Energy Efficient Short-Packet Covert Communications for Full-Duplex Wireless Systems with AoI Constraint

Yangfan Xu, Bin Yang, Lei Shao, Xiuwen Sun, Shikai Shen, Haibao Chen, Bao Gui, Tarik Taleb

***Abstract*—This paper investigates the energy efficient short-packet covert communications in a full-duplex wireless system with the constraint of age of information (AoI), where a transmitter covertly transmits information to a receiver operating on full-duplex communication mode, while a warden tries to detect the existence of this covert transmission. For supporting various security and time-sensitive applications in Internet of Things (IoT) with energy-limited devices, we focus on critical performances of short-packet communications in terms of energy efficiency, covertness and timeliness. To this end, we first analyze the covert constraint condition, AoI and covert energy efficiency (CEE) under the finite block length codewords. We then formulate CEE maximization as an optimization problem with the constraints of some parameters, i.e., the covert constraint condition, freshness of information, block length, the prior transmission probability and transmit powers of transmitter and receiver. We obtain maximum CEE by jointly optimizing the block-length, prior transmission probability and transmit powers using the interior point method. Finally, numerical results are presented to illustrate the impact of the parameters on CEE.**

***Index Terms*—Short-packet communication, covert energy efficiency, age of information.**

This work is supported in part by the National Natural Science Foundation of China under Grant No. 62372076in part by the Education Department Research Foundation of Anhui Province under Grant No. DTR2023051; in part by the Academician Weiming Shen Workstation under Grant No. 202505AF350084, Yunnan, China; in part by the Yunnan Key Laboratory of Intelligent Logistics Equipment and Systems under Grant No. 202449CE340008; in part by the Innovation Research Team on Future Network Technology of Chuzhou University; in part by the Innovation Team on Smart Home Appliance Security and Applications of Chuzhou Cityin part by the Innovative Research Team on Application of Big Data and Financial Technology of Chuzhou University; in part by Yunnan Key Laboratory of Smart City in Cyberspace Security under Grant No. 202205AG340011 , Yuxi Normal University, Yuxi, China; in part by the ICTFICIAL Oy, Finland; in part by the European Union's HE Research and Innovation Program HORIZON-JUSNS-2023 through the 6G-Path under Grant No.101139172. (Corresponding authors: Bin Yang; Shikai Shen.)

Yangfan Xu and Lei Shao are with the School of Computer and Information Engineering, Chuzhou University, Chuzhou 239000, China and the School of Computer Science and Technology of Anhui University, Hefei 230000, China (e-mail: xuyangfan1112@gmail.com, 2513175870@qq.com).

Bin Yang, Haibao Chen and Bao Gui are with the School of Computer and Information Engineering, Chuzhou University, Chuzhou 239000, China (e-mail: yangbinchi@gmail.com, chb@chzu.edu.cn, gui_bao1@163.com).

Xiuwen Sun is with the School of Computer Science and Technology of Anhui University, Hefei 230000, China (e-mail: mr.xiuwen@gmail.com).

Shikai Shen is with School of Information and Technology, Kunming University, Kunming 650214, China and Yunnan Key Laboratory of Smart City and Cyberspace Security, Yuxi 653100, China (e-mail: kmssk2000@sina.com).

Tarik Taleb is with the Faculty of Electrical Engineering and Information Technology, Ruhr University Bochum, Bochum 44801, Germany (e-mail: tarik.taleb@rub.de).

I. Introduction

Internet of Things (IoT) in upcoming sixth-generation wireless systems will offer communication services for billions of machine-type devices in extensive applications, e.g., smart transportation, smart medical care, and smart home [1–3]. Most of these applications require high system performance like high energy efficiency, low latency and high data rate [4, 5]. Short-packet communications can significantly enhance system performances, which have received widespread attention from both the industrial and academic communities [6, 7]. However, due to the inherent features of broadcast and openness of wireless channel, short-packet communications are facing a serious security threat. The existing cryptographic methods mainly rely on high computational complexity to ensure data transmission security, and the corresponding high energy consumption cannot effectively meet the security requirements for a large number of energy-limited IoT devices. As an effective supplement to the methods, covert communications aim at hiding wireless communication process to achieve data transmission security, which have been identified as a cutting-edge technology [8]. Thus, it is critical to study energy-efficient short-packet communications for supporting various security-sensitive IoT applications.

Existing works can be categorized into covert communications without and with age of information (AoI) in wireless systems (See Related Work of Section II for details), where AoI is a fundamental metric of freshness of information. For the works without AoI, these works explore covert performances in terms of covert throughput and detection error probability in various scenarios, such as noise uncertainty [9], channel uncertainty [10], interference uncertainty [11], artificial noise generated through full-duplex interferes with the monitor's detection [12–18], two-hop relay scenario [19], and unmanned aerial vehicle (UAV)-assisted scenario [20]. However, both timeliness and covertness of information transmission are of great importance for supporting time and privacy sensitive IoT applications like military communications, vehicle networking, and cybersecurity emergency response. Fortunately, AoI is widely recognized as a measure of timeliness, which is defined as the time duration from the beginning of generating the data packet [21]. For the works with AoI, the main focus is on covert performances under AoI constraint without the assistance of jamming signals [22–26]. It is notable that short-packet communications have been widely adopted to ensure the timeliness of information transmission [23, 27, 28].

TABLE I: DIFFERENCES BETWEEN OUR SCHEME AND OTHER RELATED SCHEMES

References	AoI Constraint	Full-duplex Technique	Covert Energy Efficiency	Short-Packet Communication	Transmit Probability Control	Block-Length Optimization	Interior Point Method
[22]	✓	✓	×	×	✓	×	×
[23]	✓	×	×	✓	×	×	×
[24]	✓	×	✓	×	✓	×	×
[25]	✓	✓	×	✓	×	✓	×
[26]	✓	×	×	×	×	×	×
[27]	×	×	×	✓	×	✓	×
[28]	✓	×	×	✓	×	×	×
[29]	✓	×	×	✓	×	✓	×
[30]	×	✓	✓	×	×	×	×
[31]	×	×	✓	×	×	×	×
Our Scheme	✓	✓	✓	✓	✓	✓	✓

Particularly, the work in [29] explores the short-packet covert communications for the first time.

Energy consumption poses a fundamental challenge for IoT devices, which are inherently constrained by limited power resources [30, 31]. For battery-powered nodes such as remote sensors and actuators, operational lifetime depends critically on energy efficiency. At the same time, in mission-critical applications including industrial automation and military surveillance, the timeliness of information, typically measured by AoI, is essential for effective monitoring and decision-making. In addition, in security-sensitive scenarios, maintaining covert communication is crucial to prevent detection or interception by adversaries. Simultaneously achieving high energy efficiency, data timeliness, and communication covertness remains a critical and underexplored challenge in wireless systems. Motivated by this observation, this paper investigates the energy efficient short-packet communications in a full-duplex wireless system with AoI constraint. In such a system, a receiver operating over full-duplex mode can send artificial noise to interfere with malicious node's detection. The differences between our work and existing works are shown in Table I. The main contributions of this paper are summarized as follows.

- We consider a wireless system consisting of a transmitter, a full-duplex receiver and a detector warden. The transmitter covertly transmits short-packets to the receiver, while warden detects the covert transmission process. Simultaneously, the receiver sends artificial noise to confuse warden's detection. In this system, we derive a closed-form expression for the average AoI, which is used to measure information freshness.
- We then derive the covert energy efficiency (CEE) to evaluate the system energy consumption. We also formulate CEE's maximization as an optimization problem with the constraint of the transmit powers of the transmitter and receiver, block length, and transmission probability. We further use the interior point method to solve the optimization problem.
- Finally, numerical results are presented to illustrate how system parameters affect the CEE in the wireless system.

The remaining parts of this paper are organized as follows. Section II introduces the related work. Section III presents the system model and detection performance. Section IV models the AoI and CEE. Section V formulates the CEE maximization as an optimization problem and solves it. Section VI provides numerical results. Section VII concludes the paper.

II. Related Work

1) Covert Communications without AoI: The article [9] analyzes the minimum error detection probability of the warden under noise uncertainty, which they use to measure the system's covertness. Then, they derive the approximate outage probability under the improper Gaussian signaling (IGS) scheme and solve the optimization problem by jointly designing the transmit power and IGS factor. This paper [10] investigates covert communication in an uplink non-orthogonal multiple access (NOMA) system under channel distribution information uncertainty. It derives the optimal power settings for reliable user and covert user under imperfect channel distribution information. The study [11] investigates the optimal detection performance of the adversary Willie in two extreme cases: one with perfect Channel State Information (CSI) and the other with only Channel Distribution Information (CDI). The results show that the quality of CSI does not help Willie improve his detection performance. The work in [12] proposes in their study a covert communication scheme in which the sender tries to hide its transmissions with the assistance of full-duplex receivers, while the monitoring party will use radiometers to detect such covert transmissions. This study is the first to explore the use of full-duplex receivers to control artificial noise for covert communication.

The work in [13] investigates a wireless communication system that achieves channel covertness with a full-duplex receiver under channel fading conditions. The analysis shows that although the transmission of artificial noise can cause self-interference, it effectively improves communication covertness. The article [14] investigates delay-constrained covert communications with the aid of a full-duplex receiver. The paper demonstrates that transmitting artificial noise at a fixed power does improve the performance of covert communications with delay constraints. In most covert communication studies, the primary CSI is usually assumed to be known or

partially known. However, covert users may find it difficult to perform channel estimation in practical situations, or even to obtain the CSI of the monitoring party. The work in [15] investigates the problem of realizing covert communications over incoherent Rayleigh fading channels, both fast fading and slow fading scenarios. It is shown that a positive covert rate can be achieved in both fast fading and slow fading cases. The realization of covert communications using full-duplex receivers is promising for many applications.

In work [16], they study the robust joint power and position optimization problem for full-duplex receivers in the presence of warden position uncertainty. It analyzes the effect of the warden position uncertainty on the covert communication performance. Then, the robust optimization technique is used to maximize the effective covert throughput of the transceiver by optimizing the power and position of receivers while satisfying the sufficient covertness requirement. In paper [17], research has been carried out in the area of covert communication systems for UAVs. By using classical probability theory, the optimal detection threshold and the minimum detection error probability of warden under a typical UAV covert communication model are analyzed. Considering the maneuverability of UAVs, the proposed optimal position design method can significantly improve the covert performance of UAV systems. The work in [19] focuses on a two-hop wireless relay system, where a relay operates over either full-duplex mode or half-duplex mode. The work in [20] utilizes a UAV as relay and indicates that the full-duplex UAV can enhance the transmission performance and confuse the illegal detection of the warden. The aforementioned work focuses on enhancing communication secrecy by exploiting the artificial noise generated by full-duplex communication. However, the timeliness of information is also crucial. Unfortunately, these papers did not consider the impact of information timeliness.

