

Covert Transmission in Water-to-Air Optical Wireless Communication Systems

Qingqing Hu^{ID}, Tianrui Lin^{ID}, Tianjian Wei, Nuo Huang^{ID}, Yi-Jun Zhu^{ID}, and Chen Gong^{ID}, *Senior Member, IEEE*

Abstract— Covert communications hide the information transmission from a watchful adversary (Willie) while ensuring a satisfactory decoding performance at a legitimate receiver (Bob). In this paper, we propose a covert transmission scheme based on narrow signal waveband and nonlinear suppression effect of avalanche photodiodes (APDs) for water-to-air (W2A) optical wireless communication (OWC) systems. Narrow signal waveband is selected based on a given wavelength selection pattern which is negotiated by the transmitter and receiver. Willie’s detection performance is investigated under APD characteristics and ambient radiation. Under the covertness constraints on Willie’s detection performance, the covert throughput between Alice and Bob is maximized by optimizing the power and blocklength of transmitted signals. Except for the case of Alice with knowledge of Willie’s position and optical filter passband-width, we explore a more practical scenario where the knowledge is not available to Alice. The optimal passband-width of Willie’s filter is calculated by taking into account both overlapped waveband and detection performance, and a covert region is established to limit Willie’s detection range. Numerical and experimental results indicate that strong ambient radiation and narrow signal waveband can enhance the signal covertness significantly.

Index Terms— Covert transmission, water-to-air, optical wireless communication, ambient light.

I. INTRODUCTION

WITH the development of communication technologies, there is an increasing demand for seamless global coverage and extremely low delay [1]. Direct communication across the water-air interface is a fundamental and critical issue, which can be applied to underwater resources exploration and cooperative rescue [2].

Traditional radio frequency (RF) signal experiences significant signal absorption, and acoustic signal exhibits high latency for underwater communication. In contrast, optical

Manuscript received 13 July 2023; revised 27 December 2023; accepted 22 January 2024. Date of publication 18 March 2024; date of current version 6 May 2024. This work was supported in part by the National Natural Science Foundation of China under Grant 62171428 and Grant 62101526 and in part by the Fundamental Research Funds for Central Universities under Grant KY2100000118. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Hossein Pishro-Nik. (*Corresponding authors:* Nuo Huang; Chen Gong.)

Qingqing Hu, Tianrui Lin, Tianjian Wei, Nuo Huang, and Chen Gong are with the CAS Key Laboratory of Wireless-Optical Communications, University of Science and Technology of China, Hefei 230052, China (e-mail: ruixihu@mail.ustc.edu.cn; trlin@mail.ustc.edu.cn; weitj@mail.ustc.edu.cn; huangnuo@ustc.edu.cn; cgong821@ustc.edu.cn).

Yi-Jun Zhu is with the School of Information System Engineering, Information Engineering University, Zhengzhou, Henan 450000, China (e-mail: yijunzhu@zzu.edu.cn).

Digital Object Identifier 10.1109/TIFS.2024.3376965

wireless communication (OWC) has much higher bandwidth and moderate loss in the water, becoming a promising candidate for short- and medium-range underwater communication [3]. In [4], the communication feasibility of water-to-air (W2A)-OWC system with transmission distance of tens of meters was investigated by Monte Carlo simulation. A practical W2A-OWC system was established in [5] to demonstrate the feasibility of direct W2A communication over a 2.3-m underwater and 3.5-m air link. In [6], bi-directional OWC links through the air-water interface were tested in the ocean.

The received signal is usually weak due to W2A channel fading. Compared to standard PIN photodiode, avalanche photodiode (APD) has higher sensitivity, speed, gain and larger detection dynamic range. Therefore, APDs are widely adopted for weak signal detection [5], [6], [7]. Other photodetectors including single photon avalanche diode (SPAD) and photomultiplier tube (PMT) mainly operate at the photon level, and suffer saturation under strong ambient radiation [8]. Therefore, we consider the best scheme for Bob and Willie, i.e., both Bob and Willie adopt the APD for signal detection.

In W2A-OWC systems, water surface fluctuations and large divergence of light-emitting diodes (LEDs) may cause information leakage, and the information privacy in legitimate links is under threat from a watchful adversary (Willie). Unlike physical layer security that prevents information from being decrypted in communication, covert communication with low probability of being detected aims to hide the existence of wireless transmission, which especially is a key issue for military and law enforcement agencies [9]. Therefore, covert communication emerges as a potential solution owing to its high-security level [10].

Recent researches on covert communication can be divided into two categories: 1) Fundamental limits: The square root law on the amount of covert information under additional white Gaussian noise (AWGN) was presented in [11], i.e., $\mathcal{O}(\sqrt{n})$ bits can be covertly transmitted in n channel uses. The square root law in [11] can be extended to binary symmetric channels [12], discrete memoryless channels [13] and multiple access channels [14]. The maximum achievable covert rate in Poisson channels was derived in [15], and the square root law holds in the band- and peak-limited case. The upper bounds on the maximum amount of information transmitted covertly in free-space optical (FSO) communication and indoor OWC scenarios were derived in [16] and [17] under the constraints of signal nonnegativity, average optical power and peak optical

power. A coding scheme over discrete memoryless channels was developed in [18] to achieve covert communication in the order of \sqrt{n} bits over n channel uses, requiring only an order of \sqrt{n} bits of secret key. It is possible to communicate without secret key if the receiver's channel is better than the warden's channel.

2) Covert schemes: A full-duplex receiver was adopted in [19] to generate artificial noise with varying power which causes uncertainty at Willie in Rayleigh fading channels. Work [20] considered the case of Willie with uncertain locations, where the connection throughput between a multi-antenna transmitter and a full-duplex jamming receiver was maximized under the optimal location of Willie. The covert communication between a ground transmitter and a legitimate receiver was aided by an unmanned aerial vehicle (UAV) with intelligent reflecting surface (IRS) in [21] or an UAV relay in [22]. For millimeter-wave communication systems, a framework that jointly optimizes beam training duration, training power and date transmission power was proposed in [23] to maximize the throughput under covertness constraints. Work [24] adopted public messages from reliable transmitter to protect covert messages against Willie's detection in uplink non-orthogonal multiple access (NOMA) systems. Non-equiprobable constellations modulation schemes were proposed in [25] to enhance the covert rate. An optimal low-complexity coding scheme was developed in [26] to achieve the information-theoretic limits of covert communications over binary-input discrete memoryless channels. In [27], the UAV's trajectory and transmit power were optimized to maximize the average covert transmission rate in UAV networks. Uniformly distributed random transmit power was adopted in [28] to enhance the performance of delay-intolerant covert communications in AWGN channels, and the optimal blocklength was analytically proven to be the maximum blocklength under fixed transmit power. To the best of the authors' knowledge, the covert scheme for outdoor W2A-OWC systems has not been systematically investigated in previous studies.

In this work, we propose a covert transmission scheme based on narrow signal waveband for W2A-OWC systems, taking advantage of APD characteristics and strong ambient radiation in outdoor environments. The detection performance at Willie is investigated under APD nonlinearity and shot noise. For the case of Alice with knowledge of Willie's position and optical filter passband-width, the power and blocklength of transmitted signals are optimized to maximize the covert throughput between Alice and Bob, while ensuring that Willie's detection performance satisfies given covertness constraints. We further extend to a more practical scenario where Willie's position and optical filter passband-width are not available. To address the covert optimization problem, the optimal passband-width of Willie's filter is explored to achieve the tradeoff between overlapped waveband and detection performance, and a covert region is established to constrain Willie's detection range. The contributions of this paper can be summarized comprehensively as follows:

- 1) We propose a covert transmission scheme in dynamic W2A-OWC scenarios for the first time. The proposed scheme takes advantage of secret selection of

narrow-waveband signal, APD characteristics and outdoor ambient radiation to achieve signal hiding, without the need for extra power and array to generate artificial noise or jamming signals.

- 2) We investigate the characteristics of APD receivers including APD nonlinearity and noise. We characterize Willie's detection performance under the APD characteristics and ambient radiation, and derive the covert throughput between Alice and Bob under W2A-OWC channel fading.
- 3) We formulate optimization problems that maximize the covert throughput of the proposed scheme by optimizing power and blocklength of transmitted signals. Except for the case of Alice with knowledge of Willie's position and optical filter passband-width, we further investigate a more general case where Willie's information is unknown to Alice. We compute the optimal passband-width of Willie's filter by considering both overlapped waveband and detection performance, and establish a covert region to limit Willie's detection range.
- 4) We numerically explore the maximum covert throughput under different conditions, and experimentally investigate the effects of ambient light on detection performance. It is demonstrated that strong ambient radiation and narrow signal waveband can enhance the signal covertness.

The remainder of this paper is organized as follows. In Section II, we introduce the W2A-OWC system model and APD characteristics including nonlinearity and noise. In Section III, we propose a covert transmission scheme based on narrow signal waveband as well as nonlinear suppression effect of APDs, and investigate Willie's detection performance accordingly. In Section IV, we optimize the power and blocklength of transmitted signals to maximize the covert throughput for two scenarios where Willie's position and optical filter passband-width are available or unavailable. In Section V, we numerically evaluate the maximum covert throughput with the optimized power and blocklength. In Section VI, we experimentally investigate the effects of ambient light on signal covertness. Finally, we conclude this work in Section VII.

II. SYSTEM MODEL

A. Channel Model

We consider a W2A-OWC system as shown in Fig. 1(a). An underwater vehicle (transmitter Alice) adopts visible light to communicate with an UAV (legitimate receiver Bob) in the air, while another UAV (adversary Willie) tries to detect whether the transmission from Alice to Bob happens. Due to random fluctuation of the water surface, the channel gains from Alice to both receivers vary randomly with time, modeled as lognormal block fading. Assume that Bob's channel gain $h_b \sim \text{Lognormal}(\mu_b, \sigma_b^2) \triangleq f'(h_b)$ and Willie's channel gain $h_w \sim \text{Lognormal}(\mu_w, \sigma_w^2) \triangleq f'(h_w)$ [29].

Another crucial factor affecting the transmission performance is the ambient radiation. Figure 1(b) shows the ambient light power measured by an optical power meter (Thorlabs, PM100D with sensor S130C) at central wavelength 520 nm.

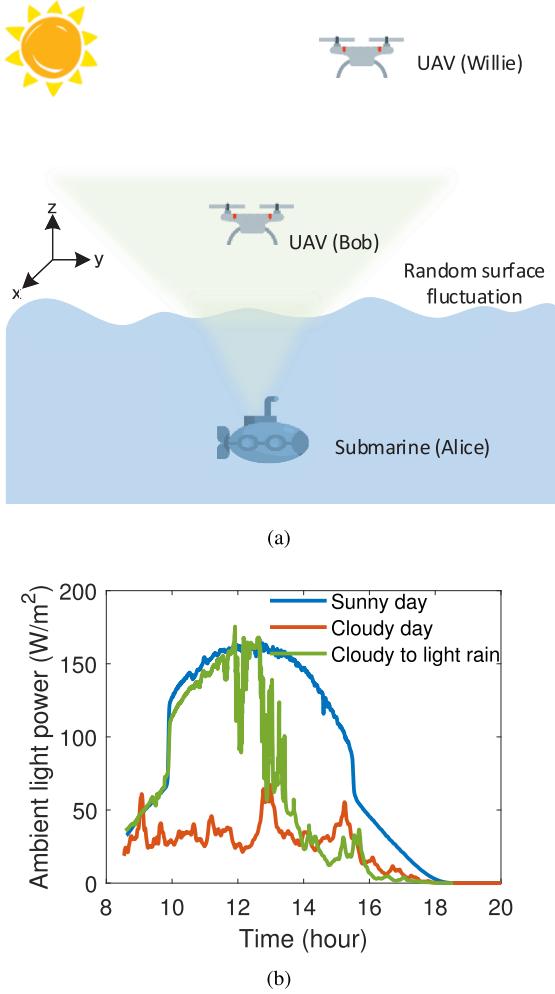


Fig. 1. (a) Illustration of the covert transmission in a W2A-OWC system. (b) Ambient light power during daytime.

It is seen that the ambient light power is mostly higher than 15 W/m^2 during daytime under three different weather conditions.

B. Transceiver Setting

Alice modulates the information bits to the emitted light intensity of an LED at the transmitter. Following existing works [30], [31], [32] on covert communication where Bob and Willie have the same hardware configuration, we assume that both Bob and Willie adopt one APD as the receiver. For covert transmission, Alice and Bob can share the secret information in advance to facilitate the covert design. Assume that Bob's wavelength range $[\lambda_{b1}, \lambda_{b2}]$ and Willie's wavelength range $[\lambda_{w1}, \lambda_{w2}]$ are selected uniformly in the range $[V, V + L]$.

C. Characteristics of APD Receivers

1) APD Nonlinearity: Note that APDs exhibit nonlinearity under strong ambient light, resulting in saturation region [33]. According to the equivalent circuit model in APDs, the relationship between incident optical power P and output

current I_o is given by [34]

$$I_o = I_{ph}M \approx \frac{S}{2M_0} \left(\sqrt{1 + \frac{4M_0^2\eta P}{S}} - 1 \right), \quad (1)$$

where $I_{ph} = \eta P$ is the generated primary photocurrent, η is APD quantum efficiency, M is the APD multiplication factor, M_0 is the initial multiplication factor of APDs, and S is the designed parameter for achieving an expected gain M_0 .

