Enhancing Wireless Communication Security with Variable Bloom Filter-Based Physical-Layer Secure Transmission

Anqi Huo, Guyue Li, *Member, IEEE,* Lilin Yang, Zilong Liu, *Senior Member, IEEE,* Aiqun Hu, *Senior Member, IEEE*

***Abstract*—This paper studies one-time pad (OTP) secure communication by leveraging the unpredictable physical layer channel characteristics. Existing OTP schemes based on physical-layer key generation (PKG) require additional transmission overhead of information reconciliation and may face security threats of information leakage under slow-varying channels. To tackle these challenges, we investigate a fault-tolerant privacy amplification method through variable bloom filters to address the underlying security problems. Specifically, the quantized bit sequence of the channel state information goes through a bloom filter to improve the randomness within the sequence while the parameters of bloom filter vary to avoid the correlations between adjacent sequences. We then optimize the parameters of the error-correcting code used during communication based on the position of the eavesdropper and the length of the quantized bits, thereby further enhancing the system security. Through comprehensive simulations, it is shown that our proposed approach can achieve a near-perfect pass rate in NIST randomness tests, and with a bit replacement rate around 0.45, whilst capable of resisting attacks under slow-varying channels. These results indicate that the proposed scheme significantly outperforms previous OTP secure transmission schemes.**

***Index Terms*—Physical-layer security, information reconciliation, one-time pad (OTP), secret key generation.**

I. Introduction

With the rapid development of quantum computing, public-key cryptography relying on complex mathematical problems may be vulnerable to various attacks [1]. The one-time pad (OTP) encryption scheme, as the only cryptographic system that offers perfect secrecy, has received increasing research attention. As early as 1919, Vernam proposed the OTP scheme, which encrypts each message bit with a different key bit via exclusive OR (XOR) [2]. In 1949, Shannon demonstrated that OTP provides "perfect secrecy", meaning it is unbreakable by any computational means [3]. Although the OTP scheme can offer perfect confidentiality, in practical applications, generating a random key of at least the same length as the confidential message for each encryption poses challenges in key generation, management, and distribution processes.

Physical-layer key generation (PKG) technology is a promising approach to address the challenges of key generation and storage. PKG utilizes the diversity of wireless channels and the uncertainty of channel characteristics in multipath environments to generate symmetric keys for encryption/decryption of transmitted information, thus addressing the problem of symmetric key distribution.

However, existing OTP secure transmission schemes may not be suitable for static or slowly-varying environments such as indoor Internet of Things (IoT) systems. This is because the joint design of OTP with PKG fundamentally relies on how the characteristics of the underlying wireless channel are fully exploited to provide a random source for OTP encryption. Generally, the secret key rate largely depends on the coherence time of the channel which is associated to the user mobility and changes of the surrounding environments. If users or their surrounding environments exhibit certain degree of mobility, it is reasonable to assume that the wireless channel is dynamic. However, when the devices operate in slowly time-varying channels, issues may arise, such as an uneven distribution of 0 and 1 bits in the quantized bit sequence[1], as well as high redundancy between adjacent frame sequences[2]. Such compromised randomness of the bit sequence may lead to reduced rates during multiple secure communications. Additionally, the decreased bit replacement rate of the bit sequence may result in the leakage of secret information and bit sequences, thus posing significant security risks to the communication process.

In addition, in most PKG schemes, the key generation process requires information reconciliation [4], [5] and privacy amplification steps to generate a pair of identical keys. In the information reconciliation step, syndrome exchange over a public channel is necessary to decide whether employing error detection protocol-based approaches or error correction code-based approaches. The problem is that it increases the system's communication overhead, whilst introducing the risk of information leakage. To solve this problem, researchers have begun to explore nonreconciled keys for OTP secure transmission, which do not involve information reconciliation and privacy

Manuscript received 23 October 2024; revised 1 March 2025; accepted 1 May 2025.*(Corresponding author: Guyue Li)*

Guyue Li, Anqi Huo and Lilin Yang are with the School of Cyber Science and Engineering, Southeast University, Nanjing 210096, China (e-mail: guyuelee@seu.edu.cn;anqihuo@seu.edu.cn;yanglilin@seu.edu.cn).