2) Covert Communications with AoI: Information freshness has been relatively little studied in covert communications. This article [22] is the first to consider the requirement of information freshness in covert communications, and proposes a covertness maximization problem under the constraint of average AoI. The work in [23] first constructs a novel covert D2D communication model, analyzes its detection performance from the perspective of the warden, and derives the minimum detection error probability to measure the covert nature of the communication. Then, under the constraints of communication covertness and information freshness, it jointly optimizes the transmission probability of data packets and transmission power to maximize the CEE of D2D. The work in [24] investigates the covert information freshness in Intelligent Reflecting Surface (IRS)-assisted multi-antenna communication. By using channel statistics information, the paper jointly designs the active beamforming of transmitter and receiver, the passive beamforming of the IRS, and the packet length of confidential packets as a non-convex optimization problem. Finally, it is concluded that the optimized packet length decreases as the quality of the covert channel improves.

The work in [25] considers the age of channel (AoC) change, which is the time duration of CSI close to static and is important for reliable covert communication in dynamic environments. Depending on the packet length or the dynamic environment, channel changes may occur faster than the optimal AoI. This paper [26] proposes a dynamic covert communication scheme in time-varying channels by jointly optimizing the AoI and AoC, a novel metric that quantifies the duration of stable channel states. Addressing the limitation of existing works that prioritize AoI but neglect AoC, the scheme dynamically adjusts transmission strategies to ensure reliable covert communication under channel dynamics. We can observe that the aforementioned work primarily focuses on minimizing the age of information and enhancing the system's covertness. We know that artificial noise generated by full-duplex techniques can further improve the system's covertness. Therefore, it is worthwhile to delve into how to incorporate full-duplex techniques to achieve a better balance between communication timeliness and covertness.

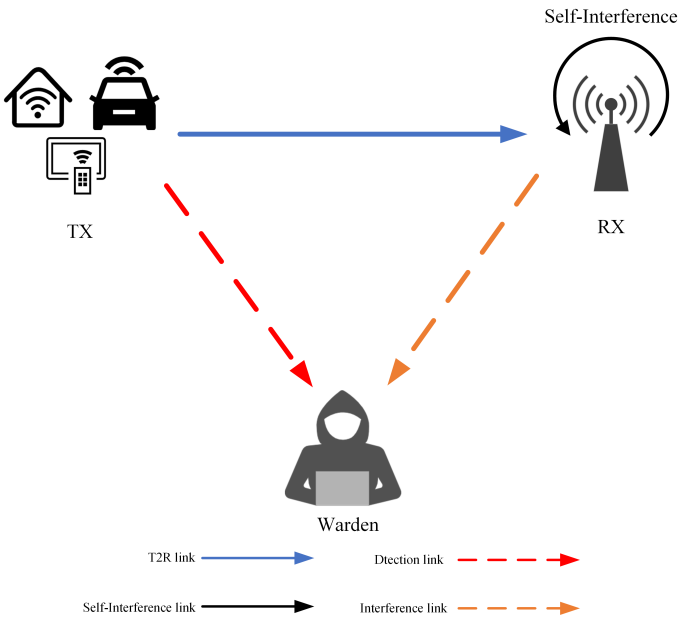


Fig. 1: System model for the full-duplex wireless covert communication system.

III. SYSTEM MODEL AND DETECTION PERFORMANCE

As shown in Fig. 1, we consider a wireless system consisting of a warden WD, a transmitter TX, and a full-duplex receiver RX. In the system, TX attempts to covertly transmit short-packets to RX, while WD detects the very existence of covert transmission. TX, RX, and WD operate on the same sub-30 GHz frequency bands, such as the 2.4 GHz ISM band and the 800 MHz narrowband IoT band. Each of TX and WD is equipped with a single omnidirectional dipole antenna, while RX utilizes two directional patch antennas to simultaneously receive short-packets and transmit artificial noise (AN) to confuse WD.

We model each channel as an independent Nakagami-m fading channel considering discrete-time slots. This means that the channel gains remain unchanged at each slot and

independently change from one slot to another. The channel gain from TX to RX, that from TX to WD and that from RX to WD are represented as H_{TR}, H_{TW}, and H_{RW}, respectively. The gain of the self-interference channel is represented as H_{RXRX}. $H_{TR}, H_{TW}, H_{RW}, H_{RXRX} \sim Nakagami(m, 1)$, where m is a fading parameter describing the severity of small-scale fading. The larger m, the milder the fading. When $m = 1$, the Nakagami-m fading channel degenerates into Rayleigh channel, corresponding to typical non-line-of-sight environments. When $m > 1$, the Nakagami-m fading channel tends to Rician channel, corresponding to environments with a strong line-of-sight path.

At each time slot, TX covertly transmits each short-packet containing K nats of information with a prior transmission probability P^{tra}. We use statistical hypothesis testing to determine whether TX is active to send the covert short-packet. The alternative hypothesis H_1 represents that TX is active, while the null hypothesis H_0 corresponds to the case where TX is silent. RX's received signal y_{RX} at each time slot is as follows

$$y_{RX} = \sqrt{P_{TX}} H_{TR} x_{TX} + \sqrt{\phi P_{RX}} H_{RXRX} v_{RX} + n_{RX}. \quad (1)$$

Here, $\phi \in [0, 1]$ denotes the self-interference cancellation coefficient, which quantifies the degree of self-interference mitigation achieved by various techniques. Compared with conventional methods such as frequency isolation, polarization isolation, and spatial isolation that rely on hardware or signal-domain separation, our approach models their effects in a unified manner through ϕ. Specifically, $\phi = 0$ represents complete cancellation of self-interference (e.g., by using frequency-division multiplexing or antenna separation), while $\phi = 1$ indicates no cancellation. By adjusting ϕ, the performance of different self-interference mitigation techniques can be effectively represented and compared in a unified manner. The signal transmit power of TX is denoted by P_{TX}, and the AN power of RX is denoted by P_{RX}. x_{TX} denotes the signal that TX transmits to RX. $v_{RX} \sim \mathcal{CN}(0, 1)$ denotes the normalized complex Gaussian noise signal with zero mean transmitted by RX, which can confuse WD's detection and cause self-interference to RX. $n_{RX} \sim \mathcal{CN}(0, \sigma^2_{RX})$ represents additive white Gaussian noise (AWGN) with mean 0 and variance σ^2_{RX}. WD's observation signal is as follows

$$H_0 : z = \sqrt{P_{RX}} H_{RW} v_{RX} + n_{WD}, \quad (2)$$

and

$$H_1 : z = \sqrt{P_{TX}} H_{TW} x_{TX} + \sqrt{P_{RX}} H_{RW} v_{RX} + n_{WD}, \quad (3)$$

where z represents received signal at WD. When H_0 is true, WD receives an interference signal of power P_{RX} from RX. When H_1 is true, WD receives a signal of power P_{TX} from TX and an interference signal of power P_{RX} from RX. Here, $n_{WD} \sim \mathcal{CN}(0, \sigma^2_{WD})$ represents AWGN with mean 0 and variance σ^2_{WD}.

A. Detection Performance at WD

We use P_0 and P_1 to denote the probability distributions of WD's channel observation under H_0 and H_1, respectively. To determine the detection error probability at WD, we define two types of errors in covert communications: false alarm and miss detection. The former represents WD making a decision favorable to H_1 when H_0 is true, while the latter represents WD making a decision favorable to H_0 when H_1 is true. The false alarm probability and miss detection probability are denoted by P_{FA} and P_{MD}, respectively. Then, the detection error probability P_{WE} at WD is determined as

$$P_{WE} = (1 - P^{tra}) P_{FA} + P^{tra} P_{MD}, \quad (4)$$

where P^{tra} is the prior transmission probability. We use a generalized bound on the total variation distance to guarantee the covert constraint, and then we have

$$P_{WE} \geq \min\{P_1^{tra}, P^{tra}\} - \max\{P_1^{tra}, P^{tra}\} V(P_0, P_1), \quad (5)$$

where $P_1^{tra} = 1 - P^{tra}$, and $V(P_0, P_1)$ is the total distance between P_0 and P_1. Covert communications require $P_{WE} \geq \min\{P_1^{tra}, P^{tra}\} - \epsilon$ to ensure communication covertness, where ϵ is an arbitrarily small constant representing the covert requirement [32]. Therefore, according to (5), requiring $\max\{P_1^{tra}, P^{tra}\} V(P_0, P_1) \leq \epsilon$ also ensures covert communications. According to the Pinsker inequality, we have

$$V(P_0, P_1) \leq \sqrt{\frac{1}{2} D(P_0 \,||\, P_1)}. \quad (6)$$

where $D(P_0 \,||\, P_1)$ is the relative entropy between P_0 and P_1. Then, we can derive

$$D(P_0 \,||\, P_1) \leq \frac{2\epsilon^2}{\max\{1 - P^{tra}, P^{tra}\}^2}. \quad (7)$$

The formula for the relative entropy between P_0 and P_1 is

$$D(P_0 \,||\, P_1) = D_L \int_X P_0(x) \ln\left(\frac{P_0(x)}{P_1(x)}\right) dx, \quad (8)$$

where D_L is the block length. The expression for the relative entropy $D(P_0 \,||\, P_1)$ can be derived easily using equation (8). Therefore, equation (7) represents our covertness requirement throughout the paper.