2) APD Noise: The APD noise typically consists of shot noise, dark noise and thermal noise. Note that the central limit theorem implies that the amplitude distribution of shot noise approaches a Gaussian distribution for a large number of photons, which has also been verified and adopted in [35], [36], [37], and [38]. In the outdoor scenario under ambient radiation, the signal-dependent shot noise can be modeled as a Gaussian random variable with variance

$$\sigma_s^2 = 2qBFM^2I_{ph} = 2qBFM^2\eta P, \quad (2)$$

where q is the magnitude of the electronic charge, B the APD bandwidth, F is the excess noise factor, which can be approximated as $F \approx M^m$ [39], and m is the excess noise index in the range $[0.2, 0.5]$ depending on different materials of APD [39]. The variances of thermal noise and dark noise are expressed as [40],

$$\sigma_t^2 = \frac{4kTB}{R_{ld}} + \frac{4kTB}{R_{sh}}, \quad (3)$$

$$\sigma_d^2 = 2qBFM^2I_{bd} + 2qI_{sd}B, \quad (4)$$

where k is Boltzmann constant and T is the temperature in Kelvin; R_{ld} is the load resistor and R_{sh} is the internal shunt resistor; thermal noise term $\frac{4kTB}{R_{sh}}$ can be omitted due to large R_{sh} in the order of $100\text{K}\Omega$ to $1\text{G}\Omega$; I_{sd} is the surface dark current; I_{bd} is the bulk dark current to be multiplied through avalanche process, relatively smaller than the ambient light generated photocurrent and thus ignored [34].

Hence, the output current signal of APD follows a Gaussian distribution with mean μ and variance σ^2 . Define the mapping function between P and μ as $f_\mu(P)$, and the mapping function between P and σ^2 as $f_\sigma(P)$. Using Eq. (1)-(4), it yields

$$\begin{aligned} \mu &= f_\mu(P) \triangleq I_o + I_{sd} \\ &= \frac{S}{2M_0} \left(\sqrt{1 + \frac{4M_0^2\eta P}{S}} - 1 \right) + I_{sd}, \end{aligned} \quad (5)$$

$$\begin{aligned} \sigma^2 &= f_\sigma(P) \triangleq 2qBFM^2\eta P + 2qI_{sd}B + \frac{4kTB}{R_{ld}} \\ &= 2qB \left(\frac{S}{2M_0} \right)^{m+2} \left(\sqrt{1 + \frac{4M_0^2\eta P}{S}} - 1 \right)^{m+2} \\ &\quad \times (\eta P)^{-m-1} + 2qI_{sd}B + \frac{4kTB}{R_{ld}}. \end{aligned} \quad (6)$$

III. COVERT TRANSMISSION SCHEME

A. Covert Transmission Scheme Design

Considering strong ambient radiation, we propose a covert transmission scheme based on narrow signal waveband and

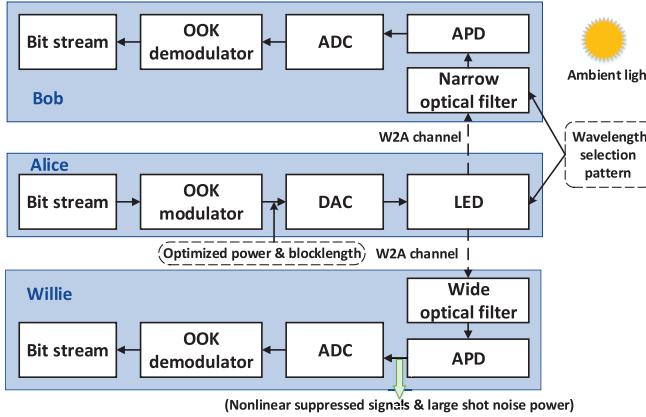


Fig. 2. The block diagram of covert W2A-OWC system.

nonlinear suppression for APDs, as shown in Fig. 2. In each transmission, Alice chooses a narrow wavelength range randomly, and modulates information symbols with optimized power and blocklength. Since Alice and Bob have common knowledge on the wavelength selection pattern, Bob adopts a filter on the corresponding wavelengths to receive signals. The signal-to-noise ratio (SNR) of Bob's received signals is improved by filtering out the ambient light outside the signal wavelength range. For Willie, due to the unknown signal wavelength, wide-passband filter is needed to ensure successful reception of transmitted signals, introducing strong ambient light. Strong ambient light not only shifts the APD operating point to the nonlinear region, but also increases the shot noise power, which significantly decreases the electrical SNR. Whereas, if Willie adopts a narrow-passband filter, there is a large probability that the passband wavelengths of Willie do not overlap with those of Alice.

B. Detection Performance at Willie

Willie detects whether Alice transmits signals to Bob based on hypothesis test, where \mathcal{H}_0 denotes the hypothesis of no transmission, and \mathcal{H}_1 denotes the hypothesis of transmission. In the i -th symbol duration, the observation of Willie is given by

$$y_w(i) = \begin{cases} f_\mu \left(\int_{\lambda_{w1}}^{\lambda_{w2}} b(\lambda) d\lambda \right) + n_e(i), & \mathcal{H}_0, \\ f_\mu \left(h_w x(i) + \int_{\lambda_{w1}}^{\lambda_{w2}} b(\lambda) d\lambda \right) + n_t(i), & \mathcal{H}_1, \end{cases} \quad (7)$$

where λ_{w1} and λ_{w2} are the minimum and maximum wavelengths of Willie's filter passband with $\Delta\lambda_w = \lambda_{w2} - \lambda_{w1}$; $b(\lambda)$ is the power spectral density of ambient light; $h_w x(i)$ represents the received i -th signal component in Willie's filter; $n_e(i)$ and $n_t(i)$ are AWGNs with variances $\sigma_e^2 = f_\sigma \left(\int_{\lambda_{w1}}^{\lambda_{w2}} b(\lambda) d\lambda \right)$ and $\sigma_t^2 = f_\sigma \left(h_w x(i) + \int_{\lambda_{w1}}^{\lambda_{w2}} b(\lambda) d\lambda \right)$, respectively; $f_\mu(\cdot)$ and $f_\sigma(\cdot)$ are defined in Eq. (5) and Eq. (6), respectively. Assume that the observations $\{y_w(i)\}$ are independent and identically distributed random variables following \mathbb{P}_0 under \mathcal{H}_0 or \mathbb{P}_1 under \mathcal{H}_1 .

Let $\mathbb{P}_0^{N_t}$ and $\mathbb{P}_1^{N_t}$ denote the probability distributions of Willie's observations $\{y_w(i)\}_{i=1}^{N_t}$ under hypotheses \mathcal{H}_0 and \mathcal{H}_1 , respectively. To achieve that the minimum sum of false alarm probability α and missed detection probability β satisfies the given constraint, i.e., $\alpha + \beta > 1 - \varepsilon$, the Kullback-Leibler (KL) divergence should satisfy the following constraint [11], [41],

$$\mathcal{D} \left(\mathbb{P}_0^{N_t} \parallel \mathbb{P}_1^{N_t} \right) = N_t \mathcal{D} (\mathbb{P}_0 \parallel \mathbb{P}_1) = N_t \int_{-\infty}^{\infty} p_0(y) \ln \frac{p_0(y)}{p_1(y)} dy < 2\varepsilon^2, \quad (8)$$

where $p_0(y)$ and $p_1(y)$ are probability density functions (PDFs) of \mathbb{P}_0 and \mathbb{P}_1 , respectively. Assume that Alice transmits On-Off Keying (OOK) signals since higher-order modulations are generally not available for weak signals. Then, \mathbb{P}_0 and \mathbb{P}_1 are given by

$$\mathbb{P}_0 = \mathcal{N} \left(\mu_0, \sigma_0^2 \right), \quad (9)$$

and

$$\mathbb{P}_1 = \frac{1}{2} \mathcal{N} \left(\mu_0, \sigma_0^2 \right) + \frac{1}{2} \mathcal{N} \left(\mu_1, \sigma_1^2 \right), \quad (10)$$

where $\mu_0 = f_\mu \left(\int_{\lambda_{w1}}^{\lambda_{w2}} b(\lambda) d\lambda \right)$ and $\sigma_0^2 = f_\sigma \left(\int_{\lambda_{w1}}^{\lambda_{w2}} b(\lambda) d\lambda \right)$; $\mu_1 = f_\mu \left(h_w C + \int_{\lambda_{w1}}^{\lambda_{w2}} b(\lambda) d\lambda \right)$ and $\sigma_1^2 = f_\sigma \left(h_w C + \int_{\lambda_{w1}}^{\lambda_{w2}} b(\lambda) d\lambda \right)$; $C = P_t \times \Delta\lambda$ is the optical power of OOK symbol "1" penetrating through Willie's filter; P_t is the power spectral density of the transmitted OOK symbol "1"; $\Delta\lambda$ is the overlapped waveband-width between the transmitted signals and Willie's filter passband.

Theorem 1: An upper bound on $\mathcal{D} (\mathbb{P}_0 \parallel \mathbb{P}_1)$ is given by

$$\begin{aligned} \mathcal{D} (\mathbb{P}_0 \parallel \mathbb{P}_1) &\leq \frac{e^r}{1+e^r} \ln \left(\frac{2e^r}{1+e^r} \right) + \frac{1}{1+e^r} \ln \left(\frac{2}{1+e^r} \right) + \frac{r}{1+e^r} \\ &\triangleq D(h_w, \Delta\lambda), \end{aligned} \quad (11)$$

$$\text{where } r = \ln \left(\frac{\sigma_1}{\sigma_0} \right) - \frac{1}{2} + \frac{\sigma_1^2 + (\mu_0 - \mu_1)^2}{2\sigma_1^2}.$$

Proof: See Appendix A.

To achieve covert communication, the covertness constraint in Eq. (8) for given fading gain h_w is transformed into

$$D(h_w, \Delta\lambda) < \frac{2\varepsilon^2}{N_t}. \quad (12)$$

Equations (9), (10), and the covertness constraint in Eq. (12) are utilized in the covert optimization at Alice in the subsequent sections.

IV. COVERT THROUGHPUT

In this section we study the covert throughput between Alice and Bob under given covertness constraints. Alice optimizes the power and blocklength of the transmitted signals to maximize the covert throughput.

The received signal of Bob in the i -th symbol duration is given by

$$y_b(i) \sim \begin{cases} \mathcal{N}\left(\mu'_0, \sigma'^2_0\right), & \text{for } X = 0, \\ \mathcal{N}\left(\mu'_1, \sigma'^2_1\right), & \text{for } X = 1, \end{cases} \quad (13)$$

where $\mu'_0 = f_\mu\left(\int_{\lambda_{b1}}^{\lambda_{b2}} b(\lambda) d\lambda\right)$ and $\sigma'^2_0 = f_\sigma\left(\int_{\lambda_{b1}}^{\lambda_{b2}} b(\lambda) d\lambda\right)$; $\mu'_1 = f_\mu\left(h_b A + \int_{\lambda_{b1}}^{\lambda_{b2}} b(\lambda) d\lambda\right)$ and $\sigma'^2_1 = f_\sigma\left(h_b A + \int_{\lambda_{b1}}^{\lambda_{b2}} b(\lambda) d\lambda\right)$; $A = P_t(\lambda_{b2} - \lambda_{b1})$ is the transmitted optical power of OOK symbol $X = 1$; λ_{b1} and λ_{b2} are the minimum and maximum wavelengths of transmitted optical signal with $\Delta\lambda_b = \lambda_{b2} - \lambda_{b1}$. The mutual information between Alice and Bob can be derived as

$$I(X; Y_b) = h(Y_b) - h(Y_b | X), \quad (14)$$

where

$$\begin{aligned} h(Y_b | X) &= \Pr(X = 0)h(Y_b | X = 0) + \Pr(X = 1)h(Y_b | X = 1) \\ &= \frac{1}{4} \log\left(2\pi e\sigma'^2_0\right) + \frac{1}{4} \log\left(2\pi e\sigma'^2_1\right), \end{aligned} \quad (15)$$

and

$$h(Y_b) = - \int p(y_b) \log(p(y_b)) dy_b, \quad (16)$$

where

$$p(y_b) = \frac{1}{2\sqrt{2\pi\sigma'^2_0}} e^{-\frac{(y_b-\mu'_0)^2}{2\sigma'^2_0}} + \frac{1}{2\sqrt{2\pi\sigma'^2_1}} e^{-\frac{(y_b-\mu'_1)^2}{2\sigma'^2_1}}. \quad (17)$$

Note that for quasi-static fading channels under finite blocklength, the fading channel can be considered constant over the transmission of a codeword, and the channel dispersion associated with finite blocklength is approximately zero [42], [43], [44]. Thus, the channel outage is dominantly caused by channel fading [9]. The transmission outage probability can be adopted to accurately describe the communication performance, given by

$$P_o = \Pr(I(X; Y_b) < F_R), \quad (18)$$

where F_R is a fixed rate threshold. Then the throughput from Alice to Bob is written as

$$I_c = N_t F_R (1 - P_o), \quad (19)$$

where N_t is the transmitted blocklength.

A. Case of Alice With Knowledge of Willie's Position and Optical Filter Passband-Width

Similar to the assumption in [41] that Bob and Willie are both legitimate users in a network, we assume that Alice has perfect knowledge about Willie's position as well as optical filter passband-width, and Willie is only hostile to Bob. We consider that instantaneous fading coefficients of Willie are unknown to Alice, and the channel distribution information (CDI) is available. Then, the probability of detection performance satisfying the covertness constraint is

constrained to be no less than $(1 - \delta)$. Assuming transmitted power spectral density $P_t \in [P_{min}, P_{max}]$ and blocklength $N_t \in [N_{min}, N_{max}]$, we seek to maximize throughput I_c with respect to P_t and N_t under covertness constraints. Mathematically, the covert optimization problem at Alice is formulated as

$$\max_{P_t, N_t} I_c = N_t F_R (1 - P_o) \quad (20a)$$

$$\text{s.t. } \Pr\left\{D(h_w, \Delta\lambda) < \frac{2\epsilon^2}{N_t}\right\} \geq 1 - \delta, \quad (20b)$$

$$P_{min} \leq P_t \leq P_{max}, \quad (20c)$$

$$N_{min} \leq N_t \leq N_{max}. \quad (20d)$$

In the case that Alice knows the optical filter passband-width of Willie, we consider the strictest constraint that Willie's detection performance satisfies given constraints as the passband of Willie's filter covers signal wavelengths, i.e., the overlapped waveband-width in Eq. (20b) is $\Delta\lambda = \Delta\lambda_b$.