Zilong Liu is with the School of Computer Science and Electronic Engineering, University of Essex, UK (e-mail:zilong.liu@essex.ac.uk).

Aiqun Hu is with the National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China (e-mail: aqhu@seu.edu.cn).

Guyue Li and Aiqun Hu are also with the Purple Mountain Laboratories for Network and Communication Security, Nanjing 210096, China.

[1]The quantized bit sequence refers to the bit sequence consisting of 0s and 1s obtained after quantization

[2]The adjacent frame sequences refer to the bit sequences obtained from two consecutive quantizations

amplification steps in the key generation process [4]–[6]. Among many others, a representative work is the unidentical key-based physical-layer secure transmission (UK-PST) in [7], whereby PKG was fully exploited to achieve secure OTP transmission. However, these OTP schemes haven't addressed security issues in static or slowly changing environments. Although existing PKG works have proposed key generation protocols for static or quasi-static environments, none have simultaneously addressed the issues of overhead redundancy in the key generation process and the security challenges posed by such environments.

Against the aforementioned challenges, we consider implementing a secure communication scheme that combines PKG technology with OTP in quasi-static or low-speed environments. Our goal is to ensure that there is no additional information transmission or privacy leakage, whilst improving the randomness and bit replacement rate of the keys to address potential security vulnerabilities in the communication process. Specifically, we focus on improving the key generation rate by processing the quantized bit sequences. The challenge lies in using unidentical keys for communication encryption. Even if the quantized bit sequences between the communicating parties differ by only one bit, the avalanche effect of privacy amplification may lead to decryption failures. Therefore, novel privacy amplification methods are needed.

To proceed, we propose passing the quantized bit sequence through a specialized **Bloom filter (BF)** before incorporating it into the OTP secure transmission scheme. A BF is a space-efficient data structure to identify the membership of a set [8]. Consisting of a long binary vector and a series of random hash function, BF has also been used as part of the encoding and perturbation methods in many privacy-preserving applications [9], [10]. In 2021, [11] utilized a specially designed BF data structure that considers sequence/order information. It projects keys onto the BF, maintaining key distance information in a non-plaintext format. Unlike traditional BFs, it passes through an additional bit before mapping to the BF, and then the internal hash function of the filter maps each element from the added bit to the BF space. Inspired by this, we conceive the idea of variable bloom filter (VBF) for processing the quantized bit sequences. This approach not only ensures a low-complexity system structure but also enhances the overall security of the system. Building upon this key idea, we aim to quantitatively investigate the joint design of key generation and OTP for secure transmission in quasi-static or low-rate environments. The main contributions of this paper are as follows:

- We propose a novel OTP secure transmission scheme named variable bloom filter-based physical-layer secure transmission (VBF-PST), by utilizing a VBF to control the randomness and consistency of the output bit sequences within a stable range. Our proposed scheme enhances the National Institute of Standards and Technology (NIST) test pass ratio and the bit replacement rate of bit sequences between adjacent frames, enabling fast and secure OTP transmission in quasi-static or low-rate environments.
- We devise an efficient algorithm to optimize the parameters of error correction code (ECC) which is used by the communicating parties based on the eavesdropper Eve's location and the length of the quantized bits. This ensures that the legitimate communication parties, Alice and Bob, can correctly recover the confidential message while Eve cannot. As a result, we reinforce the VBF-PST's resilience to attacks from nearby sources.
- We conduct a quantitative analysis and extensive computer simulations for the proposed VBF-PST scheme from three aspects: randomness, consistency, and security. The results demonstrate that compared to existing schemes, VBF-PST significantly improves the key generation NIST test pass rate to nearly 99% and increases the bit replacement rate (BRR) to 45%, thus enhancing the security of the confidentiality communication scheme more effectively.

