac{n(\sigma^2+\tan\theta_i)}{\tan\theta_i}\ln\left(\frac{\sigma^2+\tan\theta_i}{\sigma^2}\right)\right) - \gamma\left(n, \frac{n\sigma^2}{\tan\theta_i}\ln\left(\frac{\sigma^2+\tan\theta_i}{\sigma^2}\right)\right)\right]\frac{e^{-\frac{\tan\theta_i}{P_a\lambda_{aw}}}\sqrt{\theta_i\left(\frac{\pi}{2}-\theta_i\right)}}{\cos^2\theta_i}, \quad (9)$$

where B is the parameter of Gaussian-Chebyshev Quadrature, and $\theta_i = \frac{\pi}{4}\left(1+\cos\frac{(2i-1)\pi}{2B}\right)$.

Proof. By substituting (??) into (??) and considering the probability density function (PDF) of Rayleigh fading channel, the average detection error probability can be expressed as

$$\overline{\xi}(\tau^*) = 1 - \frac{1}{P_a\lambda_{aw}\Gamma(n)} \times \int_0^{+\infty}\left[\gamma\left(n, \frac{n(\sigma^2+x)}{x}\ln\left(\frac{\sigma^2+x}{\sigma^2}\right)\right) - \gamma\left(n, \frac{n\sigma^2}{x}\ln\left(\frac{\sigma^2+x}{\sigma^2}\right)\right)\right]e^{-\frac{x}{P_a\lambda_{aw}}}dx. \quad (10)$$

By substituting $x = \tan\theta$ into (??) and applying Gaussian-Chebyshev Quadrature into the above integral expression [?], (??) can be obtained, and the proof is completed. □

Due to the complicated form of (??), it is intractable to further guide the system design. Thus, a tractable lower approximation of the detection error probability in one transmission round is derived first, and then a lower approximation of the average detection error probability is derived.

Theorem 2. *A lower approximation of the minimum detection error probability in one transmission round is given by*

$$\xi^l(\tau^*) = \begin{cases} 1 - \frac{e^{-n}n^n}{\Gamma(n)}\ln\left(1+\frac{P_a|h_{aw}|^2}{\sigma^2}\right), & \frac{P_a|h_{aw}|^2}{\sigma^2} < e^{\frac{\Gamma(n)}{e^{-n}n^n}} - 1 \\ 0, & \frac{P_a|h_{aw}|^2}{\sigma^2} \geq e^{\frac{\Gamma(n)}{e^{-n}n^n}} - 1 \end{cases} \quad (11)$$

Proof. See Appendix ??. □

The lower approximation of the minimum detection error probability of (??) is tighter than the approximation based on KL divergence (i.e., $\xi^{KL} = 1 - \sqrt{\frac{1}{2}\mathcal{D}(\mathbb{P}_0||\mathbb{P}_1)}$, see Appendix ?? for the detailed definition), which is widely used to evaluate the covertness performance in the existing works [?], [?]. The detailed proof is given in Appendix ??.

The above concise approximation facilitates the performance analysis and optimization design for the covert communication system. It can be used as a metric for the system with AWGN channels [?] or the fading channels when only considering one transmission round [?]. Besides, it can also be adopted to analyze the average detection error probability in fading channels as follows [?], [?].

Based on Theorem 2, the average detection error probability at Willie is derived as follows.

$$\overline{\xi}(\tau^*) \approx \int_0^{\sigma^2\left(e^{\frac{\Gamma(n)}{e^{-n}n^n}}-1\right)} \frac{e^{\frac{-x}{P_a\lambda_{aw}}}}{P_a\lambda_{aw}}\left(1 - \frac{e^{-n}n^n}{\Gamma(n)}\ln\left(1+\frac{x}{\sigma^2}\right)\right)dx = 1 - \frac{e^{-n}n^n}{\Gamma(n)}e^{\frac{\sigma^2}{P_a\lambda_{aw}}}\left[\mathrm{E}_1\left(\frac{\sigma^2}{P_a\lambda_{aw}}\right) - \mathrm{E}_1\left(\frac{\sigma^2}{P_a\lambda_{aw}}e^{\frac{\Gamma(n)}{e^{-n}n^n}}\right)\right]. \quad (12)$$