Considering that it is difficult to directly solve optimization problem (20) under constraint (20b), Based on [45] and [46], we adopt the Markov inequality to transform the chance constraint (20b) into an expectation constraint as

$$\mathbb{E}[D(h_w, \Delta\lambda_b)] \leq \frac{2\epsilon^2\delta}{N_t}, \quad (21)$$

where $\mathbb{E}[D(h_w, \Delta\lambda_b)] = \int_0^\infty D(h_w, \Delta\lambda_b) f'(h_w) dh_w$, and $f'(h_w)$ is the PDF of channel gain following lognormal distribution, as shown in Sec. II-A. By replacing constraint (20b) by constraint (21), the covert optimization problem (20) can be rewritten as

$$\max_{P_t, N_t} I_c = N_t F_R (1 - P_o) \quad (22a)$$

$$\text{s.t. } \mathbb{E}[D(h_w, \Delta\lambda_b)] \leq \frac{2\epsilon^2\delta}{N_t}, \quad (22b)$$

$$P_{min} \leq P_t \leq P_{max}, \quad (22c)$$

$$N_{min} \leq N_t \leq N_{max}. \quad (22d)$$

Note that I_c increases with N_t . For a fixed P_t , constraint (22b) should hold with equality to obtain the maximum N_t . Consequently, we can solve the problem by performing a numerical search on P_t . The overall procedure for obtaining the optimal P_t and N_t is summarized as Algorithm 1.

B. Case of Alice Without Knowledge of Willie's Position and Optical Filter Passband-Width

Consider the scenario where Alice does not know Willie's position L_w or optical filter passband-width $\Delta\lambda_w$. If Willie adopts a wide-passband filter to improve the probability of successful reception, APD nonlinearity and shot noise caused by ambient light will degrade Willie's detection performance significantly. If Willie adopts a narrow-passband filter to alleviate the effects of ambient light, there is a large probability that Willie's passband is not overlapped with the signal wavelengths. Therefore, there exists an optimal passband-width for Willie.

Algorithm 1 Numerical Search for Problem (22)

Input: P_{\min} , P_{\max} , N_{\min} , N_{\max} , ε , δ , F_R , $f'(h_w)$,
 $f'(h_b)$, $b(\lambda)$, $\Delta\lambda_b$, $\Delta\lambda_w$

Output: P_t^* , N_t^*

```

1 for  $P_t = P_{\min} : P_{\max}$  do
2   Compute  $\mathbb{E}(D(h_w, \Delta\lambda_b))$ ;
3   Solve  $N_t = 2\varepsilon^2\delta/\mathbb{E}(D(h_w, \Delta\lambda_b))$ ;
4   if  $N_t \notin [N_{\min}, N_{\max}]$  then
5     | Delete the solution  $(P_t, N_t)$ ;
6   else
7     | Compute  $I_c$  for the solution  $(P_t, N_t)$ ;
8   end
9 end
10 Obtain the maximum  $I_c$ , and the corresponding optimal
    solution is  $(P_t^*, N_t^*)$ .
```

1) *Optimization Problem Formulation:* The overlapped waveband-width $\Delta\lambda$ varies with passband-width $\Delta\lambda_w$ of Willie's filter, and the distribution of channel gain h_w is relevant to Willie's position L_w . To limit the Willie's detection performance at the optimal $\Delta\lambda_w$ and L_w , the covert optimization problem in this case can be formulated as

$$\max_{P_t, N_t} I_c = N_t F_R (1 - P_o) \quad (23a)$$

$$\text{s.t. } \max_{L_w, \Delta\lambda_w} \mathbb{E}[D(h_w, \Delta\lambda)] \leq \frac{2\varepsilon^2\delta}{N_t}, \quad (23b)$$

$$P_{\min} \leq P_t \leq P_{\max}, \quad (23c)$$

$$N_{\min} \leq N_t \leq N_{\max}. \quad (23d)$$

2) *Average KL Divergence:* To explore the optimal Willie's passband-width $\Delta\lambda_w$, we define the average KL divergence as

$$\begin{aligned} & \mathbb{E}[D(h_w, \Delta\lambda)] \\ &= \int_0^\infty \left(\int_0^{\Delta\lambda_b} D(h_w, \Delta\lambda) f(\Delta\lambda) d\Delta\lambda \right) f'(h_w) dh_w, \end{aligned} \quad (24)$$

where $f(\Delta\lambda)$ is the PDF of overlapped waveband-width $\Delta\lambda$, which varies with Willie's filter passband-width $\Delta\lambda_w$.

We know that Bob's wavelength range $[\lambda_{b1}, \lambda_{b2}]$ and Willie's wavelength range $[\lambda_{w1}, \lambda_{w2}]$ are selected uniformly in the range $[V, V + L]$ in Sec. II-B. Denote the distance between λ_{b1} and V by L_1 , and the distance between λ_{b2} and $V + L$ by L_2 , as shown in Fig. 3. To calculate the average KL divergence in Eq. (24), we investigate the relationships among overlapped waveband-width $\Delta\lambda$, Bob's wavelength range $[\lambda_{b1}, \lambda_{b2}]$, Willie's wavelength range $[\lambda_{w1}, \lambda_{w2}]$ and candidate wavelength range $[V, V + L]$.

When $[\lambda_{w1}, \lambda_{w2}]$ shifts gradually from V to $V + L$, there are three types of curves for the initial phase in the variation of the overlapped waveband-width $\Delta\lambda$, corresponding to the following three conditions in the three columns in Fig. 4,

$$\begin{cases} \Delta\lambda_w \leq L_1 \text{ (cond. 1),} \\ L_1 < \Delta\lambda_w < L_1 + \Delta\lambda_b \text{ (cond. 2),} \\ \Delta\lambda_w \geq L_1 + \Delta\lambda_b \text{ (cond. 3).} \end{cases} \quad (25)$$

For condition 1, the overlapped waveband-width $\Delta\lambda$ initially increases from 0 to $\Delta\lambda_b$, as shown in cases {1, 2, 3} of Fig. 4. For condition 2, $\Delta\lambda$ initially increases from $\Delta\lambda_w - L_1$ to $\Delta\lambda_b$, as shown in cases {4, 5, 6}. For condition 3, $\Delta\lambda$ initially remains constant at $\Delta\lambda_b$, as shown in cases {7, 8, 9}.

Similarly, there are three types of curves for the final phase in the variation of the overlapped waveband-width $\Delta\lambda$ as $[\lambda_{w1}, \lambda_{w2}]$ shifts gradually from V to $V + L$, corresponding to the following three conditions in the three rows in Fig. 4,

$$\begin{cases} \Delta\lambda_w \leq L_2 \text{ (cond. 4),} \\ L_2 < \Delta\lambda_w < L_2 + \Delta\lambda_b \text{ (cond. 5),} \\ \Delta\lambda_w \geq L_2 + \Delta\lambda_b \text{ (cond. 6).} \end{cases} \quad (26)$$

For condition 4, $\Delta\lambda$ ultimately decreases from $\Delta\lambda_b$ to 0, as shown in cases {1, 4, 7}. For condition 5, $\Delta\lambda$ ultimately decreases from $\Delta\lambda_b$ to $\Delta\lambda_w - L_2$, as shown in cases {2, 5, 8}. For condition 6, $\Delta\lambda$ ultimately remains at $\Delta\lambda_b$, as shown in cases {3, 6, 9}.

Note that the above nine cases shown in Fig. 3 and Fig. 4 correspond to the combinations of conditions {1, 2, 3} and conditions {4, 5, 6}. Each case corresponds to a variation curve of $\Delta\lambda$ versus Willie's waveband positions. Under different relationships among the values of L , $\Delta\lambda_w$ and $\Delta\lambda_b$, some cases exist while others do not exist. The average KL divergence is given by

$$\mathbb{E}[D(h_w, \Delta\lambda)] = \sum_{\text{Case } i \in Z} P_{ci} \times \mathbb{E}[D(h_w, \Delta\lambda)]_{ci}, \quad (27)$$

where P_{ci} is the existence probability of case i , and $\mathbb{E}[D(h_w, \Delta\lambda)]_{ci}$ is the average KL divergence in case i , as derived in Appendix B. Z is the set of existence cases, given as follows:

(a) For $L \in [2\Delta\lambda_w + \Delta\lambda_b, \infty)$,

$$Z = \begin{cases} \text{Case 7 : } L_1 \in [0, \Delta\lambda_w - \Delta\lambda_b], \\ \text{Case 4 : } L_1 \in (\Delta\lambda_w - \Delta\lambda_b, \Delta\lambda_w), \\ \text{Case 1 : } L_1 \in [\Delta\lambda_w, L - \Delta\lambda_w - \Delta\lambda_b], \\ \text{Case 2 : } L_1 \in (L - \Delta\lambda_w - \Delta\lambda_b, L - \Delta\lambda_w), \\ \text{Case 3 : } L_1 \in [L - \Delta\lambda_w, L - \Delta\lambda_b] \end{cases}. \quad (28)$$

We can see that when the total bandwidth L meets the relationship with Willie's waveband-width $\Delta\lambda_w$ and Bob's waveband-width $\Delta\lambda_b$ as $L \in [2\Delta\lambda_w + \Delta\lambda_b, \infty)$, there are two variables affecting $\Delta\lambda$, namely, the selected waveband positions $[\lambda_{b1}, \lambda_{b2}]$ and $[\lambda_{w1}, \lambda_{w2}]$. Therefore, we traverse all Bob's and Willie's waveband positions in whole waveband $[V, V + L]$, i.e., distance L_1 between λ_{b1} and V varies from 0 to $L - \Delta\lambda_b$, and distance between λ_{w1} and V varies from 0 to $L - \Delta\lambda_w$, corresponding to cases {7, 4, 1, 2, 3}.

(b) For $L \in (2\Delta\lambda_w, 2\Delta\lambda_w + \Delta\lambda_b]$,

$$Z = \begin{cases} \text{Case 7 : } L_1 \in [0, \Delta\lambda_w - \Delta\lambda_b], \\ \text{Case 4 : } L_1 \in (\Delta\lambda_w - \Delta\lambda_b, L - \Delta\lambda_w - \Delta\lambda_b), \\ \text{Case 5 : } L_1 \in (L - \Delta\lambda_w - \Delta\lambda_b, \Delta\lambda_w), \\ \text{Case 2 : } L_1 \in [\Delta\lambda_w, L - \Delta\lambda_w], \\ \text{Case 3 : } L_1 \in [L - \Delta\lambda_w, L - \Delta\lambda_b] \end{cases}. \quad (29)$$

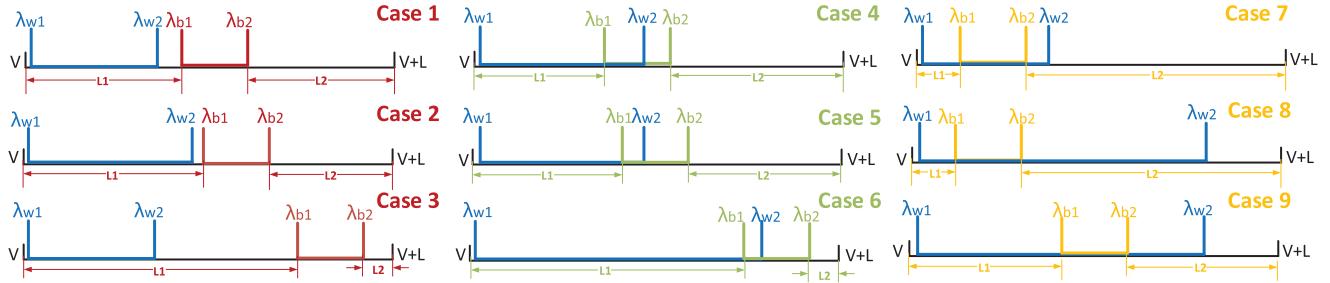
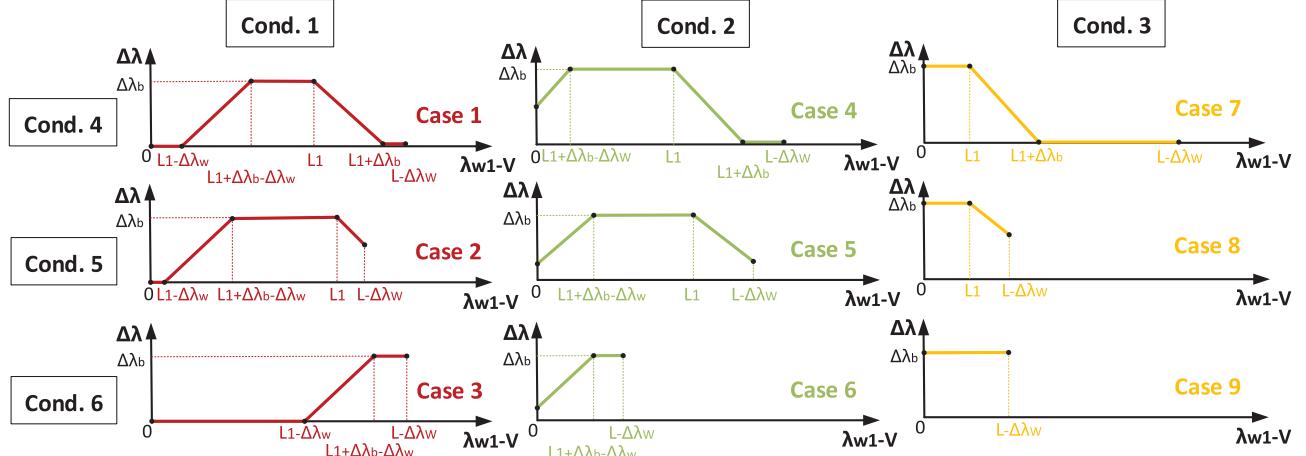


Fig. 3. Illustration of case 1 to case 9.