The rest of the paper is organized as follows: Section II introduces the related work. Section III provides a detailed description of the system model. Section IV discusses differential attacks against UK-PSK. Section V discusses the proposed scheme and its specific working details. Section VI analyzed the performance of the VBF-PST scheme. In Section VII, simulation results are presented from three perspectives: consistency, randomness, and security. Finally, Section VIII concludes the paper.

Notation and Outline: Unless otherwise specified, we use the following notations throughout the manuscript: upper (lower) bold-face letters denote matrices (column vectors). Numeral subscripts of matrices and vectors, if needed, represent their sizes. Also, matrix superscripts $(\cdot)^{\mathrm{H}}$, $(\cdot)^{\mathrm{T}}$, and $(\cdot)^*$ denote their conjugate transpose, transpose, and conjugate, respectively. We use $\{\cdot\}$ to denote the set of events in probability theory, $||\cdot||_1$ to represent the L_1-norm, and $\lceil\cdot\rceil$ to indicate the ceiling function, which rounds up to the nearest integer.

II. Related Works

To the best of our knowledge, no existing solutions have simultaneously addressed both the issue of redundant overhead in the key generation process and the security challenges in quasi-static or low-speed environments. To tackle both problems, we consider using a VBF. Therefore, we focus on studying OTP secure transmission schemes with unidentical keys, key generation schemes for quasi-static or low-speed environments, and the application of BFs.

OTP with unidentical keys. There have been a number of research attempts on OTP with unidentical keys [4]–[6]. In 2019, [7] proposed the UK-PST method that outperforms traditional OTP secure transmission schemes. This method cascades key generation with OTP secure communication, generating unidentical keys without conducting information reconciliation and privacy amplification. UK-PST benefits from its simple structure and can effectively minimizes information leakage resulting from data transmission. As a result, it is well-suited for situations that demand low data rates and high levels of security.

Key generation in static environments. Most PKG methods require dynamic environments or channels to achieve

high key rates [12]–[16]. To enable key generation in static environments, [17] proposed a low-complexity protocol in 2018 for rapidly generating keys in such settings. They utilized independently generated random bits by the legitimate parties, combined with fading parameters, as a common source of randomness. In 2020, [18] proposed a low-complexity method called induced randomness. In this method, both communicating parties independently generate local randomness and combine it with the uniqueness of wireless channel coefficients to achieve high-rate key generation. Nevertheless, these techniques unavoidably need an information reconciliation process during key generation, and its practical usage may be compromised by additional transmissions and information leakage.

Application of BFs. The BF, introduced by Bloom [8], is a space-efficient probabilistic data structure used for membership testing. BFs can query whether an element belongs to a set or not [19], [20]. For applications in networks, BF have been used in data mining and analysis [21]–[23], cloud computing [24], [25], and machine learning [26]–[29]. From database perspective, they have also been used in key-value stores [30]–[32] and privacy-preserving record linkage [33]. In addition, BF encoding is also widely used as an effective masking technique for privacy-preserving matching functions [9], [34].

III. System Overview

A. Channel Model

This paper investigates the secure transmission of partially mismatched key generation and OTP joint design in quasi-static or low-rate environments. In this scenario, a legitimate user Bob, intends to send confidential information to Alice without being intercepted by a third party. Transmission errors between Alice and Bob are corrected using **forward error correction (FEC)**[3]. FEC is a technique used to improve the reliability of data transmission by sending additional redundant data. This redundant data, commonly referred to as ECC, contains information about the original data, enabling the receiver to reconstruct the correct data even if errors occur during transmission. Bob first encodes the confidential information through FEC and then encrypts it using OTP before transmitting it to Alice. Perfect secrecy is achieved by OTP through XOR operations, where Bob encrypts the information bit by bit with a key. Alice decrypts the ciphertext using her key and then employs a decoder for error detection and correction to recover the confidential information. In this process, Alice and Bob utilize a pair of unidentical keys for secure transmission.