$P_0(x)$ and $P_1(x)$ are the densities of P_0 and P_1 , respectively. We proceed to derive the probability density functions $P_0(x)$ and $P_1(x)$ of the signals received at WD's end under the scenarios where TX does not transmit and where it does, respectively, according to equations (7) and (8). Assuming these equations correspond to the cases where H_0 and H_1 are true, respectively, the received signals at WD follow the Nakagami distribution with the scale parameters of $P_{RX}|H_{RW}|^2 + \sigma^2_{WD}$ and $P_{TX}|H_{TW}|^2 + P_{RX}|H_{RW}|^2 + \sigma^2_{WD}$, respectively. $P_0(x)$ is given by

$$P_0(x) = \frac{2m^m x^{2m-1}}{\Gamma(m)(P_{RX}|H_{RW}|^2 + \sigma^2_{WD})^m} \exp\left(-\frac{mx^2}{P_{RX}|H_{RW}|^2 + \sigma^2_{WD}}\right), \quad (9)$$

and $P_1(x)$ is given by the following expression

$$P_1(x) = \frac{2m^m x^{2m-1}}{\Gamma(m)(P_{TX}|H_{TW}|^2 + P_{RX}|H_{RW}|^2 + \sigma^2_{WD})^m} \times \exp\left(-\frac{mx^2}{P_{TX}|H_{TW}|^2 + P_{RX}|H_{RW}|^2 + \sigma^2_{WD}}\right), \quad (10)$$

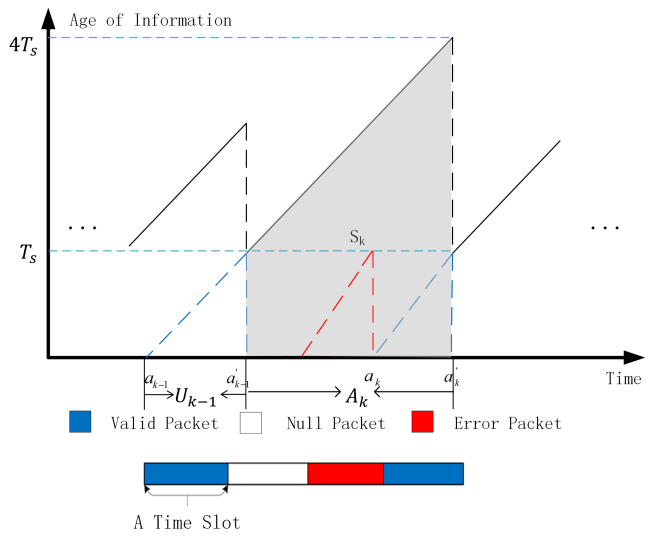


Fig. 2: The relationship between age of information and time.

where $\Gamma(m)$ represents Gamma function. Substituting equations (9) and (10) into equation (8), we obtain

$$\begin{aligned} D(P_0 \,||\, P_1) &= D_L m(\ln\left(1+\frac{P_{TX}|H_{TW}|^2}{P_{RX}|H_{RW}|^2+\sigma_{WD}^2}\right) \\ &\quad - \frac{P_{TX}|H_{TW}|^2}{P_{TX}|H_{TW}|^2+P_{RX}|H_{RW}|^2+\sigma_{WD}^2}) \\ &= D_L m(\ln(1+\zeta)-\frac{\zeta}{\zeta+1}), \end{aligned} \tag{11}$$

where $\zeta = \frac{P_{TX}|H_{TW}|^2}{P_{RX}|H_{RW}|^2+\sigma_{WD}^2}$ denotes the signal-to-interference-noise ratio (SINR) at WD. In covert communications, the transmission power is usually small resulting in a small SINR at WD. When x tends to 0, $\ln(1+x) \sim x$. Thus, the covert constraint can be simplified as

$$\frac{D_L m \zeta^2}{\zeta+1} \leq \frac{2\epsilon^2}{\max\{1-P^{tra}, P^{tra}\}^2}. \tag{12}$$

IV. Age Of Information And Covert Energy Efficiency Modeling

A. Age of Information

In this system, TX encodes the generated status updates into a short data packet and transmits the encoded short data packet to RX. We assume that TX generates a short data packet periodically within a time slot duration T_s, and TX transmits the short data packet with a certain probability P^{tra} to ensure covertness. Given that coding and propagation delays are negligible in short-packet communication, we approximate the time slot duration T_s as the transmission delay D_L/B, where B denotes the system bandwidth [29]. Moreover, to ensure the timeliness of the communication, we do not consider retransmission mechanisms [30]. Due to the impact of decoding failures, RX may fail to successfully receive the data packet. Therefore, we define the valid packet to characterize the data packet transmitted by TX and successfully decoded by RX. The AoI is defined as the time elapsed since the generation of the latest valid packet. If the latest valid packet received by RX at time τ is the k-th valid packet, generated at time a_k, then the instantaneous AoI $\triangle(\tau)$ is

$$\triangle(\tau) = \tau - a_k. \tag{13}$$

The AoI increases linearly over time and resets to the age of the most recent valid packet whenever RX receives a new one. As shown in Fig. 2, we consider an example of AoI evolution over time. When TX does not send the data packet (i.e., null packet) or when TX sends the data packet but RX fails to decode it (i.e., error packet), the AoI rises steadily over time until RX successfully receives a valid packet. We define the generation time and reception time of the k-th valid packet as a_k and $a_k^{'}$, respectively. The residence time of the k-th valid packet can be expressed as $U_k = a_k^{'} - a_k$, which is fixed and equal to the original age of the valid packet at RX's end, i.e., $U_k = T_s$. Due to the probabilistic transmission strategy and potential decoding failures of data packets, there may be multiple time slots (denoted by N) between consecutive valid packets received by RX. A_k represents the time interval between the k-th and $(k-1)$-th consecutive valid packets received by RX, i.e., $A_k = a_k^{'} - a_{k-1}^{'}$. The average AoI is a critical metric for assessing the freshness of data packet at the receiving end [22]. For a time interval $(0,T)$, the average AoI is given by

$$\overline{AoI} = \lim_{T\to\infty} \frac{1}{T}\int_0^T \triangle(\tau)\, d\tau. \tag{14}$$

From the above equation, the continuous-time integral $\int_0^T \triangle(\tau)\, d\tau$ in (14) represents the aggregate area under the AoI curve, as shown in Fig. 2. To facilitate the derivation of an analytical expression, this area is decomposed into a sum of geometric areas S_k in (15), where S_k represents the shaded area in Fig. 2 accumulated over the interval from $a_{k-1}^{'}$ to $a_k^{'}$. Therefore, the average AoI is the calculated as the sum of S_k divided by the total time T, resulting in reformulation of equation (14) as (15), which is given by

$$\overline{AoI} = \lim_{T\to\infty} \frac{1}{T}\sum_{k=1}^{K(T)} S_k = \lim_{T\to\infty} \frac{K(T)}{T} E(S). \tag{15}$$

In the equation, $K(T)$ represents the number of valid packets received by RX during the time interval $(0,T)$, $E(S)$ denotes the average value of S_k during the time interval $(0,T)$. Additionally, $\lim_{T\to\infty} \frac{K(T)}{T}$ can be understood as the average arrival rate of valid packet at RX. As shown in Fig. 2, the upper half of S_k is composed of an isosceles triangle. Assuming there are N time slots between the $(k-1)$-th and k-th valid packets, the first moment of A_k is given by

$$\begin{aligned} E(A_k) &= \sum_{n=1}^{\infty} nT_s P(N=n) \\ &= P^{tra}(1-P^{de})T_s \sum_{n=1}^{\infty} n(1-P^{tra}(1-P^{de}))^{n-1} \\ &= \frac{D_L}{BP^{tra}(1-P^{de})}, \end{aligned} \tag{16}$$

where P^{de} is the decoding error probability. Therefore, the average arrival rate of valid packet is

$$\lim_{T\to\infty}\frac{K(T)}{T}=\frac{1}{E(A_k)}=\frac{BP^{tra}}{D_L}(1-P^{de}). \tag{17}$$

Assuming the time interval between the $(l-1)$-th and the l-th transmitted data packets is $\triangle(T)$, then its first moment is given by

$$E(\triangle(T))=\frac{D_L}{B}P^{tra}\sum_{n=1}^{\infty}n(1-P^{tra})^{n-1}=\frac{D_L}{P^{tra}B}. \tag{18}$$

Then, its second moment is given by

$$E(\triangle^2(T))=\frac{D_L^2}{B^2}P^{tra}\sum_{n=1}^{\infty}n^2(1-P^{tra})^{n-1}=\frac{D_L^2}{P^{tra}B^2}(\frac{2}{P^{tra}}-1). \tag{19}$$

From equations (16), (18), and (19), the second moment of A_k can be obtained as follows

$$\begin{aligned}E(A_k^2)&=\sum_{n=1}^{\infty}E\{(\triangle(T))^2|N=n\}P(N=n)\\&=\sum_{n=1}^{\infty}E\{\sum_{i=1}^{n}\triangle(T)_i^2+\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\triangle(T)_i\triangle(T)_j\}P(N=n)\\&=(1-P^{de})\sum_{n=1}^{\infty}M(1-(1-P^{de}))^{n-1}\\&=\frac{2D_L^2}{(P^{tra}B(1-P^{de}))^2}-\frac{D_L^2}{P^{tra}(1-P^{de})B^2}.\end{aligned} \tag{20}$$

Because $\triangle(T)_i$ and $\triangle(T)_j$ are independently distributed, we have $M=nE(\triangle^2(T))+n(n-1)E(\triangle(T))^2$.

As shown in Fig. 2, the area of S_k is composed of an isosceles triangle and a rectangle, so

$$E(S)=\frac{1}{2}E(A_k^2)+T_sE(A_k). \tag{21}$$

Substituting equations (16) and (20) into (21), we can obtain

$$E(S)=\frac{2D_L^2+P^{tra}D_L^2(1-P^{de})}{2(P^{tra}B(1-P^{de}))^2}. \tag{22}$$

Therefore, by using equations (15), (17), and (22), we obtain

$$\overline{AoI}=\frac{D_L}{P^{tra}B(1-P^{de})}+\frac{D_L}{2B}. \tag{23}$$

B. Covert Energy Efficiency

For a given coding rate $R_b=K/D_L$ [33], P^{de} represents the decoding error probability, which is given by [29, 34]

$$P^{de}\approx Q(\frac{\sqrt{D_L}(1+\gamma_b)(\ln(1+\gamma_b)-R_b)}{\sqrt{\gamma_b(2+\gamma_b)}}), \tag{24}$$

where γ_b is the SINR at RX, given by the following expression

$$\gamma_b=\frac{P_{TX}|H_{TR}|^2}{\phi P_{RX}|H_{RXRX}|^2+\sigma_{RX}^2}. \tag{25}$$

The Q function is given by the following expression

$$Q(x)=\frac{1}{\sqrt{2\pi}}\int_x^{\infty}\exp\left(-\frac{t^2}{2}\right)dt. \tag{26}$$

Based on equations (24), (25), and (26), an approximate value for the decoding error probability can be derived [29, 34]

$$P^{de}\approx\begin{cases}1, & \gamma_b<\frac{K}{D_L}-\frac{\sqrt{K\pi}}{D_L},\\ \frac{1}{2}-\frac{D_L}{2\sqrt{K\pi}}\left(\gamma_b-\frac{K}{D_L}\right), & \frac{K}{D_L}-\frac{\sqrt{K\pi}}{D_L}\le\gamma_b\le\frac{K}{D_L}+\frac{\sqrt{K\pi}}{D_L},\\ 0, & \gamma_b>\frac{K}{D_L}+\frac{\sqrt{K\pi}}{D_L}.\end{cases} \tag{27}$$

The effective covert rate C can be expressed as

$$C=P^{tra}BR_b(1-P^{de}), \tag{28}$$

where B denotes system bandwidth. To maximize the utilization of limited energy at the transmitter, we introduce a new metric called CEE. This metric measures the number of covert bits that can be successfully transmitted per unit of energy. The goal of this metric is to maximize energy efficiency while ensuring the covertness of the communication. The mathematical formula is as follows

$$\eta_a=\frac{P^{tra}BR_b(1-P^{de})}{P_{TX}+P_{RX}+P_c}, \tag{29}$$

where η_a is the CEE, P_c is the circuit power consumption.