Fig. 4. Illustration of relationships between the overlapped waveband-width $\Delta\lambda$ and selected position λ_{w1} of Willie for case 1 to case 9.

(c) For $L \in (\Delta\lambda_w + \Delta\lambda_b, 2\Delta\lambda_w]$ and $\Delta\lambda_w + \Delta\lambda_b > 2\Delta\lambda_w - \Delta\lambda_b$,

$$Z = \left\{ \begin{array}{l} \text{Case 7 : } L_1 \in [0, L - \Delta\lambda_w - \Delta\lambda_b], \\ \text{Case 8 : } L_1 \in (L - \Delta\lambda_w - \Delta\lambda_b, \Delta\lambda_w - \Delta\lambda_b], \\ \text{Case 5 : } L_1 \in (\Delta\lambda_w - \Delta\lambda_b, L - \Delta\lambda_w), \\ \text{Case 6 : } L_1 \in [L - \Delta\lambda_w, \Delta\lambda_w], \\ \text{Case 3 : } L_1 \in [\Delta\lambda_w, L - \Delta\lambda_b] \end{array} \right\}. \quad (30)$$

(d) For $L \in (2\Delta\lambda_w - \Delta\lambda_b, \Delta\lambda_w + \Delta\lambda_b]$ and $\Delta\lambda_w + \Delta\lambda_b > 2\Delta\lambda_w - \Delta\lambda_b$,

$$Z = \left\{ \begin{array}{l} \text{Case 8 : } L_1 \in (0, \Delta\lambda_w - \Delta\lambda_b], \\ \text{Case 5 : } L_1 \in (\Delta\lambda_w - \Delta\lambda_b, L - \Delta\lambda_w), \\ \text{Case 6 : } L_1 \in [L - \Delta\lambda_w, L - \Delta\lambda_b] \end{array} \right\}. \quad (31)$$

(e) For $L \in [\Delta\lambda_w, 2\Delta\lambda_w - \Delta\lambda_b]$ and $\Delta\lambda_w + \Delta\lambda_b > 2\Delta\lambda_w - \Delta\lambda_b$,

$$Z = \left\{ \begin{array}{l} \text{Case 8 : } L_1 \in (0, L - \Delta\lambda_w), \\ \text{Case 9 : } L_1 \in [L - \Delta\lambda_w, \Delta\lambda_w - \Delta\lambda_b], \\ \text{Case 6 : } L_1 \in (\Delta\lambda_w - \Delta\lambda_b, L - \Delta\lambda_b) \end{array} \right\}. \quad (32)$$

(f) For $L \in (2\Delta\lambda_w - \Delta\lambda_b, 2\Delta\lambda_w]$ and $\Delta\lambda_w + \Delta\lambda_b \leq 2\Delta\lambda_w - \Delta\lambda_b$,

$$Z = \left\{ \begin{array}{l} \text{Case 7 : } L_1 \in [0, L - \Delta\lambda_w - \Delta\lambda_b], \\ \text{Case 8 : } L_1 \in (L - \Delta\lambda_w - \Delta\lambda_b, \Delta\lambda_w - \Delta\lambda_b], \\ \text{Case 5 : } L_1 \in (\Delta\lambda_w - \Delta\lambda_b, L - \Delta\lambda_w), \\ \text{Case 6 : } L_1 \in [L - \Delta\lambda_w, \Delta\lambda_w], \\ \text{Case 3 : } L_1 \in [\Delta\lambda_w, L - \Delta\lambda_b] \end{array} \right\}. \quad (33)$$

(g) For $L \in (\Delta\lambda_w + \Delta\lambda_b, 2\Delta\lambda_w - \Delta\lambda_b]$ and $\Delta\lambda_w + \Delta\lambda_b \leq 2\Delta\lambda_w - \Delta\lambda_b$,

$$Z = \left\{ \begin{array}{l} \text{Case 7 : } L_1 \in [0, L - \Delta\lambda_w - \Delta\lambda_b], \\ \text{Case 8 : } L_1 \in (L - \Delta\lambda_w - \Delta\lambda_b, L - \Delta\lambda_w), \\ \text{Case 9 : } L_1 \in [L - \Delta\lambda_w, \Delta\lambda_w - \Delta\lambda_b], \\ \text{Case 6 : } L_1 \in (\Delta\lambda_w - \Delta\lambda_b, \Delta\lambda_w), \\ \text{Case 3 : } L_1 \in [\Delta\lambda_w, L - \Delta\lambda_b] \end{array} \right\}. \quad (34)$$

(h) For $L \in [\Delta\lambda_w, \Delta\lambda_w + \Delta\lambda_b]$ and $\Delta\lambda_w + \Delta\lambda_b \leq 2\Delta\lambda_w - \Delta\lambda_b$,

$$Z = \left\{ \begin{array}{l} \text{Case 8 : } L_1 \in (0, L - \Delta\lambda_w), \\ \text{Case 9 : } L_1 \in [L - \Delta\lambda_w, \Delta\lambda_w - \Delta\lambda_b], \\ \text{Case 6 : } L_1 \in (\Delta\lambda_w - \Delta\lambda_b, L - \Delta\lambda_b) \end{array} \right\}. \quad (35)$$

3) *Covert Region:* If Alice does not know Willie's position, Alice can set a covert region where Willie's detection performance satisfies given covertness constraints. For example, we assume that the covert region corresponds to the half-space $\{t = (x, y, z)^T \mid \mathbf{a}^T t \geq z_w\}$ where $\mathbf{a} = (0, 0, 1)^T$, as shown in Fig. 5. Since Willie's average channel gains decrease with its height and horizontal distance, the optimal position for Willie's detection performance is $L_w^* = (0, 0, z_w)$ m, based on which we consider the covert constraint.

The overall procedure for obtaining the optimal P_t and N_t of optimization problem (23) is summarized as Algorithm 2.

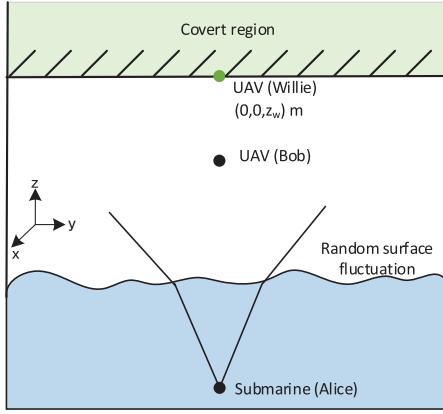


Fig. 5. Illustration of covert region.

Algorithm 2 Numerical Search for Problem (23)

Input: P_{\min} , P_{\max} , N_{\min} , N_{\max} , ε , δ , F_R , $f'(h_b)$, $b(\lambda)$, $\Delta\lambda_b$

Output: P_t^* , N_t^*

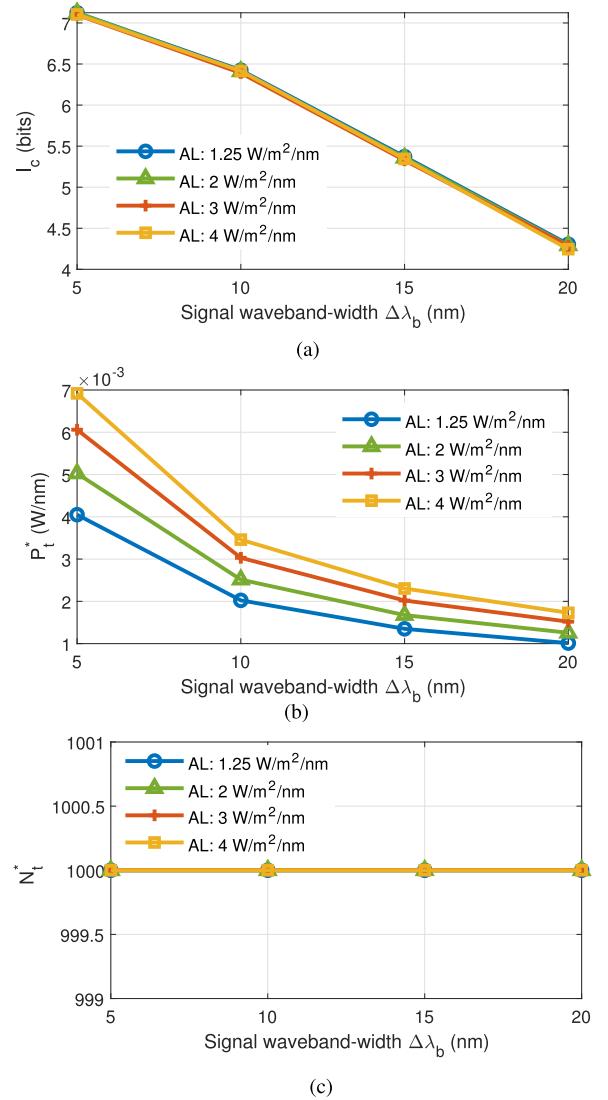
- 1 Establish a covert region and determine Willie's optimal position L_w^* and the PDF of channel gain $f'(h_w)$;
- 2 **for** $P_t = P_{\min} : P_{\max}$ **do**
- 3 **for** $\Delta\lambda_w = \Delta\lambda_b : L$ **do**
- 4 | Compute $\mathbb{E}[D(h_w, \Delta\lambda)]$;
- 5 **end**
- 6 Obtain the maximum $\mathbb{E}^*[D(h_w, \Delta\lambda)]$ and the optimal $\Delta\lambda_w^*$;
- 7 Solve $N_t = 2\varepsilon^2\delta/\mathbb{E}^*[D(h_w, \Delta\lambda)]$;
- 8 **if** $N_t \notin [N_{\min}, N_{\max}]$ **then**
- 9 | Delete the solution (P_t, N_t) ;
- 10 **else**
- 11 | Compute I_c for the solution (P_t, N_t) ;
- 12 **end**
- 13 **end**
- 14 Obtain the maximum I_c , and the corresponding optimal solution is (P_t^*, N_t^*) .

V. NUMERICAL RESULTS

In this section, we present numerical results and study communication performance of W2A-OWC systems under given covertness constraints. The APD parameters are set to typical values, i.e., $S = 1.1538$ A, $M_0 = 100$, $\eta = 1$, $B = 500$ MHz, $T = 300$ K, $I_{sd} = 10^{-9}$ A, $R_{ld} = 50$ Ω and $m = 0.3$, the same as those in [34]. Assume that Alice and Bob locate at $(0, 0, -2)$ m and $(0, 0, 1)$ m, respectively. Wave slope follows truncated logistic distribution with parameters $u_w = 0$ and $\sigma_w = 5$ [29]. The bounds of power spectral density are $P_{\max} = 0.01$ W/nm and $P_{\min} = 10^{-7}$ W/nm. The blocklength bounds are $N_{\max} = 1000$ and $N_{\min} = 500$. Other parameters are set as $\varepsilon = 0.2$, $\delta = 0.4$, and $F_R = 7.2 \times 10^{-3}$.

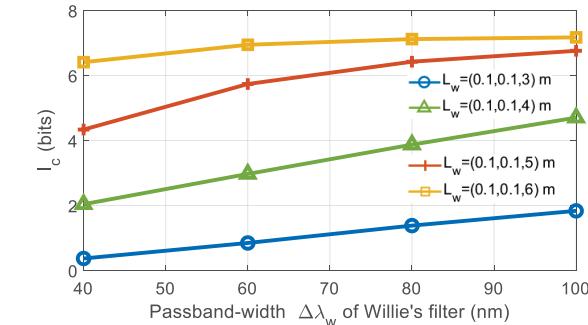
A. Case of Alice With Knowledge of Willie's Position and Optical Filter Passband-Width

Assume that Willie is located at $(0.1, 0.1, 5)$ m, and passband-width of Willie's filter is $\Delta\lambda_w = 80$ nm. In Fig. 6(a),

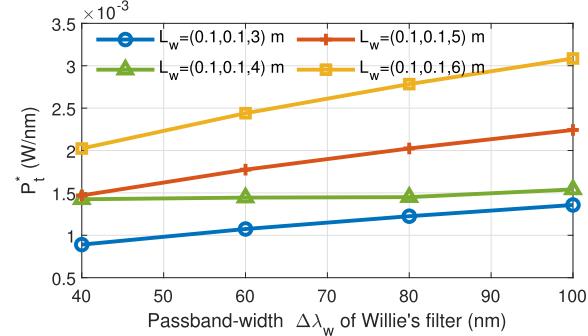
Fig. 6. (a) Covert throughput between Alice and Bob, (b) transmitted power spectral density P_t^* , (c) blocklength N_t^* versus the signal waveband-width $\Delta\lambda_b$ under different ambient light power.

the covert throughput I_c between Alice and Bob decreases as signal waveband-width $\Delta\lambda_b$ increases. The reason is that a wide-passband filter introduces more ambient light, resulting in higher noise power at Bob. Given $\Delta\lambda_b$, the covert throughputs under different ambient light powers (labeled as "AL") are close, since both the noise power at Bob and optimized transmitted signal power increase with the ambient light power. In Fig. 6(b), the optimized transmitted power spectral density P_t^* decreases with $\Delta\lambda_b$, since the total power $C = P_t\Delta\lambda = P_t\Delta\lambda_b$ satisfying covertness constraints remains the same. Moreover, P_t^* increases steadily as the ambient light power increases, which indicates that stronger ambient light improves the transmitted signal power under covertness constraints. From Fig. 6(c), the optimal blocklength N_t^* is the maximum blocklength N_{\max} , which indicates that increasing N_t has a greater impact on improving throughput.