The expression of average detection error probability (??) and its approximation (??) can be extended to the covert communication scenario with a passive warder by setting $P_w = 0$. Besides, the lower approximation of the detection error probability in one transmission round given in (??) can also be extended to the scenario with a passive warder by setting $P_w = 0$, which can replace the KL divergence approximation widely used in the existing works since it is tighter.

IV. Reliability Performance Analysis and System Design

In this section, we analyze the reliability performance of the system with AWGN channels and fading channels, and then maximize the effective throughput by jointly optimizing the system parameters. Besides, power, the transmission rate and the blocklength where both the covertness and the reliability requirements are considered.

Based on Theorem 2, the average detection error probability at Willie is derived as follows.

$$\xi(r^*) \approx \int_0^{\infty} \frac{e^{-\frac{x}{P_a\lambda_{aw}}}}{P_a\lambda_{aw}} \left(1 - \frac{e^{-n}n^n}{\Gamma(n)} \ln\left(1+\frac{x}{\sigma^2}\right)\right) dx$$
$$= 1 - \frac{e^{-n}n^n}{\Gamma(n)} e^{\frac{\sigma^2}{P_a\lambda_{aw}}} \left[\mathrm{E}_1\left(\frac{\sigma^2}{P_a\lambda_{aw}}\right) - \mathrm{E}_1\left(\frac{\sigma^2}{P_a\lambda_{aw}} e^{\frac{\Gamma(n)}{e^{-n}n^n}}\right)\right]. \quad (12)$$

The expression of average detection error probability (??) and its approximation (??) can be extended to the covert communication scenario with a passive warder by setting $P_w = 0$. Besides, the lower approximation of the detection error probability in one transmission round given in (??) can also be extended to the scenario with a passive warder by setting $P_w = 0$, which can replace the KL divergence approximation widely used in the existing works since the proposed concise approximation is tigher than the conventional one as proved in Appendix ??.

IV. Reliability Performance Analysis and System Design

In this section, to analyze the reliability performance, the decoding error probability is derived. Then, the effective throughput is maximized by jointly optimizing the transmit power, the transmission rate and the blocklength, where both the covertness and the reliability requirements are considered.

A. Average Decoding Error Probability at Bob

Since both the transmit power at Alice and the jamming power at Willie affect the received SNR at Bob, considering the fading channels, the PDF of SNR can be derived as

$$f_{\gamma_b}(t) = \frac{d\Pr(\gamma_b<t)}{dt} = \frac{d}{dt}\Pr\left(P_a|h_{ab}|^2 < tP_w|h_{wb}|^2 + t\sigma_b^2\right)$$
$$= \frac{\sigma_b^2(P_a\lambda_{ab} + P_w\lambda_{wb}t) + P_aP_w\lambda_{ab}\lambda_{wb}}{(P_a\lambda_{ab} + P_w\lambda_{wb}t)^2} e^{-\frac{\sigma_b^2 t}{P_a\lambda_{ab}}}. \quad (13)$$

Based on the linear approximation of Q-function given in [?] and the PDF of SNR given above, the average decoding error probability can be derived as

$$\bar{\delta} \approx \int_0^{\alpha-\frac{1}{2\beta}} f_{\gamma_b}(t)\, dt + \int_{\alpha-\frac{1}{2\beta}}^{\alpha+\frac{1}{2\beta}} \left[\frac{1}{2} - \beta(t-\alpha)\right] f_{\gamma_b}(t)\, dt$$
$$= 1 - \frac{P_a\lambda_{ab} e^{-\frac{\sigma_b^2}{P_a\lambda_{ab}}\left(\alpha-\frac{1}{2\beta}\right)}}{P_a\lambda_{ab} + P_w\lambda_{wb}\left(\alpha-\frac{1}{2\beta}\right)} + g\left(\alpha+\frac{1}{2\beta}\right) - g\left(\alpha-\frac{1}{2\beta}\right), \quad (14)$$