V. Covert Energy Efficiency Maximization

In this section, our aim is to maintain the freshness of information transmitted from TX to RX while ensuring communication covertness, then maximize the CEE. Specifically, we first formulate CEE maximization as a constrained optimization problem. Then, we use an interior point method with the filter line search and inertia correction to solve this optimization problem by optimizing the transmission powers at the transmitter and receiver, prior transmission probability, and block length.

A. Problem Formulation

Our goal is to optimize transmit power of TX and RX, block length, and the prior transmission probability to maximize the CEE, while satisfying covert requirements and maximum power constraints. This can be formulated as the following optimization problem

$$\begin{aligned}\max_{D_L,P_{TX},P_{RX},P^{tra}}\quad&\eta_a,\\ s.t.\quad&D(P_0\,||\,P_1)\le\frac{2\epsilon^2}{\max\{1-P^{tra},P^{tra}\}^2}, &(29a)\\&\overline{AoI}\le\delta, &(29b)\\&D_L^{min}<D_L<D_L^{max}, &(29c)\\&0<P_{TX}<P_{TX}^{max}, &(29d)\\&0<P_{RX}<P_{RX}^{max}, &(29e)\\&0\le P^{tra}\le P_{max}^{tra}, &(29f)\end{aligned}$$

where (29a) represents the covert requirement constraint, (29b) represents the AoI constraint, and (29c) represents the block

length constraint. (29d) and (29e) represent the transmit power constraints of TX and RX, respectively, and (29f) represents the prior transmission probability constraint. Here, δ denotes the maximum acceptable average AoI, D_L^{max} and D_L^{min} represent the maximum block length and minimum block length, P_{TX}^{max} and P_{RX}^{max} denote the maximum transmit power of TX and RX, and P_{max}^{tra} denotes the maximum prior transmission probability.

We consider the case that $0 < P^{de} < 1$, and we can obtain the following expression of CEE η_a by substituting equation (27) into equation (29)

$$\eta_a = \frac{P^{tra} BK(D_L \gamma_b \sqrt{\frac{1}{K\pi}} + 1 - \sqrt{\frac{K}{\pi}})}{2D_L(P_{TX} + P_{RX} + P_c)}. \tag{30}$$

When $0 < P^{de} < 1$, $\frac{K-\sqrt{K\pi}}{\gamma_b} < D_L < \frac{K+\sqrt{K\pi}}{\gamma_b}$. We define $D_L^{low} < D_L < D_L^{up}$ as the intersection of $D_L^{min} < D_L < D_L^{max}$ and $\frac{K-\sqrt{K\pi}}{\gamma_b} < D_L < \frac{K+\sqrt{K\pi}}{\gamma_b}$, where $D_L^{low} = \max\left(\frac{K-\sqrt{K\pi}}{\gamma_b}, D_L^{min}\right)$ and $D_L^{up} = \min\left(\frac{K+\sqrt{K\pi}}{\gamma_b}, D_L^{max}\right)$. The optimization problem can be rewritten as

$$\max_{D_L, P_{TX}, P_{RX}, P^{tra}} \frac{P^{tra} BK(D_L \gamma_b \sqrt{\frac{1}{K\pi}} + 1 - \sqrt{\frac{K}{\pi}})}{2D_L(P_{TX} + P_{RX} + P_c)},$$

$$s.t. \quad \frac{D_L m\zeta^2}{\zeta + 1} \le \frac{2\epsilon^2}{\max\{1 - P^{tra}, P^{tra}\}^2}, \tag{31a}$$

$$\overline{AoI} \le \delta \tag{31b}$$

$$D_L^{low} < D_L < D_L^{up}, \tag{31c}$$

$$0 < P_{TX} < P_{TX}^{max}, \tag{31d}$$

$$0 < P_{RX} < P_{RX}^{max}, \tag{31e}$$

$$0 \le P^{tra} \le P_{max}^{tra}. \tag{31f}$$

The reformulated problem (31) is a non-convex problem, which stems from variable coupling and the non-convexity of objective function (30) with respect to P_{RX} and D_L. The first-order and second-order derivatives of the objective function (30) with respect to P_{RX} and D_L are given by

$$\frac{\partial \eta_a}{\partial P_{RX}} = \frac{N(P_{RX})}{2D_L [v(P_{RX})]^2}, \tag{32}$$

$$\frac{\partial^2 \eta_a}{\partial P_{RX}^2} = \frac{N'(P_{RX}) v(P_{RX}) - 2N(P_{RX}) v'(P_{RX})}{(2D_L) \cdot [v(P_{RX})]^3}, \tag{33}$$

$$\frac{\partial \eta_a}{\partial D_L} = \frac{P^{tra} BK}{2D_L^2 (P_{TX} + P_{RX} + P_c)} (\sqrt{\frac{K}{\pi}} - 1), \tag{34}$$

$$\frac{\partial^2 \eta_a}{\partial D_L^2} = \frac{P^{tra} BK}{D_L^3 (P_{TX} + P_{RX} + P_c)} (1 - \sqrt{\frac{K}{\pi}}), \tag{35}$$

respectively, where

$$v(P_{RX}) = (P_{TX} + P_{RX} + P_c) \left(\phi P_{RX} |H_{RXRX}|^2 + \sigma_{RX}^2\right), \tag{36}$$

$$N(P_{RX}) = -P^{\text{tra}} BK \Bigg[D_L P_{TX} |H_{TR}|^2 \sqrt{\frac{1}{K\pi}} \times \left(\phi |H_{RXRX}|^2 (P_{TX} + 2P_{RX} + P_c) + \sigma_{RX}^2\right) + \left(\phi P_{RX} |H_{RXRX}|^2 + \sigma_{RX}^2\right)^2 \left(1 - \sqrt{\frac{K}{\pi}}\right) \Bigg], \tag{37}$$

Here, we note that when setting $1 - \sqrt{\frac{K}{\pi}} < 0$, $\frac{\partial^2 \eta_a}{\partial D_L^2} < 0$, which indicates that the objective function is a concave function of D_L. Moreover, we observe that the first term inside the brackets of $N(P_{RX})$ is positive, the second term is negative, and the sign of $N(P_{RX})$ is uncertain. Consequently, the signs of both the first and second derivatives of the objective function with respect to P_{RX} are difficult to determine. The overall concavity depends entirely on the relative magnitudes of other parameters in the system. Therefore, the objective function is neither convex nor concave with respect to P_{RX}. Its convexity changes with variations in P_{RX}, resulting in inflection points on its graph. Furthermore, the optimizing variables are mutually coupled in both the objective function and the constraints. Based on the above analysis, the optimization problem (31) is a non-convex problem.

B. Solution of the Optimization Problem

For this non-convex optimization problem (31), we use the interior point method with the filter line search and inertia correction to solve it. Inertia correction intelligently introduces a diagonal correction term when encountering Hessian matrix with negative eigenvalues. This term is precisely sized to eliminate negative eigenvalues, ensuring the modified matrix remains positive definite, leading to accurate and high-quality search directions. Filter Line Search accepts iteration points that sufficiently improve either the objective function value or the constraint violation, which gives the algorithm greater freedom to explore complex non-convex regions. This search strategy enhances robustness, enabling the algorithm to handle complex non-convex problems more stably and reliably.

Compared to the traditional interior point method, the filter line search and inertia correction enable interior point method to possess a more powerful and robust framework for addressing challenges arising from non-convexity. Consequently, this makes the algorithm more likely to converge to a significant and high-quality local solution, avoiding premature termination or entrapment at a poor stationary point. Moreover, compared to exhaustive search, our proposed interior point method can achieve similar accuracy with significantly fewer computational iterations.

The interior point method with the filter line search and inertia correction for covert energy efficiency maximization is shown as Algorithm 1. The method first introduces a barrier function to convert inequality constraints into an additional term in the objective function, thus preventing the solution from going out of bounds. Specifically, for the optimizing variable vector $\mathbf{x} = [D_L, P_{TX}, P_{RX}, P^{tra}]^T$, all constraints in the optimization problem (31) are rewritten in standard forms $g_i(\mathbf{x}), \forall i \in \{1, 2, ..., 10\}$, which are given by

$$g_1(\mathbf{x}) = \frac{D_L m\zeta^2}{\zeta + 1} - \frac{2\epsilon^2}{\max\{1 - P^{tra}, P^{tra}\}^2} \le 0, \tag{38a}$$

$$g_2(\mathbf{x}) = \overline{AoI} - \delta \le 0, \tag{38b}$$

$$g_3(\mathbf{x}) = D_L^{low} - D_L \le 0, \quad g_4(\mathbf{x}) = D_L - D_L^{up} \le 0, \tag{38c}$$

$$g_5(\mathbf{x}) = -P_{TX} \le 0, \qquad g_6(\mathbf{x}) = P_{TX} - P_{TX}^{max} \le 0, \tag{38d}$$

$$g_7(\mathbf{x}) = -P_{RX} \le 0, \qquad g_8(\mathbf{x}) = P_{RX} - P_{RX}^{max} \le 0, \tag{38e}$$

$$g_9(\mathbf{x}) = -P^{tra} \leq 0, \qquad g_{10}(\mathbf{x}) = P^{tra} - P^{tra}_{max} \leq 0, \quad (38f)$$

where $g_1(\mathbf{x})$ in (38a) corresponds to covert requirement constraint (31a), $g_2(\mathbf{x})$ in (38b) corresponds to AoI constraint (31b), $g_3(\mathbf{x})$ and $g_4(\mathbf{x})$ in (38c) correspond to block length constraint (31c), $g_5(\mathbf{x})$ and $g_6(\mathbf{x})$ in (38d) correspond to transmit power constraint of TX (31d), $g_7(\mathbf{x})$ and $g_8(\mathbf{x})$ in (38e) correspond to transmit power constraint of RX (31e), $g_9(\mathbf{x})$ and $g_{10}(\mathbf{x})$ in (38f) correspond to prior transmission probability constraint (31f). Thus, the new objective function obtained by introducing the barrier function is given by

$$f(\mathbf{x}) = -\eta_a(\mathbf{x}) - t \cdot \sum_{i=1}^{10} \ln(-g_i(\mathbf{x})), \quad (39)$$

where $t > 0$ is the barrier parameter. The smaller t is, the stronger the restriction imposed by the barrier function on the constraint boundary, and the closer the solution approaches the optimal solution of the original problem. When t approaches 0, the solution converges to the solution of the original problem.