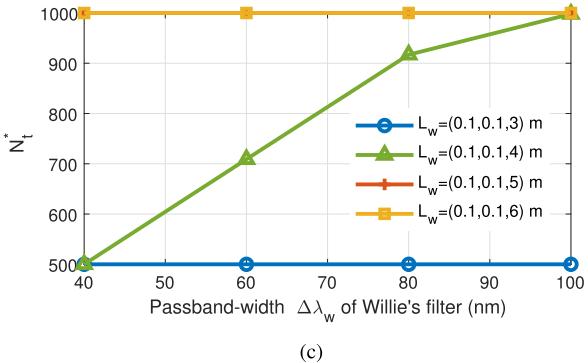
Assume $\Delta\lambda_b = 10$ nm, and the ambient light power is 1.25 W/m²/nm. In Fig. 7(a), the covert throughput increases with $\Delta\lambda_w$ and Willie's height, since wider-passband filters



(a)



(b)



(c)

Fig. 7. (a) Covert throughput between Alice and Bob, (b) transmitted power spectral density P_t^* , (c) blocklength N_t^* versus the passband-width $\Delta\lambda_w$ of Willie's filter under different Willie's positions.

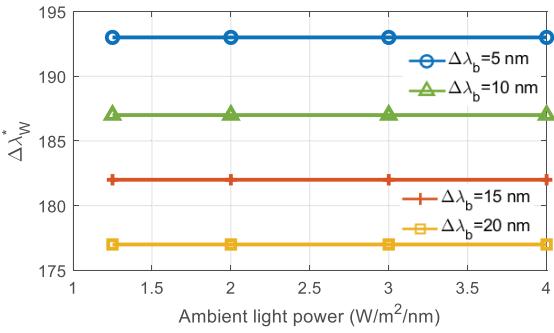
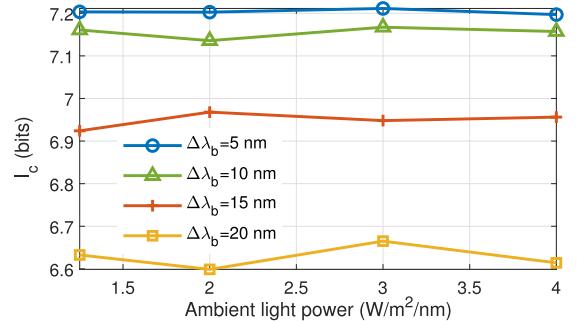
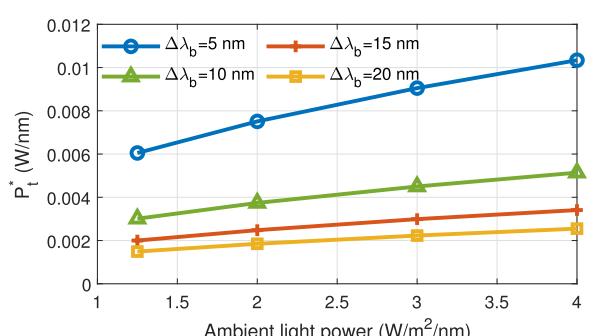


Fig. 8. The optimal Willie's passband-width $\Delta\lambda_w^*$ versus ambient light power under different signal waveband-width $\Delta\lambda_b$.

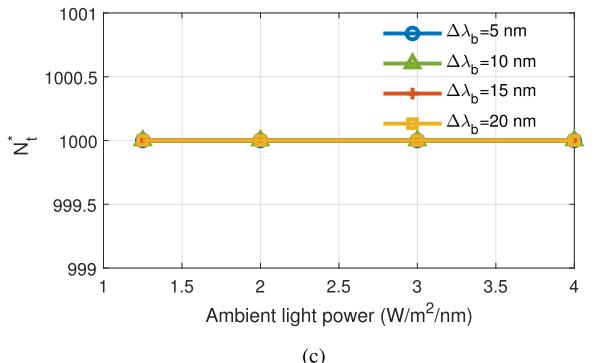
at Willie introduce more ambient light, and higher locations correspond to smaller channel gains, which both result in worse detection performance at Willie and are beneficial for signal-hiding. In Fig. 7(b), P_t^* increases with $\Delta\lambda_w$. In Fig. 7(c), N_t^* remains to N_{min} with Willie at (0.1, 0.1, 3) m; N_t^* increases from N_{min} to N_{max} as $\Delta\lambda_w$ increases from



(a)



(b)



(c)

Fig. 9. (a) Covert throughput between Alice and Bob, (b) transmitted power spectral density P_t^* , (c) blocklength N_t^* versus the ambient light power under different signal waveband-width $\Delta\lambda_b$.

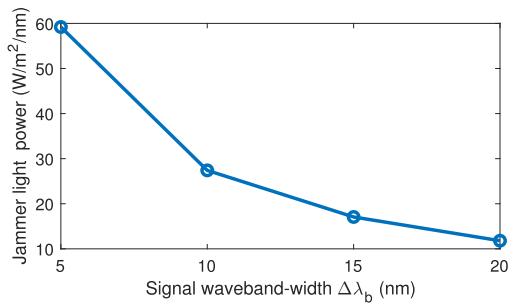


Fig. 10. The jammer light power versus signal waveband-width.

40 nm to 100 nm with Willie at (0.1, 0.1, 4) m; N_t^* remains to N_{max} with Willie at (0.1, 0.1, 5) m and (0.1, 0.1, 6) m. It indicates that the optimized N_t^* and P_t^* become larger at higher positions of Willie.

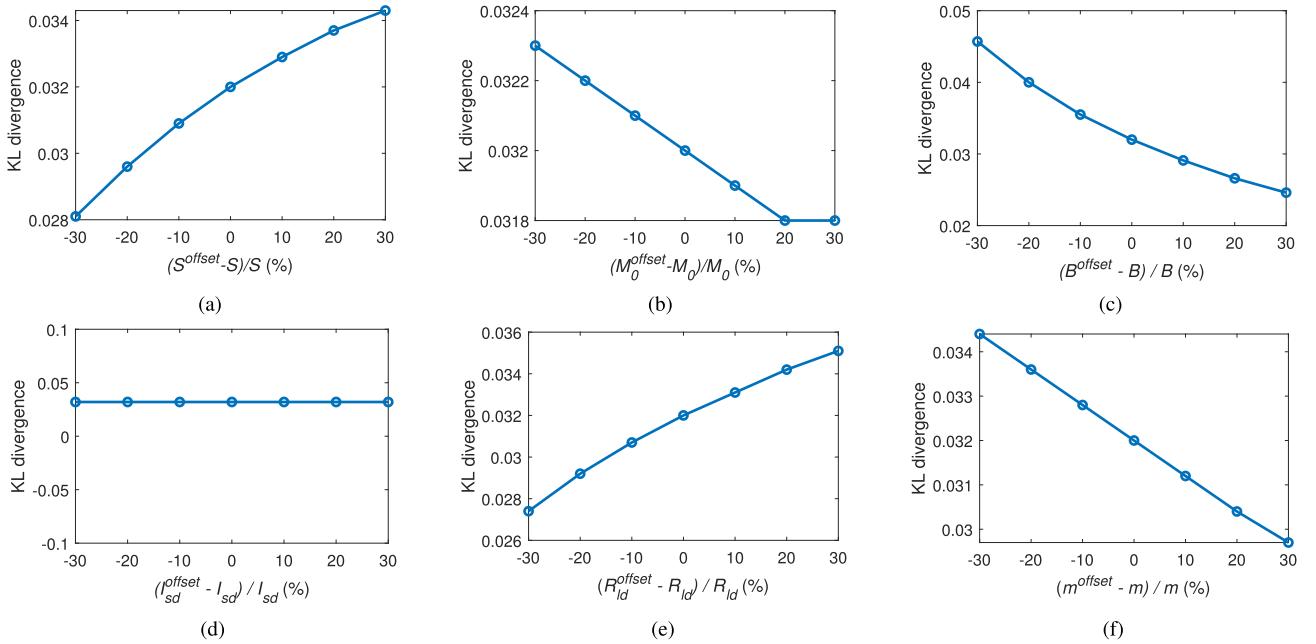


Fig. 11. The KL divergence at Willie versus parameter offset of (a) S , (b) M_0 , (c) B , (d) I_{sd} , (e) R_{ld} , and (f) m .

B. Case of Alice Without Knowledge of Willie's Position and Optical Filter Passband-Width

The covert region is set to be $\{t = (x, y, z)^T \mid \mathbf{a}^T t \geq 5\}$, and the bandwidth range $[V, V + L] = [400, 600]$ nm. Figure 8 shows the optimal Willie's bandwidth $\Delta\lambda_w^*$ versus ambient light power under different signal waveband-width $\Delta\lambda_b$. It is seen that the optimal Willie's bandwidth $\Delta\lambda_w^*$ remains the same when the ambient light power varies from 1.25 to 4 W/m²/nm, and $\Delta\lambda_w^*$ decreases from 193 nm to 177 nm as $\Delta\lambda_b$ increases from 5 nm to 20 nm.

In Fig. 9(a), as ambient light power increases from 1.25 to 4 W/m²/nm, the covert throughput I_c changes slightly since both the ambient light noise and optimized transmitted signal power increase, which is similar to the result in Fig. 6(a). Larger signal waveband-width $\Delta\lambda_b$ introduces stronger ambient light, leading to lower I_c . In Fig. 9(b), P_t^* decreases with $\Delta\lambda_b$, which is similar to the result in Fig. 6(b). As shown in Fig. 9(c), the optimized blocklength N_t^* remains to N_{\max} under different ambient light powers and $\Delta\lambda_b$.

C. Comparison With Covert Schemes Based on Jammer

Consider a covert scheme where a jammer transmits strong light to interfere with Willie. Alice chooses a fixed waveband to transmit signals without random selection. Assume that Willie is located at (0.1, 0.1, 5) m and the ambient light power is 1.25 W/m²/nm. The transmitted signal power and blocklength are set as the optimized results in Fig. 6 to ensure that Bob's throughput remains unchanged. Fig. 10 shows the required jammer light power versus the signal waveband-width under the same covert constraint (22b). It is seen that the required jammer light power reaching Willie decreases from 59.2 W/m²/nm to 11.8 W/m²/nm as the signal waveband-width increases from 5 nm to 20 nm. However, larger signal waveband-width leads to lower covert throughput as shown in

Fig. 6(a), since stronger ambient light is introduced at Bob. Compared with the proposed scheme, the jammer-based covert scheme requires more power and additional transmitters to satisfy the same covertness constraint.

D. Scheme Robustness to APD Parameters of Willie

Assume $P_t^* = 2.02 \times 10^{-3}$ W/nm, $N_t^* = 1000$, $\Delta\lambda_b = 10$ nm, $\Delta\lambda_w = 80$ nm, and $L_w = (0.1, 0.1, 5)$ m. The ambient light power is 1.25 W/m²/nm and the constraint on KL divergence is $N_t \mathbb{E}[D(h_w, \Delta\lambda_b)] \leq 2\varepsilon^2\delta = 0.032$ as shown in Section V-A. To investigate the robustness of the covert scheme to APD parameters of Willie, we vary the parameters within the range from 70% to 130% of the original value, and calculate the KL divergence at Willie as shown in Fig. 11. According to Eq. (5), APD nonlinearity is determined by the two parameters S and M_0 . It is seen from Figs. 11(a) and 11(b) that as S and M_0 vary, the maximum increase of KL divergence relative to the constraint value (0.032) is 0.0023, which is relatively small. According to Eq. (6), parameters B , I_{sd} , R_{ld} , m are related to noise variance. It is seen from Figs. 11(c)-11(f) that the offset of B has the greatest impact on KL divergence where the maximum increase of KL divergence is 0.0137 under the offset -30% in Fig. 11(c), but smaller bandwidths can cause intersymbol interference. In Fig. 11(d), the offset of I_{sd} has almost no effect on KL divergence.

E. Scheme Covertness to Multiple APDs of Willie

Assuming Willie has a parallel bank of APDs covering different bands, Fig. 12 shows the KL divergence at Willie versus the number of APDs in Willie's APD bank. For example, if Willie has 10 APDs covering bandwidth range $[V, V + L] = [400, 600]$ nm, each APD is equipped with an optical filter with passband-width of 20 nm. We consider the

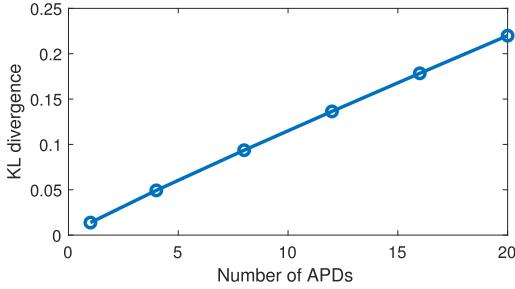


Fig. 12. The KL divergence at Willie versus the number of APDs.

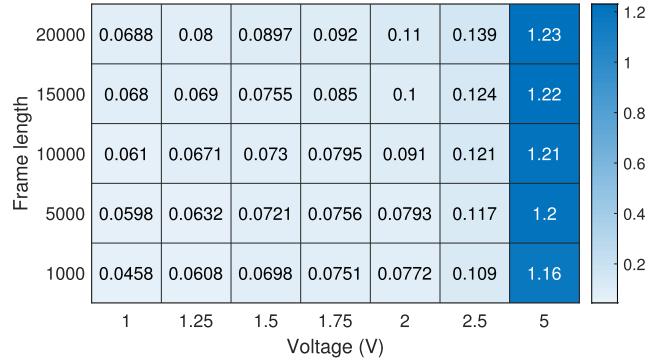


Fig. 13. The experimental system of W2A-OWC.