where $\alpha = 2^R - 1$, $\beta = \sqrt{\frac{n}{2\pi(4^R-1)}}$ and $g(x) = e^{-\frac{\sigma_b^2 x}{P_a\lambda_{ab}}} \frac{P_a\lambda_{ab}}{P_w\lambda_{wb}} \left(\frac{2\beta P_a\lambda_{ab} + 2\alpha\beta P_w\lambda_{wb}x - P_w\lambda_{wb}}{2(P_a\lambda_{ab}+P_w\lambda_{wb}x)}\right) + e^{\frac{\sigma_b^2}{P_w\lambda_{wb}}} \frac{\beta P_a\lambda_{ab}}{P_w\lambda_{wb}} \mathrm{E}_1\left(\frac{x\sigma_b^2}{P_a\lambda_{ab}} + \frac{\sigma_b^2}{P_w\lambda_{wb}}\right)$.

The above result can be extended to the scenario with a passive warder by replacing $P_w = 0$, and the corresponding average decoding error probability can be simplified as

$$\bar{\delta} \approx 1 + \frac{P_a\lambda_{ab}\beta}{\sigma_b^2} e^{-\frac{\sigma_b^2\alpha}{P_a\lambda_{ab}}} \left(e^{-\frac{\sigma_b^2}{2P_a\lambda_{ab}\beta}} - e^{\frac{\sigma_b^2}{2P_a\lambda_{ab}\beta}}\right). \quad (15)$$

B. Effective Throughput Maximization Optimization

Based on the above analysis, an optimization problem can be formulated to maximize the effective throughput of the system subject to the combined constraints of covertness, reliability, blocklength, and transmit power.

$$(\mathrm{P1}): \max_{P_a, R, n} \ \eta \quad (16)$$
$$\mathrm{s.t.}\ \xi(r^*) \geq 1-\varepsilon, \quad (16a)$$
$$\bar{\delta} \leq \kappa, \quad (16b)$$
$$P_a \leq P_a^{max}, \quad (16c)$$
$$n_{min} \leq n \leq n_{max}, n \in \mathbb{N}^+, \quad (16d)$$

where (16a) and (16b) denote the covertness and reliability constraints with predetermined covertness and reliability requirements ε, κ, respectively. (16c) denotes the transmit power constraints at Alice with the maximum power P_a^{max}. Besides, (16d) denotes the blocklength constraint due to delay requirements and channel coding requirements with the maximum blocklength n_{max} and the minimum blocklength n_{min}.

To solve this optimization problem with coupled optimization variables P_a, R and n, a joint optimization framework is proposed as follows, which involves a two-layer process. In the inner-layer, the optimal transmit power P_a^* and the optimal transmission rate R^* are derived with a given blocklength n. In the outer-layer, the optimal blocklength n^* can be obtained via an exhaustive search over $[n_{\min}, n_{\max}]$, where P_a^* and R^* are calculated with each value of n. Finally, the globally optimal solutions P_a^*, R^* and n^* can be obtained by the above framework. The details are elaborated below.

Inner-layer stage: When n is given, the optimal transmit power can be derived by $P_a^* = \min\{P_a^{\max}, P_a^o\}$, where P_a^o is the solution of $\xi(r^*) = 1-\varepsilon$. This is because $\xi(r^*)$ and $\bar{\delta}$ decrease with P_a, and the effective throughput increases with P_a. After the optimal transmit power P_a^* is obtained, we can obtain the optimal transmission rate R^* as follows. It is verified that the effective throughput is a first-increasing and then-deceasing function of R [?]. Consequently, the optimal R^o that maximizes η can be effectively calculated using the bisection method. Considering the reliability constraint (??), the maximum transmission rate R^{max} can be derived by solving $\bar{\delta} = \kappa$. Thus, the optimal transmission rate $R^* = \min\{R^o, R^{\max}\}$.