Next, we construct the Lagrangian function by introducing Lagrange multipliers, and further obtain the Karush-Kuhn-Tucker (KKT) conditions. The Lagrangian function of optimization problem (31) is given by

$$\mathcal{L}(\mathbf{x}, \lambda) = -\eta_a(\mathbf{x}) + \sum_{i=1}^{10} \lambda_i g_i(\mathbf{x}), \quad (40)$$

where $\lambda = [\lambda_1, \lambda_2, ..., \lambda_{10}]^T$ is the Lagrange multipliers corresponding to the inequality constraints in (38), satisfying $\lambda_i > 0, \forall i \in \{1, 2, \ldots, 10\}$. The gradient conditions in the KKT conditions involve taking the partial derivative with respect to each variable and setting it equal to zero, which are given by

$$\nabla_x \mathcal{L}(\mathbf{x}, \lambda) = -\nabla_x \eta_a(\mathbf{x}) + \sum_{i=1}^{10} \lambda_i \cdot \nabla_x g_i(\mathbf{x}) = 0, \quad \forall x \in \mathbf{x}. \quad (41)$$

Moreover, the perturbed complementary slackness conditions with the barrier parameter t in KKT conditions are given by

$$-\lambda_i g_i(\mathbf{x}) = t, \quad \forall i \in \{1, 2, \ldots, 10\}, \quad (42)$$

where t approaches 0 as the algorithm converges, at which point the standard complementary slackness conditions $\lambda_i g_i(\mathbf{x}) = 0$ are satisfied. By combining the gradient condition (41), complementary slackness condition (42), feasibility condition (38), and $\lambda_i > 0$, we obtain the non-linear system of equations constituting the KKT conditions. We can then further employ numerical methods to solve the KKT conditions to get the search direction. In each iteration, the optimizing variables $\mathbf{x} = [D_L, P_{TX}, P_{RX}, P^{tra}]^T$ and Lagrange multipliers λ_i are updated, and the barrier parameter t is gradually reduced to approach the actual constraint boundaries until the convergence condition is met, thereby outputting the optimal solution.

Algorithm 1 Interior Point Method for Covert Energy Efficiency Maximization

Require: Objective function $f(\mathbf{x})$ in (39), constraint functions $g_i(\mathbf{x})$ in (38), initial point $\mathbf{x}_0$, Lagrange multipliers λ_i for constraints, tolerance ϵ_{tol}, barrier parameter t, barrier update factor t_{factor}.
Ensure: Optimal solution $\mathbf{x}_{\text{opt}}$, maximum covert energy efficiency $\eta_{\max}$.

Initialize: Set iteration index $k = 0$. Set $\mathbf{x} = \mathbf{x}_0$, $\lambda_i = 1$, $t = 1$, $\epsilon_{tol} = 0.00001$, $\mu_{\text{factor}} = 0.2$. Initialize filter $\mathcal{F}_k = \{(\theta_i(\mathbf{x}_k), f(\mathbf{x}_k))\}$, where $\theta_i(\mathbf{x}_k) = \|\lambda_i g_i(\mathbf{x}_k)\|$ is the constraint violation. Set line search backtrack factor $\beta \in (0, 1)$, filter margins γ_θ, γ_f.
for iter = 1, 2, 3... **do**
  Compute gradient $\nabla(-\eta_a(\mathbf{x}_k))$ and Hessian $\nabla^2(-\eta_a(\mathbf{x}_k))$ for $-\eta_a(\mathbf{x}_k)$,
  Compute $\nabla g_i(\mathbf{x}_k)$ and $\nabla^2 g_i(\mathbf{x}_k)$,
  Construct the KKT system for the search direction,
  Compute Hessian matrix $H_k = \nabla^2(-\eta_a(\mathbf{x}_k)) + \sum_i \lambda_i \nabla^2 g_i(\mathbf{x}_k)$,
  Compute grad = $\nabla(-\eta_a(\mathbf{x}_k)) + \sum_i \lambda_i \nabla g_i(\mathbf{x}_k)$,
  Inertia Correction: Set correction parameter μ_k,
  while Hessian matrix H_k is not positive definite **do**
    Modify the Hessian matrix $H_k = H_k + \mu_k I$,
  end while
  Solve the KKT system to get search direction $\Delta\mathbf{x}_k$,$\Delta\lambda_i$,
  Perform filter line search and determine step size α_k,
  loop
    Compute trial point $\mathbf{x}_{\text{trial}} = \mathbf{x}_k + \alpha_k \Delta\mathbf{x}_k$,
    Compute trial constraint violation $\theta_{i,\text{trial}} = \theta_i(\mathbf{x}_{\text{trial}})$ and trial objective value $f_{\text{trial}} = f(\mathbf{x}_{\text{trial}})$,
    Set flag=false, where flag indicates rejection status,
    for each $(\theta_{i,j}, f_j)$ in $\mathcal{F}_k$ **do**
      if $\theta_{i,\text{trial}} \geq (1-\gamma_\theta)\theta_{i,j}$ and $f_{\text{trial}} \geq f_j - \gamma_f$ **then**
        flag = true, **break**
      end if
    end for
    if not flag **then**
      break loop,
    else
      $\alpha_k = \beta \cdot \alpha_k$,
    end if
  end loop
  Update variables $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \Delta\mathbf{x}_k$,
  Update Lagrange multipliers $\lambda_i = \lambda_i + \alpha_k \Delta\lambda_i$,
  Update filter $\mathcal{F}_{k+1}$ by adding $(\theta_i(\mathbf{x}_{k+1}), f(\mathbf{x}_{k+1}))$,
  Update barrier parameter $t = t_{\text{factor}} \cdot t$, $k = k + 1$,
  if $\|\Delta\mathbf{x}_k\| < \epsilon_{\text{tol}}$ **then**
    break
  end if
end for
return $\mathbf{x}_{\text{opt}} = \mathbf{x}_{k+1}$, $\eta_{\max} = \eta_a(\mathbf{x}_{\text{opt}})$.

C. Algorithm Analysis

For each iteration with a given barrier parameter t, the algorithm updates the optimization variable $\mathbf{x}_{k+1}$. The set of all optimization variables collectively forms a central path. As the barrier parameter t gradually decreases toward 0, the perturbed complementary slackness conditions $-\lambda_i g_i(\mathbf{x}) = t$ in (42) ultimately converge to the standard complementary

TABLE II: SYSTEM PARAMETERS

Parameters	Values
System bandwidth (B)	1 MHz
Information amount in each short-packet (K)	10 nats
Noise variance at RX and WD (σ_{RX}^2, σ_{WD}^2)	-174 dBm
Maximum transmit power of TX and RX (P_{TX}^{max}, P_{RX}^{max})	2 mW
Circuit power consumption (P_c)	0.05 mW
Fading parameter (m)	1
Minimum block length (D_L^{min})	10
Maximum block length (D_L^{max})	1000
Maximum acceptable average AoI (δ)	10 ms

relaxation condition $\lambda_i g_i(\mathbf{x}) = 0$. Furthermore, since the objective function and each constraint function in (38) are all continuously differentiable with respect to variables $\mathbf{x}$, the algorithm guarantees convergence to the solution satisfying the KKT conditions of the original optimization problem (31).

Additionally, filter line search and inertia correction ensure the algorithm's global convergence. Filter line search checks both the objective function and constraint violation to guarantee each iteration makes meaningful progress, preventing the algorithm accepting points that perform worse in both dimensions and avoiding cycles. Inertia correction guarantees that the computed search direction $\Delta\mathbf{x}_k$ is a positive descent direction to help the algorithm escape saddle point regions, enabling filter line search to find an acceptable step size α_k and thus ensuring robustness for non-convex problem (31). Therefore, with the filter line search and inertia correction, the algorithm converges from any initial feasible point $\mathbf{x}_0$ to the solution satisfying the KKT conditions of the original problem. Moreover, Fig. 8 in Numerical Results demonstrates the algorithm's convergence.

In general, the solution obtained by the algorithm is a reliable, high-quality approximation close to the global optimization. Furthermore, we compare the algorithm with exhaustive search in Fig. 9 of Numerical Results and find that it achieves nearly identical performance as exhaustive search, which demonstrates its effectiveness.

The computational complexity of the algorithm is primarily determined by the number of iterations and the computational cost per iteration. The computational cost per iteration mainly stems from constructing and solving the KKT system to obtain the search direction. The complexity per iteration is $O(a^{1/2}b^2(a+b))$ [35], where a is the variable dimension and b is the number of inequality constraints. For our optimization problem, due to its low variable dimension and the finite number of inequality constraints in (38), each iteration has a constant-level computation. In this case, the overall computational complexity is primarily determined by the number of iterations N_{ite} and is denoted as $O(N_{\text{ite}})$, which indicates the algorithm has low computational complexity.

VI. Numerical Results

This section provides extensive numerical results to explore the influence of various system parameters on the CEE. The maximization of CEE is determined by the optimization problem formulated in Equations (29) and (30). Without loss of generality, we set TX's transmit power and RX's interference power to 0.2 and 2 mW, respectively, and the values of parameters are set to 1 for channel gains. Unless otherwise stated, we set the system parameters listed in Table II.