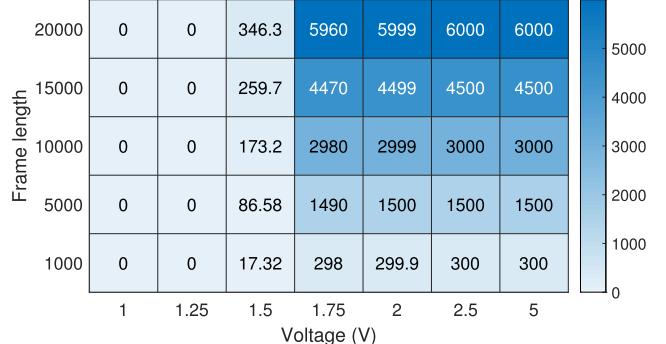
worst scenario for Alice where the entire signal waveband is within the passband of one APD at Willie. P_t^* , N_t^* , $\Delta\lambda_b$, L_w , and the ambient light power are the same as those in Section V-D. It is seen from Fig. 12 that the KL divergence at Willie increases approximately linearly from 0.0138 to 0.2199 as the number of APDs increases from 1 to 20. It indicates that more APDs and narrow-band optical filters can improve the detection performance of Willie. Whereas, if we further reduce signal waveband-width and increase the wavelength selection range, Willie needs more APDs and narrow-band optical filters to capture the signal and avoid the effect of ambient radiation, which leads to a dramatic increase in hardware complexity, device size, and further the probability of Willie being detected.

VI. EXPERIMENTAL RESULTS

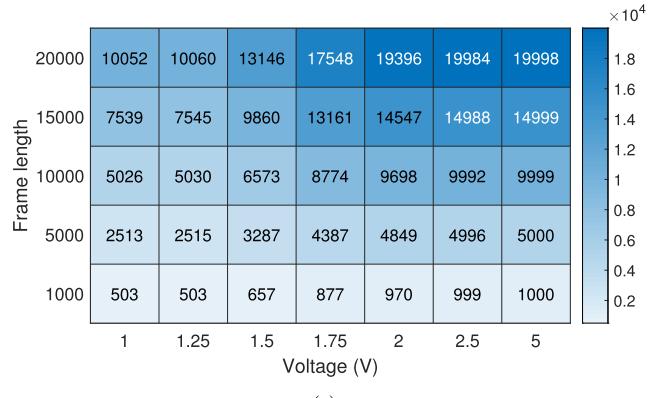
To verify the effectiveness of the proposed covert scheme, a W2A-OWC system is established under ambient light, as shown in Fig. 13. A green LED (LED1, Cree XP-E, 520~535 nm) is placed under depth 0.17 m as the transmitter Alice with coordinate (0, 0, -0.17) m. An APD (APD1, Hamamatsu, S5344) with a filter (Semrock, 513~533 nm) is placed directly above Alice at (0, 0, 0.93) m as the legitimate receiver Bob. Another APD (APD2, Hamamatsu, S5344) acting as the adversary Willie is placed at (0.39, 0,



(a)



(b)



(c)

Fig. 14. (a) The KL divergence at Willie, (b) Bob's throughput, and (c) the number of correctly received bits versus voltages of OOK symbol “1” and frame length under weak ambient light.

1.87) m. A white LED (LED2, Kcent KC-SLCOB15) is deployed as the ambient light source. An arbitrary waveform generator (AWG, keysight 33600A) is adopted to generate OOK signals of frequency 200 KHz for driving LED1. The received signal is sampled by a digital storage oscilloscope (DSO, Agilent MSO-X 6004A) with sampling rate 20 MHz for offline signal processing.

For the case that we adjust the ambient light to be relatively weak, the ambient light power at Willie is 0.621 W/m^2 and the ambient light power after passing through the filter at Bob is 0.039 W/m^2 . For different voltages of OOK symbol “1” and frame length, we calculate the KL divergence between PDFs with information symbols and without information symbols at Willie, Bob’s throughput ($F_R = 0.3$), and the number of correctly received bits, as shown in Fig. 14. It is seen that the

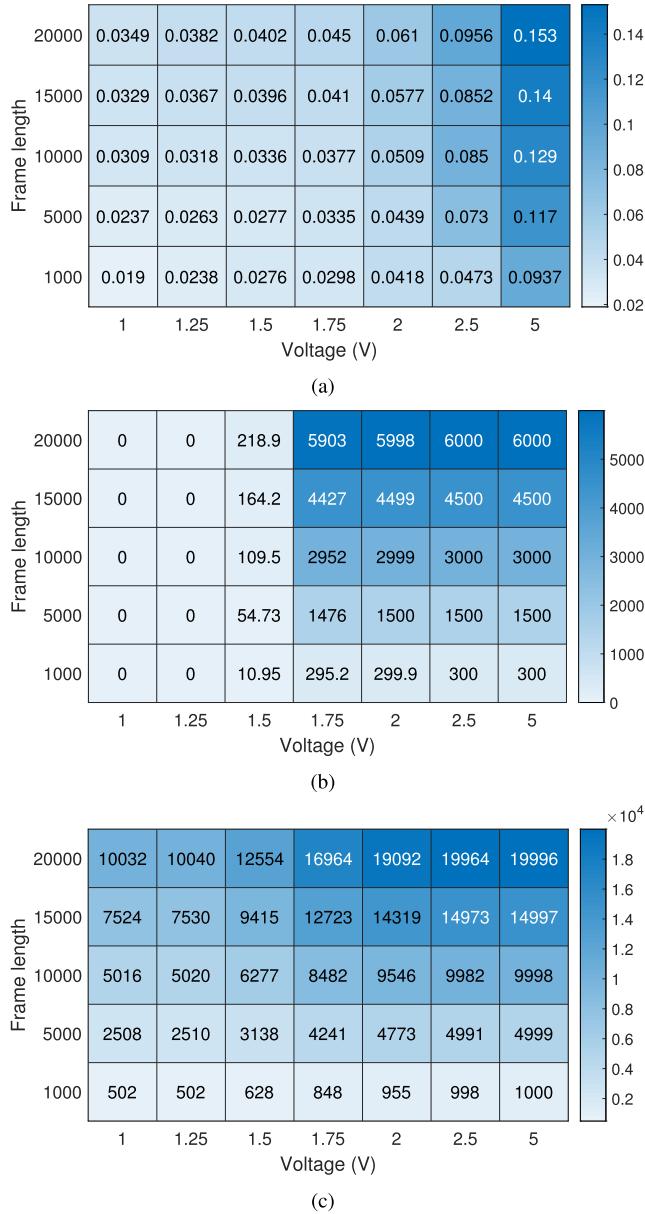


Fig. 15. (a) The KL divergence at Willie, (b) Bob’s throughput, and (c) the number of correctly received bits versus voltages of OOK symbol “1” and frame length under strong ambient light.

KL divergence, the throughput and the number of correctly received bits increase with the signal voltage and frame length, and there exists a tradeoff between the voltage and frame length to satisfy a given covertness constraint.

When the ambient light is adjusted to be strong, the ambient light power at Willie is 5.784 W/m^2 and the ambient light power after passing through filter at Bob is 0.066 W/m^2 . Figure 15 shows the KL divergence, the throughput and the number of correctly received bits versus signal voltage and frame length. Compared with Fig. 14(a), the KL divergence in Fig. 15(a) decreases significantly under the same voltage and frame length, which indicates that stronger ambient light can enhance the signal covertness. The throughput in Fig. 15(b) only shows a slight decrease compared to that in Fig. 14(b), since Bob’s narrow-passband filter blocks the ambient light effectively.

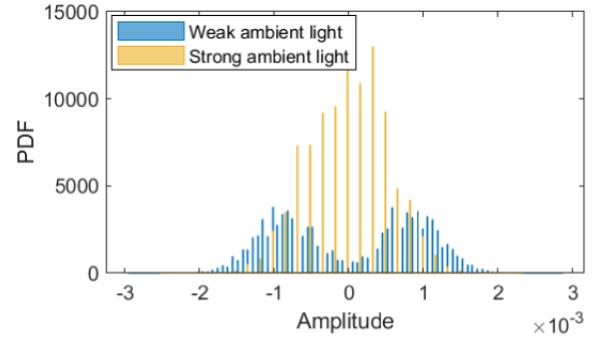


Fig. 16. The PDFs of received signals for voltage 5 V and frame length 1000 under weak and strong ambient light.

Figure 16 shows the PDF of received signal samples at Willie when the voltage of symbol “1” is 5 V and frame length is 1000. There are clearly two separate Gaussian distributions of signal samples under weak ambient light, but only one peak under strong ambient light, which is closer to the PDF without information symbols. It indicates that strong ambient light can undermine Willie’s detection performance effectively.

VII. CONCLUSION

In this work, we have investigated a covert transmission scheme for W2A-OWC systems. The proposed scheme takes advantage of secret selection of narrow-waveband signal, APD characteristics and ambient radiation to achieve signal hiding. The detection performance of Willie has been analyzed under APD nonlinearity, shot noise and ambient light. The maximum throughput of the proposed covert scheme has been explored by optimizing the power and blocklength of transmitted signals. For a more general case where Willie’s position and optical filter passband-width are unknown to Alice, the optimal passband-width of Willie’s filter has been calculated by taking into account both overlapped waveband and detection performance, and a covert region has been established to limit Willie’s detection range. Numerical and experimental results have shown that strong ambient radiation and narrow signal waveband can enhance the signal covertness significantly. It remains for our future work to investigate the case where Willie adopts multiple APDs covering different bands.

APPENDIX A PROOF OF THEOREM 1

The KL divergence between two Gaussian mixture models $\sum_{i=1}^I w_i f_i$ and $\sum_{i=1}^I \tilde{w}_i \tilde{f}_i$ is upper-bounded by [47], and [48]

$$\begin{aligned} & \mathcal{D}\left(\sum_{i=1}^I w_i f_i \middle\| \sum_{i=1}^I \tilde{w}_i \tilde{f}_i\right) \\ & \leq \sum_{i=1}^I w_i \ln\left(\frac{w_i}{\tilde{w}_i}\right) + \sum_{i=1}^I w_i \mathcal{D}(f_i \parallel \tilde{f}_i). \end{aligned} \quad (36)$$

We utilize Eq. (36) to obtain an upper bound of $\mathcal{D}(\mathbb{P}_0 \parallel \mathbb{P}_1)$ for $\mathbb{P}_0 = \mathcal{N}(\mu_0, \sigma_0^2)$ and $\mathbb{P}_1 = \frac{1}{2}\mathcal{N}(\mu_0, \sigma_0^2) + \frac{1}{2}\mathcal{N}(\mu_1, \sigma_1^2)$. Specifically, we have

$$\begin{cases} w_1 = a, f_1 = \mathcal{N}(\mu_0, \sigma_0^2) \\ w_2 = 1 - a, f_2 = \mathcal{N}(\mu_0, \sigma_0^2), \end{cases} \quad (37)$$

and

$$\begin{cases} \tilde{w}_1 = 0.5, \tilde{f}_1 = \mathcal{N}\left(\mu_0, \sigma_0^2\right) \\ \tilde{w}_2 = 0.5, \tilde{f}_2 = \mathcal{N}\left(\mu_1, \sigma_1^2\right), \end{cases} \quad (38)$$

where $a \in [0, 1]$. Substituting Eq. (37) and Eq. (38) into Eq. (36), we have

$$\begin{aligned} \mathcal{D}(\mathbb{P}_0 \parallel \mathbb{P}_1) &\leq \left(a \ln(2a) + (1-a) \ln(2-2a) + \right. \\ &\quad \left. a \mathcal{D}(\mathcal{N}(\mu_0, \sigma_0^2) \parallel \mathcal{N}(\mu_0, \sigma_0^2)) + \right. \\ &\quad \left. (1-a) \mathcal{D}(\mathcal{N}(\mu_0, \sigma_0^2) \parallel \mathcal{N}(\mu_1, \sigma_1^2)) \right) \\ &= a \ln(2a) + (1-a) \\ &\quad \left(\ln\left(\frac{2\sigma_1(1-a)}{\sigma_0}\right) + \frac{\sigma_0^2 - \sigma_1^2 + (\mu_0 - \mu_1)^2}{2\sigma_1^2} \right) \\ &= \mathcal{D}^U. \end{aligned} \quad (39)$$

The second derivative of the upper bound in Eq. (39) with respect to a is

$$\frac{\partial^2 \mathcal{D}^U}{\partial a^2} = \frac{1}{a} + \frac{1}{1-a} > 0. \quad (40)$$

Thus, there exists the minimal \mathcal{D}^U where the corresponding a satisfies

$$\begin{aligned} \frac{\partial \mathcal{D}^U}{\partial a} &= \ln(2a) - \ln(2-2a) - \left(\ln\left(\frac{\sigma_1}{\sigma_0}\right) - \frac{1}{2} + \right. \\ &\quad \left. \frac{\sigma_0^2 + (\mu_0 - \mu_1)^2}{2\sigma_1^2} \right) \\ &= 0. \end{aligned} \quad (41)$$

Equation (41) leads to

$$a = \frac{e^{\left(\ln\left(\frac{\sigma_1}{\sigma_0}\right) - \frac{1}{2} + \frac{\sigma_0^2 + (\mu_0 - \mu_1)^2}{2\sigma_1^2}\right)}}{1 + e^{\left(\ln\left(\frac{\sigma_1}{\sigma_0}\right) - \frac{1}{2} + \frac{\sigma_0^2 + (\mu_0 - \mu_1)^2}{2\sigma_1^2}\right)}}. \quad (42)$$

Substituting Eq. (42) into Eq. (39), we have

$$\begin{aligned} \mathcal{D}(\mathbb{P}_0 \parallel \mathbb{P}_1) &\leq \frac{e^r}{1+e^r} \ln\left(\frac{2e^r}{1+e^r}\right) + \frac{1}{1+e^r} \ln\left(\frac{2}{1+e^r}\right) + \frac{r}{1+e^r}, \end{aligned} \quad (43)$$

$$\text{where } r = \ln\left(\frac{\sigma_1}{\sigma_0}\right) - \frac{1}{2} + \frac{\sigma_0^2 + (\mu_0 - \mu_1)^2}{2\sigma_1^2}.$$

Figure 17 compares the derived upper bound (43) with the Monte Carlo simulation result of $\mathcal{D}(\mathbb{P}_0 \parallel \mathbb{P}_1)$ under different transmitted power spectral densities. Monte Carlo results are obtained by generating 1×10^7 samples from \mathbb{P}_0 . It is seen that the derived upper bound is closer to the Monte Carlo simulation result under smaller transmitted power spectral density.