B. Attack Model

Concerning the attacker Eve (details are provided in Section IV), this paper focuses on passive attacks. For the differential attack, We assume that both parties in communication are trusted users, and that there exists only one passive eavesdropper Eve, equipped with a single antenna, located at a distance of at least half a wavelength from the legitimate users. Consequently, the wireless channel between Eve and the legitimate users is independent. Additionally, this paper considers scenarios where, in unattended areas, many IoT devices or wireless sensors are deployed to select and transmit information. Since no one is monitoring these devices, Eve can easily position herself near Alice or Bob and perform a proximity attack [35]. Therefore, the proximity attacks is a big security hazard in secret key generation. For proximity attacks, we assume that the passive eavesdropper Eve is located near one of the legitimate users, such as Bob. She can easily obtain a channel state information similar to Bob's and attempt to compromise the key shared between the legitimate users. For both of the aforementioned attacks, Eve cannot obtain the legitimate users' measurement values through observation, however, she can eavesdrop on the information transmitted over the public channel and is aware of all the protocols being used.

C. Problem Statement and System Model

In quasi-static or low-rate environments, the use of PKG faces challenges such as uneven distribution of 0 and 1 bits in the quantized bit sequence, as well as high repetition of sequences between adjacent frames. These issues pose security risks to the system. Additionally, the information reconciliation process requires communication over public channels, which not only increases system overhead but also introduces the risk of information leakage.

To reduce the number of interactions during the key generation process, we opt for generating unidentical keys for OTP encryption. In the proposed scheme, the key generation process does not include the information reconciliation and privacy amplification steps. Additionally, to enhance security in quasi-static or low-rate environments, we apply a VBF to the quantized bit sequence, which improves the randomness of the bit sequence.

We consider generating unidentical keys for both communication parties using PKG methods derived from wireless channels. The specific process is shown in Fig. 1. The process of PKG typically involves four steps: channel probing, quantization, information reconciliation and privacy amplification.

In the channel probing stage, we consider a narrowband communication system operating in time division duplexing (TDD) mode, where the legitimate communication users, Alice and Bob, are each equipped with a single antenna operating at the same frequency. Our study focuses on quasi-static or low-speed environments, where the wireless channel between Alice and Bob is modeled as a narrowband multipath fading channel. Therefore, the uplink and downlink channels between Alice and Bob are reciprocal. In PKG, channel probing is used to acquire reciprocal channel state information, such as received signal strength (RSS) and channel state information (CSI). Alice and Bob independently collect channel measurements $\mathbf{r}_{a,b}$ and $\mathbf{r}_{b,a}$, respectively. After completing N_T measurements within time T, Bob estimates the channel amplitude as $\mathbf{r}_{a,b} = [r_{a,b}(1), r_{a,b}(2), \ldots, r_{a,b}(N_T)]^T$,

[3]The core of FEC technology lies in the encoding process, where ECCs are added before transmitting the data, and the decoding process, where these codes are used upon data reception to detect and correct errors.

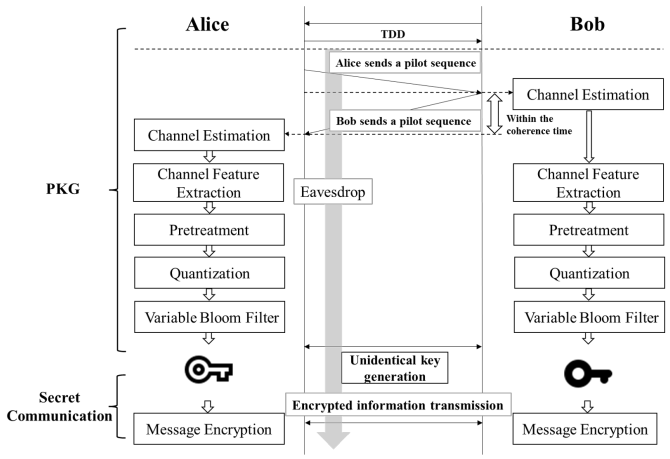


Fig. 1. The basic process of secure communication based on unidentical keys.