Outer-layer stage: Considering that the blocklength n affects the constraints (??), (??) and the objective function, it is difficult to derive the optimal solutions directly. By exhaustive search over n over $[n_{\min}, n_{\max}]$, the globally optimal solutions for (P1) and the maximum effective throughput η^* can be obtained.

V. Numerical Simulation

In this section, we provide numerical results to show the covert and reliable performance of the short-packet communication system against a proactive warder. The parameter settings are as follows, unless specified otherwise: the fading parameters $\lambda_{ab} = 5\times10^{-2}$, $\lambda_{aw} = \lambda_{wb} = 10^{-3}$, the AWGN variances $\sigma_b^2 = \sigma_w^2 = 10^{-1}$ (W), the self-interference cancellation coefficient $\phi = 10^{-4}$, the blocklength $n = 100$, the minimum blocklength $n_{min} = 50$, the maximum blocklength $n_{max} = 200$, the maximum transmit power $P_a^{max} = 5$ (W), the covertness requirement $\varepsilon = 10^{-1}$ and the reliability requirement $\kappa = 10^{-1}$. All the simulation results shown in this paper are obtained by averaging over 10^6 channel realizations.

In Fig. 2, the impact of the transmit (jamming) power on the average detection error probability. The curves with "Sim", "Eq. (9)", "Appr. (12)", and "KL appr" denote the results obtained by numerical simulations, the analytical expression (??), the approximation (??) and the KL divergence, respectively. It can be seen that the curves with numerical simulations and (??) coincide. However, a significant gap exists between the KL approximation and simulation results. Besides, the results validate the analysis given in Section III, implying that the proposed approximation expression (??) can be used as a metric to replace

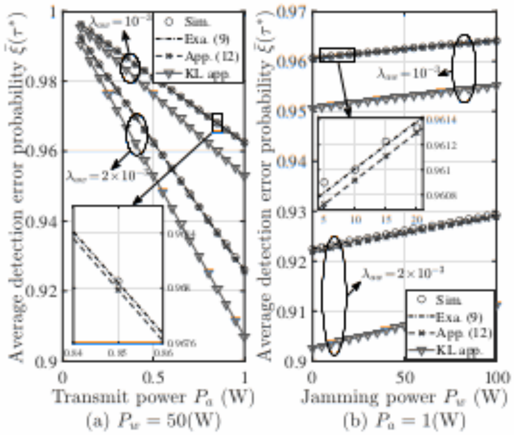

Fig. 2. The detection error probability versus the transmit/jamming power.

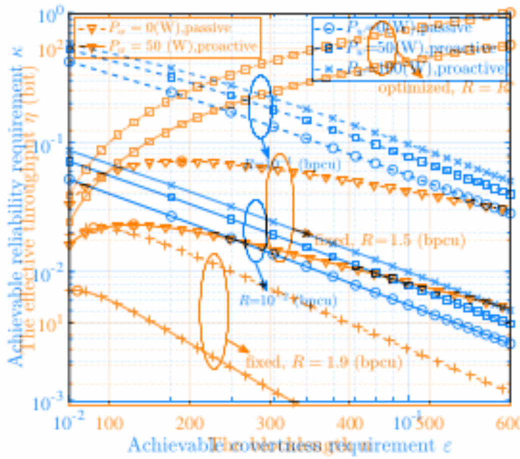

Fig. 3. The achievable reliability requirement versus the achievable covertness requirement.