APPENDIX B AVERAGE KL DIVERGENCE IN EACH CASE

A. Case 1

When condition 1 and condition 4 are satisfied, the relationship between the overlapped waveband-width $\Delta\lambda$ and

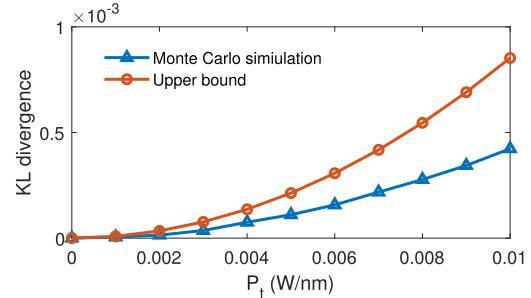


Fig. 17. The comparison of the Monte Carlo simulation result and derived upper bound on $\mathcal{D}(\mathbb{P}_0 \parallel \mathbb{P}_1)$ under different transmitted power spectral densities P_t .

the value of $\lambda_{w1} - V$ is shown in Fig. 4. The selection range of L_1 is given by

$$L_1 \in [L_1^{n_1}, L_1^{m_1}] \triangleq [\Delta\lambda_w, L - \Delta\lambda_w - \Delta\lambda_b], \quad (44)$$

for $L \in [2\Delta\lambda_w + \Delta\lambda_b, \infty)$. Thus, the existence probability of case 1 is $P_{c1} = (L_1^{m_1} - L_1^{n_1}) / (L - \Delta\lambda_b)$. In this case, the average KL divergence is written as

$$\mathbb{E}[D(h_w, \Delta\lambda)]_{c1} = \int_0^\infty D_1 \times f'(h_w) dh_w, \quad (45)$$

where

$$D_1 = \frac{\Delta\lambda_w - \Delta\lambda_b}{L - \Delta\lambda_w} D(h_w, \Delta\lambda_b) + \int_0^{\Delta\lambda_b} \frac{2}{L - \Delta\lambda_w} D(h_w, \Delta x) d\Delta x. \quad (46)$$

B. Case 2

When condition 1 and condition 5 are satisfied, the selection range of L_1 is given by

$$L_1 \in \begin{cases} (L_1^{n_2}, L_1^{m_2}) \triangleq (L - \Delta\lambda_w - \Delta\lambda_b, L - \Delta\lambda_w), \\ \text{for } L \in (2\Delta\lambda_w + \Delta\lambda_b, \infty); \\ [L_1^{n_2}, L_1^{m_2}] \triangleq [\Delta\lambda_w, L - \Delta\lambda_w], \\ \text{for } L \in [2\Delta\lambda_w, 2\Delta\lambda_w + \Delta\lambda_b]. \end{cases} \quad (47)$$

Thus, $P_{c2} = (L_1^{m_2} - L_1^{n_2}) / (L - \Delta\lambda_b)$. In this case, the average KL divergence is

$$\mathbb{E}[D(h_w, \Delta\lambda)]_{c2} = \int_0^\infty D_2 \times f'(h_w) dh_w, \quad (48)$$

where

$$\begin{aligned} D_2 &= \frac{\Delta\lambda_w - \Delta\lambda_b}{L - \Delta\lambda_w} D(h_w, \Delta\lambda_b) \\ &+ \int_0^{\Delta\lambda_b} \frac{1}{L - \Delta\lambda_w} D(h_w, \Delta x) d\Delta x \\ &+ \int_{L_1^{n_2}}^{L_1^{m_2}} \int_{L_1 + \Delta\lambda_w + \Delta\lambda_b - L}^{\Delta\lambda_b} \left\{ \frac{1}{L - \Delta\lambda_w} D(h_w, \Delta x) \times \frac{1}{L_1^{m_2} - L_1^{n_2}} \right\} d\Delta x dL_1. \end{aligned} \quad (49)$$

C. Case 3

When condition 1 and condition 6 are satisfied, the selection range of L_1 is given by

$$L_1 \in \begin{cases} [L_1^{n_3}, L_1^{m_3}] \triangleq [L - \Delta\lambda_w, L - \Delta\lambda_b], \\ \text{for } L \in (2\Delta\lambda_w, \infty); \\ [L_1^{n_3}, L_1^{m_3}] \triangleq [\Delta\lambda_w, L - \Delta\lambda_b], \\ \text{for } L \in [\Delta\lambda_w + \Delta\lambda_b, 2\Delta\lambda_w]. \end{cases} \quad (50)$$

Thus, $P_{c3} = (L_1^{m_3} - L_1^{n_3}) / (L - \Delta\lambda_b)$. In this case, the average KL divergence is written as

$$\mathbb{E}[D(h_w, \Delta\lambda)]_{c3} = \int_0^\infty D_3 \times f'(h_w) dh_w, \quad (51)$$

where

$$D_3 = \int_{L_1^{n_3}}^{L_1^{m_3}} \frac{L - L_1 - \Delta\lambda_b}{L - \Delta\lambda_w} D(h_w, \Delta\lambda_b) \frac{1}{L_1^{m_3} - L_1^{n_3}} dL_1 + \int_0^{\Delta\lambda_b} \frac{1}{L - \Delta\lambda_w} D(h_w, \Delta x) d\Delta x. \quad (52)$$

D. Case 4

When condition 2 and condition 4 are satisfied, the selection range of L_1 is given by

$$L_1 \in \begin{cases} (L_1^{n_4}, L_1^{m_4}) \triangleq (\Delta\lambda_w - \Delta\lambda_b, \Delta\lambda_w), \\ \text{for } L \in (2\Delta\lambda_w + \Delta\lambda_b, \infty); \\ (L_1^{n_4}, L_1^{m_4}] \triangleq (\Delta\lambda_w - \Delta\lambda_b, L - \Delta\lambda_w - \Delta\lambda_b], \\ \text{for } L \in [2\Delta\lambda_w, 2\Delta\lambda_w + \Delta\lambda_b]. \end{cases} \quad (53)$$

Thus, $P_{c4} = (L_1^{m_4} - L_1^{n_4}) / (L - \Delta\lambda_b)$. In this case, the average KL divergence is written as

$$\mathbb{E}[D(h_w, \Delta\lambda)]_{c4} = \int_0^\infty D_4 \times f'(h_w) dh_w, \quad (54)$$

where

$$D_4 = \frac{\Delta\lambda_w - \Delta\lambda_b}{L - \Delta\lambda_w} D(h_w, \Delta\lambda_b) + \int_0^{\Delta\lambda_b} \frac{1}{L - \Delta\lambda_w} D(h_w, \Delta x) d\Delta x + + \int_{L_1^{n_4}}^{L_1^{m_4}} \int_{\Delta\lambda_w - L_1}^{\Delta\lambda_b} \left(\frac{1}{L - \Delta\lambda_w} D(h_w, \Delta x) \frac{1}{L_1^{m_4} - L_1^{n_4}} \right) d\Delta x dL_1. \quad (55)$$

E. Case 5

When condition 2 and condition 5 are satisfied, the selection range of L_1 is given by

$$L_1 \in \begin{cases} (L_1^{n_5}, L_1^{m_5}) \triangleq (L - \Delta\lambda_w - \Delta\lambda_b, \Delta\lambda_w), \\ \text{for } L \in (2\Delta\lambda_w, 2\Delta\lambda_w + \Delta\lambda_b); \\ (L_1^{n_5}, L_1^{m_5}) \triangleq (\Delta\lambda_w - \Delta\lambda_b, L - \Delta\lambda_w), \\ \text{for } L \in (2\Delta\lambda_w - \Delta\lambda_b, 2\Delta\lambda_w]. \end{cases} \quad (56)$$

Thus, $P_{c5} = (L_1^{m_5} - L_1^{n_5}) / (L - \Delta\lambda_b)$. In this case, the average KL divergence is written as

$$\mathbb{E}[D(h_w, \Delta\lambda)]_{c5} = \int_0^\infty D_5 \times f'(h_w) dh_w, \quad (57)$$

where

$$D_5 = \frac{\Delta\lambda_w - \Delta\lambda_b}{L - \Delta\lambda_w} D(h_w, \Delta\lambda_b) + \int_{L_1^{n_5}}^{L_1^{m_5}} \int_{\Delta\lambda_w - L_1}^{\Delta\lambda_b} \frac{1}{L - \Delta\lambda_w} D(h_w, \Delta x) \frac{1}{L_1^{m_5} - L_1^{n_5}} d\Delta x dL_1 + \int_{L_1^{n_5}}^{L_1^{m_5}} \int_{L_1 + \Delta\lambda_w + \Delta\lambda_b - L}^{\Delta\lambda_b} \left\{ \frac{1}{L - \Delta\lambda_w} D(h_w, \Delta x) \times \frac{1}{L_1^{m_5} - L_1^{n_5}} \right\} d\Delta x dL_1. \quad (58)$$

F. Case 6

When condition 2 and condition 6 are satisfied, the selection range of L_1 is given by

$$L_1 \in \begin{cases} [L_1^{n_6}, L_1^{m_6}] \triangleq [L - \Delta\lambda_w, \Delta\lambda_w], \\ \text{for } L \in (2\Delta\lambda_w - \Delta\lambda_b, 2\Delta\lambda_w); \\ (L_1^{n_6}, L_1^{m_6}) \triangleq (\Delta\lambda_w - \Delta\lambda_b, \Delta\lambda_w), \\ \text{for } L \in [\Delta\lambda_w, 2\Delta\lambda_w - \Delta\lambda_b]. \end{cases} \quad (59)$$

Thus, $P_{c6} = (L_1^{m_6} - L_1^{n_6}) / (L - \Delta\lambda_b)$. In this case, the average KL divergence is written as

$$\mathbb{E}[D(h_w, \Delta\lambda)]_{c6} = \int_0^\infty D_6 \times f'(h_w) dh_w, \quad (60)$$

where

$$D_6 = \frac{\Delta\lambda_w - \Delta\lambda_b}{L - \Delta\lambda_w} D(h_w, \Delta\lambda_b) \frac{1}{L_1^{m_6} - L_1^{n_6}} dL_1 + \int_{L_1^{n_6}}^{L_1^{m_6}} \int_{\Delta\lambda_w - L_1}^{\Delta\lambda_b} \frac{1}{L - \Delta\lambda_w} D(h_w, \Delta x) \frac{1}{L_1^{m_6} - L_1^{n_6}} d\Delta x dL_1. \quad (61)$$

G. Case 7

When condition 3 and condition 4 are satisfied, the selection range of L_1 is given by

$$L_1 \in \begin{cases} [L_1^{n_7}, L_1^{m_7}] \triangleq [0, \Delta\lambda_w - \Delta\lambda_b], \\ \text{for } L \in (2\Delta\lambda_w, \infty); \\ [L_1^{n_7}, L_1^{m_7}] \triangleq [0, L - \Delta\lambda_w - \Delta\lambda_b], \\ \text{for } L \in [\Delta\lambda_w + \Delta\lambda_b, 2\Delta\lambda_w]. \end{cases} \quad (62)$$

Thus, $P_{c7} = (L_1^{m_7} - L_1^{n_7}) / (L - \Delta\lambda_b)$. In this case, the average KL divergence is written as

$$\mathbb{E}[D(h_w, \Delta\lambda)]_{c7} = \int_0^\infty D_7 \times f'(h_w) dh_w, \quad (63)$$

where

$$D_7 = \int_{L_1^{n_7}}^{L_1^{m_7}} \frac{L_1}{L - \Delta\lambda_w} D(h_w, \Delta\lambda_b) \frac{1}{L_1^{m_7} - L_1^{n_7}} dL_1 \\ + \int_0^{\Delta\lambda_b} \frac{1}{L - \Delta\lambda_w} D(h_w, \Delta x) d\Delta x. \quad (64)$$

H. Case 8

When condition 3 and condition 5 are satisfied, the selection range of L_1 is given by

$$L_1 \in \begin{cases} (L_1^{n_8}, L_1^{m_8}] \triangleq (L - \Delta\lambda_w - \Delta\lambda_b, \Delta\lambda_w - \Delta\lambda_b], \\ \text{for } L \in (2\Delta\lambda_w - \Delta\lambda_b, 2\Delta\lambda_w]; \\ (L_1^{n_8}, L_1^{m_8}) \triangleq (L - \Delta\lambda_w - \Delta\lambda_b, L - \Delta\lambda_w), \\ \text{for } L \in [\Delta\lambda_w, 2\Delta\lambda_w - \Delta\lambda_b]. \end{cases} \quad (65)$$

Thus, $P_{c8} = (L_1^{m_8} - L_1^{n_8}) / (L - \Delta\lambda_b)$. In this case, the average KL divergence is written as

$$\mathbb{E}[D(h_w, \Delta\lambda)]_{c8} = \int_0^\infty D_8 \times f'(h_w) dh_w, \quad (66)$$

where

$$D_8 \\ = \int_{L_1^{n_8}}^{L_1^{m_8}} \frac{L_1}{L - \Delta\lambda_w} D(h_w, \Delta\lambda_b) \frac{1}{L_1^{m_8} - L_1^{n_8}} dL_1 \\ + \int_{L_1^{n_8}}^{L_1^{m_8}} \int_{L_1 + \Delta\lambda_w + \Delta\lambda_b - L}^{\Delta\lambda_b} \left\{ \frac{1}{L - \Delta\lambda_w} D(h_w, \Delta x) \times \frac{1}{L_1^{m_8} - L_1^{n_8}} \right\} d\Delta x dL_1. \quad (67)$$