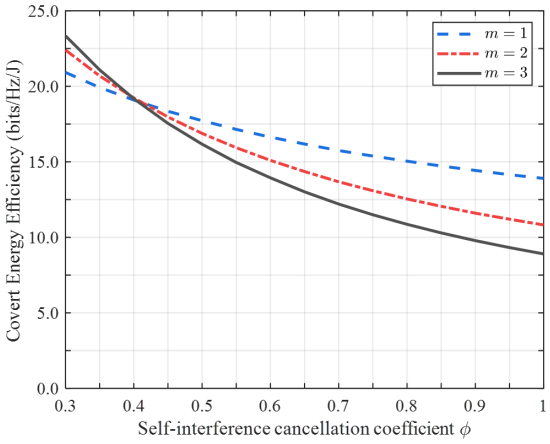
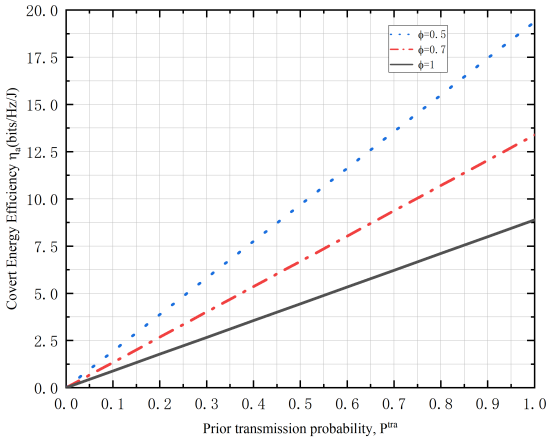


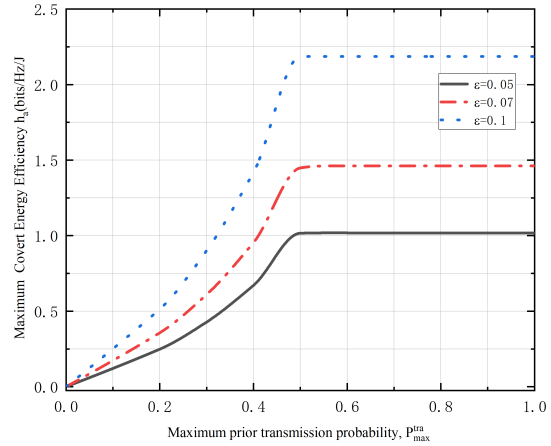
Fig. 3: The influence of self-interference cancellation coefficient ϕ on the simulation results of covert energy efficiency under different values of fading parameter m.

In Fig.3, we explore the impact of self-interference cancellation coefficient ϕ and fading parameter m on the simulation results of CEE. From the figure, it can be observed that the CEE decreases as the self-interference cancellation coefficient increases. This is because the larger ϕ causes more severe self-interference, which leads to the lower effective covert rate at RX, and thus decreases CEE. Moreover, we observe that CEE increases as the fading parameter m increases at the small self-interference cancellation coefficient ϕ, and the CEE decreases with m when ϕ is large. The reason is as follows. The larger fading parameter m means the lower small-scale fading degrees of self-interference channel and short-packet communication channel from TX to RX. When ϕ is small, the self-interference is weak, and CEE is mainly influenced by the quality of short-packet communication channel. The increase in fading parameter m causes the lower fading degree of short-packet communication channel, which leads to the increase in CEE. However, with the increase in ϕ, the self-interference enhances, and CEE is increasingly affected by self-interference channel. In this case, the quality of the self-interference channel increases with m, which leads to the higher decoding error probability, and thus CEE decreases.

In Fig. 4, we explore the impact of prior transmission probability on CEE and maximum CEE. Fig. 4(a) summarizes how the CEE changes with the prior transmission probability under the settings of $P_{TX} = 0.2$ mW, $P_{RX} = 2$ mW, and $D_L = 200$. From the figure, it can be observed that as the prior transmission probability increases, η_a increases. This can be explained as follows. The higher prior transmission probability indicates that TX transmits more information to RX, leading to an increase in η_a. Moreover, for the same prior transmission probability, as the self-interference cancellation coefficient ϕ

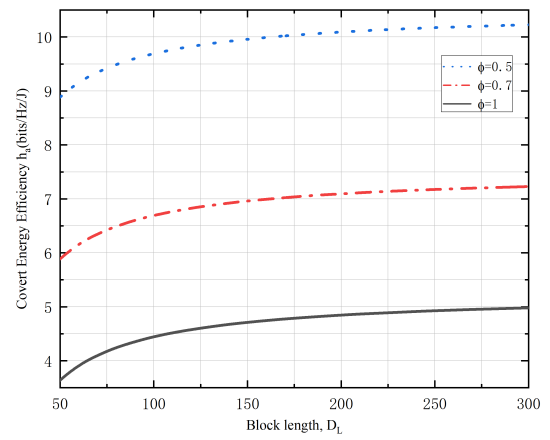


(a) Covert energy efficiency versus the prior transmission probability.

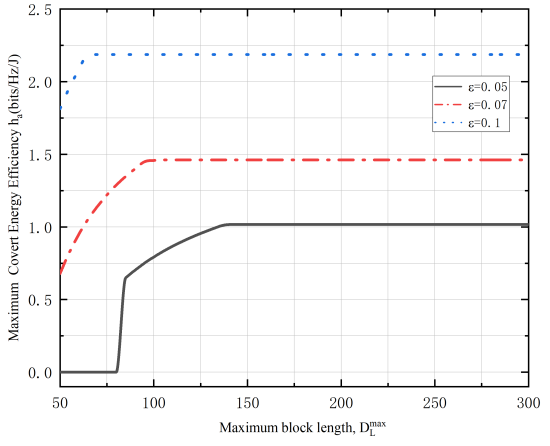


(b) Maximum covert energy efficiency versus the maximum prior transmission probability.

Fig. 4: The influence of prior transmission probability on covert energy efficiency. (a) η_a vs. P^{tra}. (b) maximum η_a vs. P^{tra}_{max}.



(a) Covert energy efficiency versus the block length.

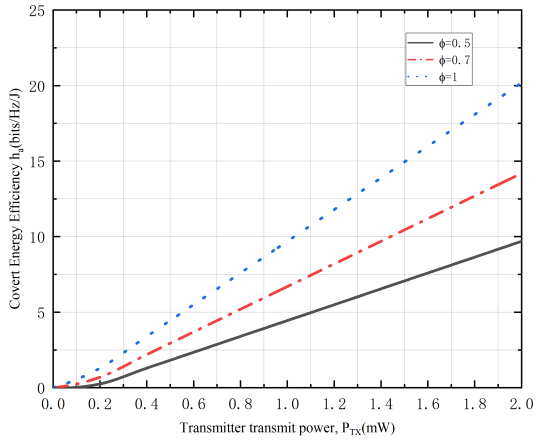


(b) Maximum covert energy efficiency versus the maximum block length.

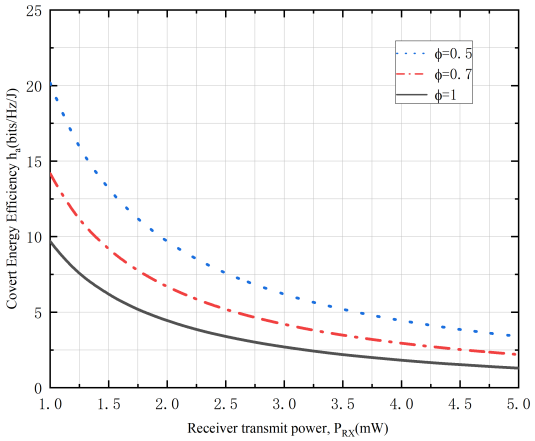
Fig. 5: The influence of block length on covert energy efficiency. (a) η_a vs. D_L. (b) maximum η_a vs. D_L^{max}.

decreases, the CEE increases. This is because a smaller self-interference coefficient results in less self-interference in the full-duplex system, thereby increasing the CEE. Fig. 4(b) summarizes how the maximum CEE changes with the maximum prior transmission probability. From the figure, it can be observed that as the maximum prior transmission probability increases, η_a increases and then remains constant after reaching a certain value. This can be explained as follows: CEE is an increasing function of the prior transmission probability, and as the maximum prior transmission probability increases, the optimal prior transmission probability equals the maximum prior transmission probability, leading to an increase in CEE. However, to ensure covert requirements, when the maximum prior transmission probability further increases, the optimal prior transmission probability remains constant. Under the same maximum prior transmission probability, as the covert constraint becomes looser, the optimal transmission power can be larger, leading to an increase in maximum CEE.

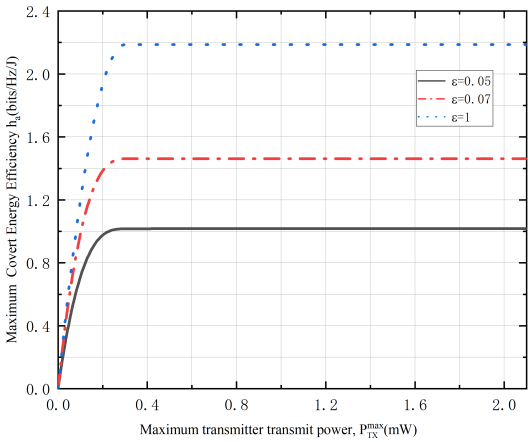
In Fig. 5, we investigate the influence of block length on CEE and maximum CEE. Fig. 5(a) summarizes how the CEE changes with block length under the settings of P_{TX} = 0.2 mW, P_{RX} = 2 mW, and P^{tra} = 0.5. From the figure, it can be observed that as the block length increases, η_a also increases. This can be explained as follows. A larger block length reduces decoding error probability, leading to an increase in η_a. Furthermore, under the same block length, a smaller self-interference cancellation coefficient ϕ results in a larger CEE. This is because a smaller self-interference coefficient leads to less self-interference in the full-duplex system, thereby increasing the CEE. Fig. 5(b) summarizes how the maximum CEE changes with the maximum block length. From the figure, it can be observed that as the maximum block length increases, η_a first increases and then remains constant after reaching



(a) Covert energy efficiency versus the transmit power of TX.

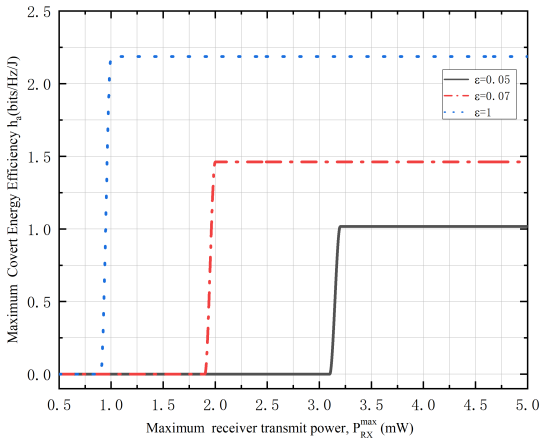


(a) Covert energy efficiency versus the receiver transmission power.



(b) Maximum covert energy efficiency versus the maximum transmit power of TX.

Fig. 6: The influence of TX's transmit power on covert energy efficiency. (a) η_a vs. P_{TX}. (b) maximum η_a vs. P_{TX}^{max}.



(b) Maximum covert energy efficiency versus the maximum receiver transmission power.