V. Numerical Simulation

In this section, we provide numerical results to show the covert and reliable performance of the short-packet communication system against a proactive warder. The parameter settings are as follows, unless specified otherwise: the fading parameters $\lambda_{ab} = 5 \times 10^{-2}, \lambda_{aw} = \lambda_{wb} = 10^{-3}$, the AWGN variances $\sigma_b^2 = \sigma_w^2 = 10^{-1}$ (W), the self-interference cancellation coefficient $\phi = 10^{-4}$, the blocklength $n = 100$, the minimum blocklength $n_{min} = 50$, the maximum blocklength $n_{max} = 200$, the maximum transmit power $P_a^{max} = 5(W)$ the covertness requirement $\varepsilon = 10^{-1}$ and the reliability requirement $\kappa = 10^{-1}$. All the simulation results shown in this paper are obtained by averaging over 10^6 channel realizations.

In Fig.2, the impact of the transmit (jamming) power on the average detection error probability is investigated. The curves with "Sim.", "Exa. (9)", "App. (12)", and "KL app." denote the results obtained by numerical simulations, the exact analytical expression of (??), the approximation (??), and the numerical integration combined with the KL divergence of (??), respectively. It can be seen that the curves with numerical simulations, (??) and (??) almost coincide. However, a significant gap exists between the curves with the KL divergence approximation and the simulation results. Besides, the average detection error probability decreases with P_a and increases with P_w. These results validate the results given in Section III, implying that the proposed approximation expression (??) can be adopted as a covertness metric to replace the widely used KL divergence metric due to its conciseness and tightness.

In Fig. 3, the relationship among the achievable covertness requirement ε and reliability requirement κ is investigated. It can be seen from the figure, larger ε is tolerant, and smaller κ can be achieved. Conversely, larger κ is tolerant, and smaller ε can be achieved. Besides, it can be seen that the system performance (covertness and reliability performance) is degraded by the proactive warder $P_w = 50, 100$ (W) compared with the passive one $P_w = 0$ (??). These results show that the tradeoff between transmission covertness and reliability is changed by the proactive warder, and the proposed performance evalutions in Sections III and IV can be adopted to guide the system design in this case.

In Fig. 4, the impact of blocklength on the effective throughput is investigated. The transmission rate is either fixed or optimized. The curves with marked solid lines and marked dotted lines denote the system performance in the system with a proactive warder and that with a passive warder, respectively. The red dots in the figure indicate the optimal blocklength that maximizes the throughput. It can be seen that in the system with fixed transmission rates, the effective throughput first increases and then decreases with n. This is because when n is too small, η is directly limited by the blocklength. On the contrary, when n is too large, the transmit power is limited by the covertness constraint and the decoding error is too large, resulting in the reduction of effective throughput. In addition, the effective throughput with an optimized transmission rate is always higher than that with a fixed transmission rate, which demonstrates the feasibility of the proposed optimization framework. These results imply that for the system with optimized rates, a longer blocklength is always beneficial to improve the effective throughput. However, for the system with a fixed rate, the optimal blocklength is not necessarily the maximum one, which is critical for the system design.

Fig. 4. The effective throughput with optimized/fixed transmission rates versus the blocklength.

VI. Conclusion

In this paper, we investigated the reliable and covert performance of short-packet communication systems against a proactive warder ... to maximize the effective throughput. Numerical results ...

Appendix A
Proof of Theorem 2

We denote $f^u(x) = n\left(\frac{1}{x}+1\right)\ln(1+x)$, $f^l(x) = \frac{n}{x}\ln(1+x)$ ... $\frac{f^u(x)}{f^l(x)}$ ... $\frac{P_a|h_{aw}|^2}{\sigma^2}$ ...

the covertness constraint. In addition, the effective throughput with optimized transmission rate is always higher than that with a fixed transmission rate, which demonstrates the feasibility of the proposed optimization framework. These results imply that for the system with optimized rates, a longer blocklength is always beneficial to improve the effective throughput. However, for the system with a fixed rate, the optimal blocklength is not necessarily the maximum one, which is critical for the system design.

VI. Conclusion

In this paper, we investigated the reliable and covert performance of short-packet communication systems against a proactive warden. Specifically, the average detection error probability and its approximation were derived to evaluate the covertness performance. In addition, the average decoding error probability was derived to evaluate the reliability performance. Based on the analysis above, an optimization framework was proposed to maximize the effective throughput. Numerical results verified the feasibility of the proposed approximations and the optimization framework. The performance loss brought by a proactive warden was investigated compared with the passive one, and the optimal blocklength to maximize the effective throughput was elaborated with different systems.