I. Case 9

When condition 3 and condition 6 are satisfied, the selection range of L_1 is given by

$$L_1 \in [L_1^{n_9}, L_1^{m_9}] \triangleq [L - \Delta\lambda_w, \Delta\lambda_w - \Delta\lambda_b], \quad (68)$$

for $L \in [\Delta\lambda_w, 2\Delta\lambda_w - \Delta\lambda_b]$. Thus, $P_{c9} = (L_1^{m_9} - L_1^{n_9}) / (L - \Delta\lambda_b)$. In this case, the average KL divergence is written as

$$\mathbb{E}[D(h_w, \Delta\lambda)]_{c9} = \int_0^\infty D(h_w, \Delta\lambda_b) f'(h_w) dh_w. \quad (69)$$

REFERENCES

- [1] Z. Xu, W. Liu, Z. Wang, and L. Hanzo, "Petahertz communication: Harmonizing optical spectra for wireless communications," *Digit. Commun. Netw.*, vol. 7, no. 4, pp. 605–614, Nov. 2021.
- [2] H. Luo, J. Wang, F. Bu, R. Ruby, K. Wu, and Z. Guo, "Recent progress of air/water cross-boundary communications for underwater sensor networks: A review," *IEEE Sensors J.*, vol. 22, no. 9, pp. 8360–8382, May 2022.
- [3] L.-K. Chen, Y. Shao, and Y. Di, "Underwater and water-air optical wireless communication," *J. Lightw. Technol.*, vol. 40, no. 5, pp. 1440–1452, Mar. 1, 2022.
- [4] T. Lin, C. Gong, J. Luo, and Z. Xu, "Dynamic optical wireless communication channel characterization through air-water interface," in *Proc. IEEE/CIC Int. Conf. Commun. China (ICCC Workshops)*, Aug. 2020, pp. 173–178.
- [5] X. Sun et al., "Field demonstrations of wide-beam optical communications through water-air interface," *IEEE Access*, vol. 8, pp. 160480–160489, 2020.
- [6] K. Enhos, D. Unal, E. Demirors, and T. Melodia, "Breaking through the air-water interface with software-defined visible light networking," *IEEE Internet Things Mag.*, vol. 5, no. 4, pp. 10–16, Dec. 2022.
- [7] Md. M. Hossain et al., "Low-noise speed-optimized large area CMOS avalanche photodetector for visible light communication," *J. Lightw. Technol.*, vol. 35, no. 11, pp. 2315–2324, Jun. 1, 2017.
- [8] Z. Jiang, C. Gong, and Z. Xu, "Achievable rates and signal detection for photon-level photomultiplier receiver based on statistical non-linear model," *IEEE Trans. Wireless Commun.*, vol. 18, no. 12, pp. 6015–6029, Dec. 2019.
- [9] K. Shahzad and X. Zhou, "Covert wireless communications under quasi-static fading with channel uncertainty," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 1104–1116, 2020.
- [10] X. Chen et al., "Covert communications: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 25, no. 2, pp. 1173–1198, 2nd Quart., 2023.
- [11] B. A. Bash, D. Goeckel, and D. Towsley, "Limits of reliable communication with low probability of detection on AWGN channels," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1921–1930, Sep. 2013.
- [12] P. H. Che, M. Bakshi, and S. Jaggi, "Reliable deniable communication: Hiding messages in noise," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2013, pp. 2945–2949.
- [13] L. Wang, G. W. Wornell, and L. Zheng, "Fundamental limits of communication with low probability of detection," *IEEE Trans. Inf. Theory*, vol. 62, no. 6, pp. 3493–3503, Jun. 2016.
- [14] K. S. K. Arumugam and M. R. Bloch, "Keyless covert communication over multiple-access channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2016, pp. 2229–2233.
- [15] L. Wang, "Covert communication over the Poisson channel," *IEEE J. Sel. Areas Inf. Theory*, vol. 2, no. 1, pp. 23–31, Mar. 2021.
- [16] J.-W. Shi, X.-H. Wang, P.-F. Yu, and J.-Y. Wang, "Fundamental performance limit of covert free-space optical communications," *Phys. Commun.*, vol. 54, pp. 1–9, Oct. 2022.
- [17] J. Wang, P. Yu, J. Shi, M. Lin, and J. Wang, "Research on fundamental performance limit of covert visible light communications," *CAS J. Electron. Inf. Technol.*, vol. 44, no. 8, pp. 1–10, Aug. 2022.
- [18] M. R. Bloch, "Covert communication over noisy channels: A resolvability perspective," *IEEE Trans. Inf. Theory*, vol. 62, no. 5, pp. 2334–2354, May 2016.
- [19] K. Shahzad, X. Zhou, S. Yan, J. Hu, F. Shu, and J. Li, "Achieving covert wireless communications using a full-duplex receiver," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 8517–8530, Dec. 2018.
- [20] X. Chen et al., "Multi-antenna covert communication via full-duplex jamming against a warden with uncertain locations," *IEEE Trans. Wireless Commun.*, vol. 20, no. 8, pp. 5467–5480, Aug. 2021.
- [21] C. Wang et al., "Covert communication assisted by UAV-IRS," *IEEE Trans. Commun.*, vol. 71, no. 1, pp. 357–369, Jan. 2023.
- [22] X. Chen, M. Sheng, N. Zhao, W. Xu, and D. Niyato, "UAV-relayed covert communication towards a flying warden," *IEEE Trans. Commun.*, vol. 69, no. 11, pp. 7659–7672, Nov. 2021.
- [23] J. Zhang et al., "Joint beam training and data transmission design for covert millimeter-wave communication," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 2232–2245, 2021.
- [24] Z. Duan, X. Yang, Y. Gong, D. Wang, and L. Wang, "Covert communication in uplink NOMA systems under channel distribution information uncertainty," *IEEE Commun. Lett.*, vol. 27, no. 5, May 2023.
- [25] S. Ma et al., "Optimal probabilistic constellation shaping for covert communications," *IEEE Trans. Inf. Forensics Security*, vol. 17, pp. 3165–3178, 2022.
- [26] I. A. Kadampot, M. Tahmasbi, and M. R. Bloch, "Multilevel-coded pulse-position modulation for covert communications over binary-input discrete memoryless channels," *IEEE Trans. Inf. Theory*, vol. 66, no. 10, pp. 6001–6023, Oct. 2020.
- [27] X. Zhou, S. Yan, J. Hu, J. Sun, J. Li, and F. Shu, "Joint optimization of a UAV's trajectory and transmit power for covert communications," *IEEE Trans. Signal Process.*, vol. 67, no. 16, pp. 4276–4290, Aug. 2019.
- [28] S. Yan, B. He, X. Zhou, Y. Cong, and A. L. Swindlehurst, "Delay-intolerant covert communications with either fixed or random transmit power," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 1, pp. 129–140, Jan. 2019.

- [29] Q. Hu, C. Gong, T. Lin, J. Luo, and Z. Xu, "Secrecy performance analysis for water-to-air visible light communication," *J. Lightw. Technol.*, vol. 40, no. 14, pp. 4607–4620, Jul. 15, 2022.
- [30] M. Forouzesh, P. Azmi, A. Kuhestani, and P. L. Yeoh, "Covert communication and secure transmission over untrusted relaying networks in the presence of multiple wardens," *IEEE Trans. Commun.*, vol. 68, no. 6, pp. 3737–3749, Jun. 2020.
- [31] D. Wang, Q. Fu, J. Si, N. Zhang, and Z. Li, "Improper Gaussian signaling based covert wireless communication in IoT networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2021, pp. 1–6.
- [32] X. Zhang, J. Liu, and L. Zhi, "Design and optimization on covert communication-oriented short packet D2D," in *Proc. 14th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Nov. 2022, pp. 257–263.
- [33] S. Santran, M. E. Martinez-Rosas, L. Canioni, and L. Sarger, "Characterization of optical nonlinearity in semiconductor photodiodes using cross-polarized autocorrelation," *IEEE J. Quantum Electron.*, vol. 40, no. 12, pp. 1687–1694, Dec. 2004.
- [34] W. Liu and Z. Xu, "APD nonlinearity and its impact on PAM-based visible light communication," *IEEE Commun. Lett.*, vol. 24, no. 5, pp. 1057–1061, May 2020.
- [35] W. B. Davenport and W. L. Root, *An Introduction to the Theory of Random Signals and Noise*, vol. 159. New York, NY, USA: McGraw-Hill, 1958.
- [36] Z. Wang, Q. Wang, W. Huang, and Z. Xu, *Visible Light Communications: Modulation and Signal Processing*. Hoboken, NJ, USA: Wiley, 2017.
- [37] G. P. Agrawal, *Fiber-Optic Communication Systems*. Hoboken, NJ, USA: Wiley, 2012.
- [38] D. A. Luong, T. C. Thang, and A. T. Pham, "Effect of avalanche photodiode and thermal noises on the performance of binary phase-shift keying-subcarrier-intensity modulation/free-space optical systems over turbulence channels," *IET Commun.*, vol. 7, no. 8, pp. 738–744, May 2013.
- [39] *Characteristics and Use of Si APD (Avalanche Photodiode)*, Hamamatsu Photon., Shizuoka, Japan, 2004.
- [40] F. Xu, M.-A. Khalighi, and S. Bourennane, "Impact of different noise sources on the performance of PIN- and APD-based FSO receivers," in *Proc. 11th Int. Conf. Telecommun.*, Jun. 2011, pp. 211–218.
- [41] S. Ma et al., "Robust beamforming design for covert communications," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 3026–3038, 2021.
- [42] W. Yang, G. Durisi, T. Koch, and Y. Polyanskiy, "Block-fading channels at finite blocklength," in *Proc. Int. Symp. Wireless Commun. Sys.*, Aug. 2013, pp. 1–4.
- [43] W. Yang, G. Durisi, T. Koch, and Y. Polyanskiy, "Quasi-static SIMO fading channels at finite blocklength," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2013, pp. 1531–1535.
- [44] W. Yang, G. Durisi, T. Koch, and Y. Polyanskiy, "Quasi-static multiple-antenna fading channels at finite blocklength," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4232–4265, Jul. 2014.
- [45] O. Ibe, *Markov Processes for Stochastic Modeling*. Amsterdam, The Netherlands: Elsevier, 2013.
- [46] X. Zhou, S. Yan, Q. Wu, F. Shu, and D. W. K. Ng, "Intelligent reflecting surface (IRS)-aided covert wireless communications with delay constraint," *IEEE Trans. Wireless Commun.*, vol. 21, no. 1, pp. 532–547, Jan. 2022.
- [47] M. N. Do, "Fast approximation of kullback-leibler distance for dependence trees and hidden Markov models," *IEEE Signal Process. Lett.*, vol. 10, no. 4, pp. 115–118, Apr. 2003.
- [48] J. R. Hershey and P. A. Olsen, "Approximating the Kullback Leibler divergence between Gaussian mixture models," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, vol. 4, Apr. 2007, pp. 317–320.



Qingqing Hu received the B.S. degree from Dalian University of Technology, Dalian, China, in 2016, and the M.S. degree from the University of Science and Technology of China, Hefei, China, in 2019, where she is currently pursuing the Ph.D. degree. Her research interests include information security, signal processing, and visible light communication.



Tianrui Lin received the B.S. degree from the University of Electronic Science and Technology of China, Chengdu, China, in 2019. He is currently pursuing the Ph.D. degree in optical wireless communication with the University of Science and Technology of China, Hefei, China. His research interests include water-to-air visible light communication and MIMO communication.



Tianjian Wei received the B.S. degree from Northwestern Polytechnical University, Xi'an, China, in 2021. He is currently pursuing the Ph.D. degree with the University of Science and Technology of China, Hefei, China. His research interests include water-to-air visible light communication and multi-user scheduling.



Nuo Huang received the B.S. degree in electronics and information engineering from Huazhong University of Science and Technology, Wuhan, China, in 2012, and the Ph.D. degree in information and communication engineering from the National Mobile Communications Research Laboratory, Southeast University, Nanjing, China, in 2019. From December 2015 to June 2017, he was a Visiting Student with the Department of Electrical Engineering, Columbia University, New York, NY, USA. He was selected by the National Postdoctoral Program for Innovative Talents in 2019. He is currently a Research Associate with the Department of Electronic Engineering and Information Science, University of Science and Technology of China. His research interests include resource allocation and transceiver design in wireless (optical) communications.



Yi-Jun Zhu received the B.Eng., M.Sc., and Ph.D. degrees from the National Digital Switching System Engineering and Technological Research Center (NDSC), Zhengzhou, Henan, China, in 1999, 2002, and 2010, respectively. In 2011, he visited the Department of Electrical and Computer Engineering, McMaster University. He is currently with the School of Information Engineering University. His research interests include the areas of wireless communication theory, visible light communications, and signal processing.



Chen Gong (Senior Member, IEEE) received the B.S. degree in electrical engineering and mathematics (minor) from Shanghai Jiaotong University, Shanghai, China, in 2005, the M.S. degree in electrical engineering from Tsinghua University, Beijing, China, in 2008, and the Ph.D. degree from Columbia University, New York, NY, USA, in 2012. He was a Senior Systems Engineer with Qualcomm Research, San Diego, CA, USA, from 2012 to 2013. He is currently a Faculty Member with the University of Science and Technology of China. His research interests include wireless communications, optical wireless communications, and signal processing. He was awarded by Hong Kong Qiushi Outstanding Young Researcher Award in 2016.