while Alice obtains the channel amplitude as $\mathbf{r}_{b,a} = [r_{b,a}(1), r_{b,a}(2), \ldots, r_{b,a}(N_T)]^T$.

Subsequently, both communication parties independently quantize the channel amplitudes into 01-bit sequences. For instance, using a two-bit quantizer, the quantization can be expressed as follows:

$$q\big(r(t_k)\big) = \begin{cases} 11, F^{-1}(0.75) < r(t_k) \leq F^{-1}(1); \\ 10, F^{-1}(0.5) < r(t_k) \leq F^{-1}(0.75); \\ 01, F^{-1}(0.25) < r(t_k) \leq F^{-1}(0.5); \\ 00, F^{-1}(0) \leq r(t_k) \leq F^{-1}(0.25), \end{cases} \tag{1}$$

where $F(x) = \Pr(X < x)$ is the cumulative distribution function (CDF), $\Pr(\cdot)$ is a probability function, and $F^{-1}(\cdot)$ is the inverse function of $F(x)$.

In this paper, we denote $\mathbf{q}_{a,b}(T) = [q_{a,b}(1), q_{a,b}(2), ..., q_{a,b}(L_q)]^T$ as the quantization result for Bob, where the length of $\mathbf{q}_{a,b}(T)$ is $L_{\mathbf{q}} = N_T Q_{Level}$. Here, Q_{Level} represents the number of quantization levels. Similarly, the quantization result for Alice is denoted as $\mathbf{q}_{a,b}(T)$, and for Eve as $\mathbf{q}_{a,e}(T)$. In this paper, the Bit Disagreement Rate (BDR) is defined as the number of different bits between Alice and Bob divided by the total number of bits of the raw key, which is calculated using the raw bits after quantization. Let ϵ_q be the BDR between $\mathbf{q}_{b,a}(T)$ and $\mathbf{q}_{a,b}(T)$, and ϵ_e be the BDR between $\mathbf{q}_{b,e}(T)$ and $\mathbf{q}_{e,b}(T)$, which are defined as:

$$\epsilon_q = \frac{\| q_{a,b} - q_{b,a} \|_1}{L_q}, \tag{2}$$

$$\epsilon_e = \frac{\| q_{e,b} - q_{b,e} \|_1}{L_q}. \tag{3}$$

Due to the presence of the bit disagreement rate ϵ_q between the quantized bit sequences, we employ FEC to correct errors in transmitted information. It treats the bit inconsistencies between the quantized bit sequences and the actual transmitted channel errors as the errors generated by an equivalent channel during transmission. Assuming the actual transmission channel error rate is ϵ_0, then the error rate ϵ_{eq} of the equivalent channel [36] is given by:

$$\epsilon_{eq} = \epsilon_0 + \epsilon_q - 2\epsilon_0\epsilon_q. \tag{4}$$

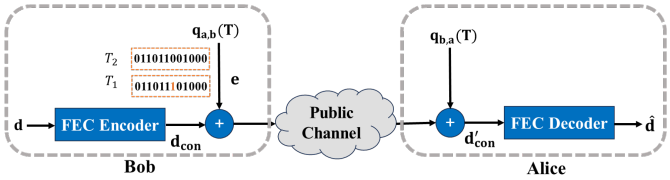


Fig. 2. An illustration of the flowchart of the UK-PST scheme.