Fig. 7: The influence of RX's transmit power on covert energy efficiency. (a) η_a vs. P_{RX}. (b) maximum η_a vs. P_{RX}^{max}.

a certain value. This can be explained as follows: The CEE increases with block length, and as the maximum block length increases, the optimal block length equals the maximum block length, leading to an increase in CEE. However, when the block length increases to a certain value, further increases in block length fail to meet the information age constraint. Therefore, when the maximum block length further increases, the optimal block length remains constant. It is worth noting that the transmitter needs to use a larger transmission power with a small block length to meet the information age requirement. However, when the covert constraint is very strict, this may lead to a decrease in covertness, thus failing to meet the covert constraint. As a result, η_a becomes a constant 0. Under the same maximum block length, as the covert constraint becomes looser, the optimal transmission power can increase, leading to a higher maximum CEE.

Fig. 6 investigates the influence of transmit power of TX on CEE and maximum CEE. Fig. 6(a) summarizes how CEE changes with transmit power of TX under the settings of $D_L = 200$, $P_{RX} = 2$ mW, and $P^{tra} = 0.5$. From the figure, it can be observed that as the transmit power of TX increases, η_a increases. This can be explained as follows. Increasing P_{TX} significantly improves the SINR and the effective covert rate. Although the power consumption also increases, since the transmit power of TX is relatively smaller compared to the sum of transmit power of RX P_{RX} and circuit power P_c, the gain in covert rate outweighs the cost in power consumption, resulting in a corresponding increase in η_a. Moreover, consistent with previous results, a smaller self-interference cancellation coefficient ϕ leads to a higher CEE due to reduced self-interference. Fig. 6(b) summarizes how the maximum CEE changes with the maximum transmit power of TX. From the figure, it can be observed that as the maximum transmit power of TX increases, η_a increases and then remains

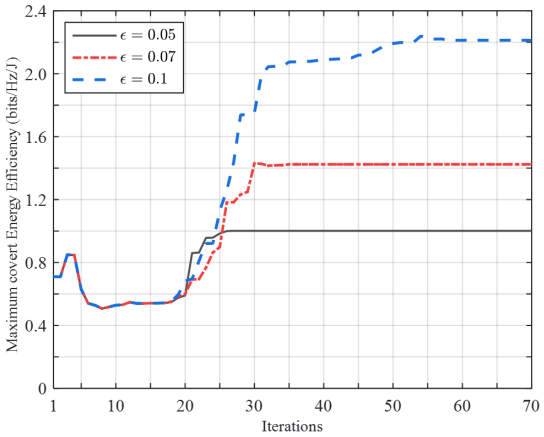


Fig. 8: Maximum covert energy efficiency versus iterations.

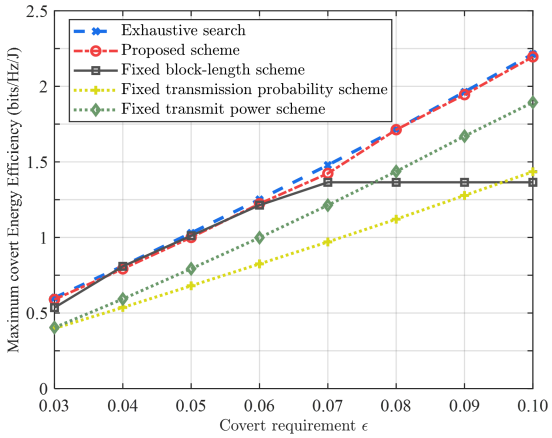


Fig. 9: Maximum covert energy efficiency versus covert requirement ϵ with different schemes.

constant after reaching a certain value. This can be explained as follows. When P_{TX}^{max} is small, as the maximum transmit power of TX increases, the optimal transmit power of TX equals the maximum transmit power of TX. As a result, CEE increases. However, when the maximum transmit power of TX further increases to a certain threshold, the transmit power becomes constrained by the covertness requirements to avoid detection. To ensure covert requirements, further increasing the power budget does not help, and the optimal transmit power of TX remains constant, resulting in the maximum CEE remaining unchanged. Under the same maximum transmit power of TX, as the covert constraint becomes looser, the optimal prior transmission probability can be larger and the receiver transmission power can be smaller, leading to a higher maximum CEE.

In Fig. 7, we explore the effect of the transmit power of RX on the CEE and the maximum CEE. Fig. 7(a) summarizes how the CEE varies with the transmit power of receiver in the settings of $D_L = 200$, $P_{TX} = 0.2$mW and $P^{tra} = 0.5$. From the figure, it can be observed that η_a decreases as transmit power of RX increases. This can be explained as follows. Increasing transmit power of RX increases the total energy consumption and causes stronger self-interference, which decreases the effective covert rate. Both factors contribute to the decrease in CEE. Moreover, a smaller self-interference cancellation coefficient ϕ leads to a higher CEE due to the less self-interference generated by the full-duplex system. Fig. 7(b) summarizes how the maximum CEE changes with the maximum transmit power of RX. It can be observed from the figure that as the maximum receiver transmission power increases, η_a is initially zero and then remains constant after reaching a certain value. This can be explained as follows. When the covert constraint is relatively strict, a higher transmit power of RX is required to meet higher covert requirements. Therefore, when the maximum transmit power of RX is relatively low, the receiver cannot generate sufficient artificial noise to hide the transmission and satisfy the covert requirement, resulting in η_a being zero. As the maximum transmit power of RX increases to satisfy the covert requirement, the optimal transmit power of RX equals the minimum necessary jamming power to meet the constraint for the CEE. Therefore, further increasing the maximum transmit power of RX does not change the optimal receiver transmission power, resulting in η_a remaining constant. Under the same maximum transmit power of RX, as the covert constraint becomes loose, the optimal prior transmission probability and the optimal transmit power of TX can be larger, while the optimal transmit power of RX can be smaller, leading to a larger maximum CEE.

In Fig. 8, we show the convergence of our algorithm under different covert requirement ϵ. It can be seen that the maximum CEE can converge within dozens of iterations, which indicates the high convergence performance and low complexity of the proposed algorithm. Moreover, as ϵ increases, the number of iterations to reach convergence increases. This is because the larger ϵ has looser covert constraints, which expands the search range of the algorithm, resulting in the need for more iterations.

In order to evaluate the performance of our scheme, we employ exhaustive search and three other schemes to compare these schemes with the proposed scheme, as shown in Fig. 9. The first of the other three schemes is the fixed block length scheme for the settings of $D_L = 200$ in [24], which optimizes the transmit probability and power without considering block length. The second is the fixed transmission probability scheme for the settings of $P^{tra} = 0.4$ in [36], which optimizes the block length and transmit power, but does not consider transmission probability. The third is the fixed transmit power scheme for the settings of $P_{TX} = 0.2$ mW in [22], which optimizes the transmission probability, but power control is not taken into account.

From Fig. 9, it is clearly observed that the proposed scheme can achieve similar performance to the exhaustive search while outperforming other schemes. Although the exhaustive search has the optimal performance on maximum CEE, its computational cost is extremely high and rises exponentially with accuracy. For example, assuming each variable in the optimization problem has hundreds of sampling points, the exhaustive search would require hundreds of millions of iterations. In contrast, the proposed scheme converges after just dozens of

iterations from Fig. 8. Therefore, the proposed scheme has a computational complexity significantly lower than that of the exhaustive search, which indicates its greater effectiveness in practical applications compared to other schemes. Moreover, Fig. 9 shows that as ϵ increases, the maximum CEE increases for the proposed, exhaustive search, fixed transmission probability, and fixed transmit power schemes. This is because a larger ϵ indicates a more relaxed covert constraint, which means that TX can use larger transmit block length and power, resulting in an increase in the maximum CEE. For the fixed block length scheme, the maximum CEE first increases similarly to the proposed scheme, then remains unchanged. This can be explained as follows. When ϵ is relatively low, adjusting transmission probability and power can achieve favorable CEE performance while satisfying information age and covert constraint. However, when ϵ increases to a certain value, the information age constraint limits the transmission probability and power, causing the CEE to remain unchanged. The proposed scheme is significantly better than these schemes with fixed system parameters, indicating the necessity of joint optimization and showing that proper system parameter setting can enhance CEE performance.

VII. CONCLUSION

This paper explored the CEE maximization in a full-duplex wireless system by the joint optimization of the block length, prior transmission probability and transmit powers. Specifically, we first modeled the covert constraint condition, AoI and CEE. Then, we formulated the CEE maximization as a constrained optimization problem, and further solved it using the interior point method. Numerical results show that when the maximum transmit powers of TX and RX increase up to a threshold, the maximum covert energy efficiency remains unchanged. A proper setting of system parameters (e.g., block length, prior transmission probability) can improve covert energy efficiency performance.

In future studies, our work can be extended in the following directions. First, the single-antenna system considered in this paper can be extended to a multi-antenna configuration with beamforming capabilities. By employing beamforming, the transmitter can directionally transmit covert signals to the intended receiver, and thus this can reduce unintended signal leakage to the warden. Simultaneously, the receiver can generate artificial noise to directionally jam the warden. This approach is expected to significantly improve both the covertness and energy efficiency of the system. Second, the proposed system can be further developed into an integrated sensing and communication scenario. By designing artificial noise in the form of radar beams, the receiver can not only interfere with the warden but also sense the warden's position or activity status. Such a design would enable more proactive and robust covert communication strategies in dynamic environments.

REFERENCES

[1] D. C. Nguyen, M. Ding, P. N. Pathirana, A. Seneviratne, J. Li, D. Niyato, O. Dobre, and H. V. Poor, "6G Internet of Things: A comprehensive survey," *IEEE Internet of Things Journal*, vol. 9, no. 1, pp. 359–383, 2022.

[2] M. Harounabadi and T. Heyn, "Toward integration of 6G-NTN to terrestrial mobile networks: Research and standardization aspects," *IEEE Wireless Communications*, vol. 30, no. 6, pp. 20–26, 2023.

[3] H. Chamkhia, A. Erbad, A. Mohamed, A. R. Hussein, A. K. Al-Ali, and M. Guizani, "Stochastic geometry-based physical-layer security performance analysis of a hybrid NOMA-PDM-based IoT system," *IEEE Internet of Things Journal*, vol. 11, no. 2, pp. 2027–2042, 2024.

[4] S. A. Soleymani, S. Goudarzi, M. H. Anisi, H. Cruickshank, A. Jindal, and N. Kama, "TRUTH: Trust and authentication scheme in 5G-IIoT," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 1, pp. 880–889, 2023.

[5] T. B. da Silva, R. P. dos Santos Chaib, C. S. Arismar, R. da Rosa Righi, and A. M. Alberti, "Toward future Internet of Things experimentation and evaluation," *IEEE Internet of Things Journal*, vol. 9, no. 11, pp. 8469–8484, 2022.