Appendix A
Proof of Theorem 2

We denote $f^u(x) = n\left(\frac{1}{x}+1\right)\ln(1+x)$, $f^l(x) = \frac{n}{x}\ln(1+x)$ and $g(x) = \int_{f^l(x)} \frac{e^{-t}t^{n-1}}{\Gamma(n)} dt$ where $x = \frac{P_a|h_{aw}|^2}{\sigma^2}$. In the high covertness scenarios, x always approaches zero to meet the covertness requirement, resulting in $f^u(x) \to n$ and $f^l(x) \to n$. Thus, $g(x)$ can be approximated as $g(x) \approx e^{-n}n^n \ln\left(\frac{\sigma^2+P_a|h_{aw}|^2}{\sigma^2}\right)$ and the approximation (??) is obtained.

Below, we prove (??) is a lower bound of (??).

When $x \geq e^{\frac{\Gamma(n)}{e^{-n}n^n}} - 1$, it holds that $\xi^l(\tau^*) = 0 < 1 - \frac{g(x)}{\Gamma(n)}$.

When $0 \leq x < e^{\frac{\Gamma(n)}{e^{-n}n^n}} - 1$, we define the function $f_1(x) = g(x) - e^{-n}n^n \ln(1+x)$ with $f_1(0) = 0$ and $\frac{df_1(x)}{dx} = \frac{1}{(x+1)}\left[\left(e^{\frac{\ln(1+x)}{(1+x)^{\frac{1}{x}}x}}\right)^n - 1\right]$. We denote the function $f_2(x)$ with $f_2(0) = \frac{1}{e}$ and $\frac{df_2(x)}{dx} = \frac{(\log(x+1)-x)((x+1)\log(x+1)-x)}{(x+1)^{1/x+1}x^3} < 0$. Thus, the first-order derivative of $f_1(x)$ is small than 0, and $f_1(x) \leq f_1(0) = 0$.

Thus, for $0 \leq x < e^{\frac{\Gamma(n)}{e^{-n}n^n}} - 1$, we can obtain $\xi^l(\tau^*) = 1 - \frac{e^{-n}n^n}{\Gamma(n)}\ln(1+x) < 1 - \frac{g(x)}{\Gamma(n)}$, and Theorem 2 is proved.

Appendix B
Comparison between (11) and KL Divergence

By adopting the Pinsker's inequality, a lower bound of minimum detection error probability is given by [?], [?]

$$\xi^{KL} = 1 - \sqrt{\frac{1}{2}\mathcal{D}\left(\mathbb{P}_0||\mathbb{P}_1\right)} \quad (17)$$

where $\mathbb{P}_0$ and $\mathbb{P}_1$ denote the probability distributions of the observations at Willie under $\mathcal{H}_0$ and $\mathcal{H}_1$, respectively. $\mathcal{D}\left(\mathbb{P}_0||\mathbb{P}_1\right)$ is the KL divergence from $\mathbb{P}_0$ to $\mathbb{P}_1$ as $\mathcal{D}(\mathbb{P}_0||\mathbb{P}_1) = n\left(\ln\left(1+\frac{P_a|h_{aw}|^2}{\sigma^2}\right)+\frac{\sigma^2}{\sigma^2+P_a|h_{aw}|^2}-1\right)$.

Then, we prove that (??) is tighter than the KL divergence approximation, i.e., $\xi\left(\tau^*\right) > \xi^l(\tau^*) > \xi^{KL}$.