IV. Attack Against UK-PSK

A. The Protocol Description of UK-PSK

Fig. 2 illustrates secure transmission using a pair of unidentical keys. We assume that $\mathbf{q}_{b,a}(T)$ and $\mathbf{q}_{a,b}(T)$ represent the quantized results of the channel probing values between Alice and Bob. T_1 and T_2 denote adjacent time intervals, with $\mathbf{q}_{b,a}(T_1)$ and $\mathbf{q}_{a,b}(T_2)$ representing the quantized results for Bob's adjacent frames, where the lengths of the quantized bits in the two adjacent frames are the same. After quantization, the legitimate parties obtain unidentical bit sequences, which are then used to encrypt and decrypt the confidential message $\mathbf{d}$. The UK-PST scheme treats the XOR encryption and decryption module along with the physical channel as an equivalent cascade channel, with tiny differences between keys regarded as part of the transmission error, and thus can be corrected by the off-the shelf FEC with a stronger correction capability [7].

The UK-PST scheme avoids the information reconciliation step, offering a simple structure for secure OTP transmission. This reduces system communication overhead and minimizes the risk of information leakage. However, in quasi-static or low-speed environments, where the user's Doppler frequency is low, it thus requires a long time to establish long enough keys. Additionally, there always exists inevitable and unknown temporal correlation between adjacent channels samples, resulting in a large proportion of repeated bit segments in the quantized bit sequence. The refresh rate and randomness of adjacent frame quantized bits are low, leading to greater security risks. Consequently, in slow-varying environments, this scheme may be vulnerable to differential cryptanalysis attacks during multiple ciphertext transmissions.

B. Proposed Differential Attack

In a static environment, the attacker Eve can infer the correlation between $\mathbf{q}_{a,b}(T_1)$ and $\mathbf{q}_{a,b}(T_2)$ by calculating the correlation of the quantized bits $\mathbf{q}_{a,e}(T_1)$ and $\mathbf{q}_{a,e}(T_2)$ obtained at adjacent times T_1 and T_2. By observing the ciphertext of secret communication in adjacent instances, a differential attack can be executed to obtain confidential message and the key. In this work, $\aleph_{\mathbf{q}}(T_1, T_2)$ denotes the Bit Replacement Rate (BRR) between $\mathbf{q}_{a,b}(T_1)$ and $\mathbf{q}_{a,b}(T_2)$, defined as:

$$\aleph_{\mathbf{q}}(T_1, T_2) = \frac{\| \mathbf{q}_{a,b}(T_1) - \mathbf{q}_{a,b}(T_2) \|_1}{L_{\mathbf{q}}}, \tag{5}$$

where, $0 \leq \aleph_{\mathbf{q}}(T_1, T_2) \leq 1$.

In a static environment, where the channel changes slowly, there unavoidably exists correlation in the key generation based on wireless channel characteristics. Between adjacent time instances T_1 and T_2, there is a strong correlation between the estimated channel amplitudes obtained by the legitimate communication parties. Consequently, multiple bits in the quantized key $\mathbf{q}_{a,b}(T_1)$ and $\mathbf{q}_{a,b}(T_2)$ remain unchanged. Exploiting this characteristic, Eve can execute a differential attack.

For the convenience of describing and analyzing this attack, let us assume two instances of secure communication denoted by $\mathbf{e}(T_1)$ and $\mathbf{e}(T_2)$, corresponding to adjacent times T_1 and T_2, respectively. The ciphertexts adhere to the following equation:

$$\mathbf{e}(T_1) = \mathbf{d}_{\text{con}}(T_1) \oplus \mathbf{q}_{a,b}(T_1), \tag{6}$$

$$\mathbf{e}(T_2) = \mathbf{d}_{\text{con}}(T_2) \oplus \mathbf{q}_{a,b}(T_2). \tag{7}$$