[6] D. Xu, H. Zhao, and H. Zhu, "Resource allocation for secure short packet communications in wireless powered IoT networks," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 8, pp. 11000–11005, 2023.

[7] D. Xu and H. Zhu, "Energy efficient resource allocation for wireless powered short packet communication networks," *IEEE Communications Letters*, vol. 28, no. 8, pp. 1880–1884, 2024.

[8] X. Chen, J. An, Z. Xiong, C. Xing, N. Zhao, F. R. Yu, and A. Nallanathan, "Covert communications: A comprehensive survey," *IEEE Communications Surveys Tutorials*, vol. 25, no. 2, pp. 1173–1198, 2023.

[9] D. Wang, P. Qi, Y. Zhao, C. Li, W. Wu, and Z. Li, "Covert wireless communication with noise uncertainty in space-air-ground integrated vehicular networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 3, pp. 2784–2797, 2022.

[10] Z. Duan, X. Yang, Y. Gong, D. Wang, and L. Wang, "Covert communication in uplink NOMA systems under channel distribution information uncertainty," *IEEE Communications Letters*, vol. 27, no. 5, pp. 1282–1286, 2023.

[11] K. Shahzad and X. Zhou, "Covert wireless communications under quasi-static fading with channel uncertainty," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 1104–1116, 2021.

[12] J. Hu, K. Shahzad, S. Yan, X. Zhou, F. Shu, and J. Li, "Covert communications with a full-duplex receiver over wireless fading channels," in *2018 IEEE International Conference on Communications (ICC)*, pp. 1–6, IEEE, 2018.

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

[14] F. Shu, T. Xu, J. Hu, and S. Yan, "Delay-constrained covert communications with a full-duplex receiver,"

IEEE Wireless Communications Letters, vol. 8, no. 3, pp. 813–816, 2019.

[15] M. Zheng, A. Hamilton, and C. Ling, "Covert communications with a full-duplex receiver in non-coherent Rayleigh fading," *IEEE Transactions on Communications*, vol. 69, no. 3, pp. 1882–1895, 2020.

[16] R. Xu, L. Guan, Y. Zhao, Z. Li, and D. Wang, "Robust power and position optimization for the full-duplex receiver in covert communication," in *2021 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, IEEE, 2021.

[17] Y. Zhang, L. Yang, X. Li, K. Guo, and H. Liu, "Covert communications for STAR-RIS-assisted industrial networks with a full duplex receiver and RSMA," *IEEE Internet of Things Journal*, vol. 11, no. 12, pp. 22483–22493, 2024.

[18] Z. Guo, S. Zhao, J. Wang, H. Lit, and Y. Shen, "Optimal location design for UAV covert communications with a full-duplex receiver," in *2022 International Conference on Networking and Network Applications (NaNA)*, pp. 35–40, IEEE, 2022.

[19] R. Sun, B. Yang, S. Ma, Y. Shen, and X. Jiang, "Covert rate maximization in wireless full-duplex relaying systems with power control," *IEEE Transactions on Communications*, vol. 69, no. 9, pp. 6198–6212, 2021.

[20] R. Zhang, X. Chen, M. Liu, N. Zhao, X. Wang, and A. Nallanathan, "UAV relay assisted cooperative jamming for covert communications over Rician fading," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 7, pp. 7936–7941, 2022.

[21] M. E. Ildiz, O. T. Yavascan, E. Uysal, and O. T. Kartal, "Pull or wait: How to optimize query age of information," *IEEE Journal on Selected Areas in Information Theory*, vol. 4, pp. 794–807, 2023.

[22] Y. Wang, S. Yan, W. Yang, and Y. Cai, "Covert communications with constrained age of information," *IEEE Wireless Communications Letters*, vol. 10, no. 2, pp. 368–372, 2021.

[23] Z. Tang, N. Yang, X. Zhou, and J. Lee, "Average age of information penalty of short-packet communications with packet management," in *ICC 2023-IEEE International Conference on Communications*, pp. 1670–1675, IEEE, 2023.

[24] J. Li, D. Wu, C. Yue, Y. Yang, M. Wang, and F. Yuan, "Energy-efficient transmit probability-power control for covert D2D communications with age of information constraints," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 9, pp. 9690–9704, 2022.

[25] C. Wang, Z. Li, T.-X. Zheng, D. W. K. Ng, and N. Al-Dhahir, "Intelligent reflecting surface-aided full-duplex covert communications: Information freshness optimization," *IEEE Transactions on Wireless Communications*, vol. 22, no. 5, pp. 3246–3263, 2023.

[26] S. S. Hosseini, P. Azmi, and N. Mokari, "Minimizing average age of information in reliable covert communication on time-varying channels," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 1, pp. 651–659, 2024.

[27] G. Durisi, T. Koch, J. Östman, Y. Polyanskiy, and W. Yang, "Short-packet communications over multiple-antenna Rayleigh-fading channels," *IEEE Transactions on Communications*, vol. 64, no. 2, pp. 618–629, 2016.

[28] Y. Zhang, Y. Chen, B. Yu, X. Diao, and Y. Cai, "Minimizing age of information based on predictions and short packet communications in UAV relay systems," in *2021 13th International Conference on Wireless Communications and Signal Processing (WCSP)*, pp. 1–5, 2021.

[29] W. Yang, X. Lu, S. Yan, F. Shu, and Z. Li, "Age of information for short-packet covert communication," *IEEE Wireless Communications Letters*, vol. 10, no. 9, pp. 1890–1894, 2021.

[30] Z. Fu, B. Ju, J. Moon, S. Hwang, and I. Lee, "Covert communications in two-way relay systems with energy harvesting," in *2022 13th International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 985–989, 2022.

[31] Y. Wang, S. Yan, W. Yang, Y. Huang, and C. Liu, "Energy-efficient covert communications for bistatic backscatter systems," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 3, pp. 2906–2911, 2021.

[32] X. Lu, W. Yang, S. Yan, Z. Li, and D. W. K. Ng, "Covertness and timeliness of data collection in UAV-aided wireless-powered IoT," *IEEE Internet of Things Journal*, vol. 9, no. 14, pp. 12573–12587, 2022.

[33] Y. Ma, R. Ma, Z. Lin, R. Zhang, Y. Cai, W. Wu, and J. Wang, "Improving age of information for covert communication with time-modulated arrays," *IEEE Internet of Things Journal*, vol. 12, no. 2, pp. 1718–1731, 2025.

[34] G. Durisi, T. Koch, and P. Popovski, "Toward massive, ultrareliable, and low-latency wireless communication with short packets," *Proceedings of the IEEE*, vol. 104, no. 9, pp. 1711–1726, 2016.

[35] A. Meng, X. Gao, Y. Zhao, and Z. Yang, "Three-dimensional trajectory optimization for energy-constrained UAV-enabled IoT system in probabilistic LoS channel," *IEEE Internet of Things Journal*, vol. 9, no. 2, pp. 1109–1121, 2022.

[36] M. Wang, Y. Yao, B. Xia, Z. Chen, and J. Wang, "Covert and reliable short-packet communications over fading channels against a proactive warder: Analysis and optimization," *IEEE Transactions on Wireless Communications*, vol. 23, no. 5, pp. 3932–3945, 2024.

Yangfan Xu received the B.S. degree from Huazhong Agricultural University in 2022. He is currently pursuing a M.S. degree in computer technology at Anhui University. His research interest focuses on the covert communication in the physical layer.

Bin Yang received his Ph.D. degree in systems information science from Future University Hakodate, Japan in 2015. He was a research fellow with the School of Electrical Engineering, Aalto University, Finland, from Nov. 2019 to Nov. 2021. He is currently a professor with the School of Computer and Information Engineering, Chuzhou University, China. His research interests include unmanned aerial vehicle networks, cyber security and Internet of Things.

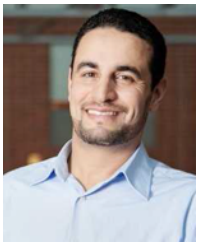
Lei Shao received the B.S. degree from Anhui University of Science and Technology in 2021. He is currently pursuing a M.S. degree in network and information security at Anhui University. His research interest focuses on the covert communication in the physical layer.

Xiuwen Sun received his Ph.D. degree in computer science from Xi'an Jiao tong University in 2019 and is now an assistant professor in the school of computer science and technology at Anhui University. His research interests are computer networks and network security.

Shikai Shen received his B.S. and M.S. Degrees from Yunnan Normal University in 1984 and from Yunnan University in 2003, respectively. He is currently a professor with the School of Information Engineering, Kunming University, Yunnan Province Key Laboratory of Smart City and Cyberspace Security, and Dianchi College, Kunming, China. Senior Member of China Computer Federation. His research interests include wireless sensor networks, network coding, the Internet of Things, etc.

Haibao Chen received his PhD in Computer Architecture at Huazhong University of Science and Technology, China. He is currently an associate professor with the School of Computer and Information Engineering, Chuzhou University, China. His research interests include covert communication and cybersecurity.

Bao Gui received the M.S. degree from Anhui University of Science and Technology, China, in 2022. Currently, he is with the School of Computer and Information Engineering, Chuzhou University, Anhui, China. His research interests include vehicular networks and wireless communication.



Tarik Taleb received the B.E. degree (with distinction) in information engineering and the M.Sc. and Ph.D. degrees in information sciences from Tohoku University, Sendai, Japan, in 2001, 2003, and 2005, respectively. He is currently a Full Professor at Ruhr University Bochum, Germany. He was a Professor with the Center of Wireless Communications, University of Oulu, Oulu, Finland. He is the founder of ICTFICIAL Oy, and the founder and the Director of the MOSA!C Lab. From October 2014 to December 2021, he was an Associate Professor with the School of Electrical Engineering, Aalto University, Espoo, Finland. Prior to that, he was working as a Senior Researcher and a 3GPP Standards Expert with NEC Europe Ltd., Heidelberg, Germany. Before joining NEC and till March 2009, he worked as an Assistant Professor with the Graduate School of Information Sciences, Tohoku University, in a lab fully funded by KDDI. From 2005 to 2006, he was a Research Fellow with the Intelligent Cosmos Research Institute, Sendai. Taleb has been directly engaged in the development and standardization of the Evolved Packet System as a member of the 3GPP System Architecture Working Group. His current research interests include AI-based network management, architectural enhancements to mobile core networks, network softwarization and slicing, mobile cloud networking, network function virtualization, software-defined networking, software-defined security, and mobile multimedia streaming.