We denote $f_3(x) = -2e^{-2n}n^{2n-1}(\Gamma(n))^{-2}\ln^2 x + \ln x + \frac{1}{x} - 1$ with $\frac{df_3(x)}{dx} = \frac{1}{x}\left(-2e^{-2n}n^{2n-1}(\Gamma(n))^{-2}\ln x + 1 - \frac{1}{x}\right)$. In addition, we denote $f_4(x) = -2e^{-2n}n^{2n-1}(\Gamma(n))^{-2}\ln x + 1 - \frac{1}{x}$ with $\frac{df_4(x)}{dx} = x^{-2} - 2e^{-2n}n^{2n-1}(\Gamma(n))^{-2}x^{-1}$, where $2e^{-2n}n^{2n-1}(\Gamma(n))^{-2} < 1$, proved as follows.

By denoting $M_1(n) = \frac{\Gamma(n)}{\sqrt{2}ne^{-n}n^{n-1}}$, we can obtain $\frac{\partial M_1(n)}{\partial n} = -\frac{\sqrt{2}}{4}e^n n^{-n-3/2} n! \left(2n\log(n) - 2n\psi^{(0)}(n) - 1\right) < 0$, and $M_1(1) = \frac{e}{\sqrt{2}} > 1$, $\lim_{n\to\infty} M_1(n) = \sqrt{\pi} + \mathcal{O}\left(\frac{1}{n}\right) > 1$. Thus, $2e^{-2n}n^{2n-1}(\Gamma(n))^{-2} \in \left(\frac{1}{\pi}, \frac{2}{e^2}\right)$.

Therefore, $f_4(x)$ increases in $\left[1, \frac{1}{2}e^{2n}n^{1-2n}(\Gamma(n))^2\right)$ and decreases in $\left(\frac{1}{2}e^{2n}n^{1-2n}(\Gamma(n))^2, +\infty\right)$. In addition, $f_4(1) = 0$ and $f_4(x)$ is larger than 0 in the interval $[1, \kappa_1)$ and less than 0 in the interval $(\kappa_1, +\infty)$, where κ_1 is the solution to $-2e^{-2n}n^{2n-1}(\Gamma(n))^{-2}x\ln x + x = 1$ except 1. Futhermore, $f_3(x)$ increases in $[1, \kappa_1)$ and decreases in $(\kappa_1, +\infty)$. When $x = e^{\frac{\Gamma(n)}{e^{-n}n^n}}$, we can obtain $f_3(e^{\frac{\Gamma(n)}{e^{-n}n^n}}) = \frac{1}{n}\left(n\left(\frac{\Gamma(n)}{e^{-n}n^n} + e^{-\frac{\Gamma(n)}{e^{-n}n^n}} - 1\right) - 2\right)$, where the sequence $M_2(n) = n\left(\frac{\Gamma(n)}{e^{-n}n^n} + e^{-\frac{\Gamma(n)}{e^{-n}n^n}} - 1\right)$ is monotonically increasing with n and $M_2(1) = 1.78 < 2$, $M_2(2) = 2.01 > 2$. Therefore, $f_3(x) > 0$ in $\left[1, e^{\frac{\Gamma(n)}{e^{-n}n^n}}\right]$ with $n \geq 2$. And by substituting $x = 1 + \frac{P_a|h_{aw}|^2}{\sigma^2}$ into $f_3(x)$, and adopting the results derived above, we can obtain $\xi^l(\tau^*) = 1 - \frac{e^{-n}n^n}{\Gamma(n)}\ln\left(1+\frac{P_a|h_{aw}|^2}{\sigma^2}\right) > 1 - \sqrt{\frac{1}{2}\mathcal{D}\left(\mathbb{P}_0||\mathbb{P}_1\right)}$ for $0 < \frac{P_a|h_{aw}|^2}{\sigma^2} \leq \left(e^{\frac{\Gamma(n)}{e^{-n}n^n}} - 1\right)$.

Besides, when $\frac{P_a|h_{aw}|^2}{\sigma^2} > e^{\frac{\Gamma(n)}{e^{-n}n^n}} - 1$, $\xi^l(\tau^*) = 0 > \xi^{KL}$. Note that the above results hold with the assumption $n \geq 2$, which is always true in short-packet communications.