By observing $\mathbf{e}(T_1)$ and $\mathbf{e}(T_2)$, the eavesdropper Eve can obtain the following equation:

$$\begin{aligned} \mathbf{e}_{attack} &= \mathbf{e}(T_1) \oplus \mathbf{e}(T_2) \\ &= (\mathbf{d}_{\text{con}}(T_1) \oplus \mathbf{d}_{\text{con}}(T_1)) \oplus (\mathbf{d}_{\text{con}}(T_2) \oplus \mathbf{q}_{a,b}(T_2)) \\ &= (\mathbf{d}_{\text{con}}(T_1) \oplus \mathbf{d}_{\text{con}}(T_2)) \oplus (\mathbf{q}_{a,b}(T_1) \oplus \mathbf{q}_{a,b}(T_2)) \\ &= (\mathbf{d}_{\text{con}}(T_1) \oplus \mathbf{d}_{\text{con}}(T_2)) \oplus \triangle\mathbf{q}_{ab}(T_1, T_2), \end{aligned} \tag{8}$$

where $\triangle\mathbf{q}_{ab}(T_1, T_2)$ represents the difference between $\mathbf{q}_{a,b}(T_1)$ and $\mathbf{q}_{a,b}(T_2)$ in the binary bit sequence.

In the worst-case scenario, the key strings $\mathbf{q}_{a,b}(T_1)$ and $\mathbf{q}_{a,b}(T_2)$ generated at two adjacent times T_1 and T_2 are completely identical, i.e., $\aleph_{\mathbf{q}}(T_1, T_2) = 0$. In this case, $e_{attack} = \mathbf{d}_{\text{con}}(T_1) \oplus \mathbf{d}_{\text{con}}(T_2)$. Suppose Eve has already obtained $\mathbf{d}_{\text{con}}(T_1)$ through ciphertext analysis at time T_1. Then, the transmission of $\mathbf{d}_{\text{con}}(T_2)$ at time T_2 will also be fully exposed to Eve, leaving the data vulnerable. Therefore, in UK-PST, where channel randomness is insufficient and the length of channel keys is limited, multiple secure transmission processes may be vulnerable to differential attacks, leading to confidential message leakage and posing significant security risks.

C. Proposed Proximity Attack

Under the proximity attack, Eve is located near to either two legitimate users, e.g., Bob, without loss of generality, as shown in Fig. 3. Thus, **the channel of Eve is correlated to that of Bob.** When Alice and Bob transmit probing signals to estimate channel state information, Eve can also obtain a similar channel observation and generate secret keys similar to that between Alice and Bob [37]. Eve has the capability to eavesdrop on all information transmitted over the public channel and is aware of all protocols involved in the bit sequence generation process during channel probing and quantization.

In this paper, let h_{ab} represent the channel coefficient between Alice and Bob, and h_{ae} represent the channel coefficient between Alice and Eve. The parameter ρ_{ae} denotes the correlation coefficient between h_{ab} and h_{ae}, while ρ_{ab} is

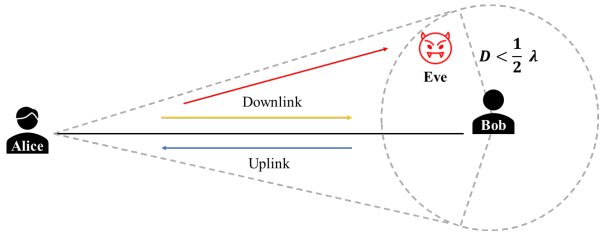


Fig. 3. Proximity Attack Scenario.

the autocorrelation coefficient of h_{ab}. Since the distance D between the eavesdropper Eve and Bob is within half a wavelength, and the correlation coefficient of the eavesdropping channel should be less than the autocorrelation coefficient of the legitimate channel, we have

$$0 < \rho_{ae} = \mathcal{J}_0(2\pi \triangle D) < \rho_{ab} = \mathcal{J}_0(2\pi f_d \tau), \tag{9}$$

where $J_0(\cdot)$ is the zero-order Bessel function of the first kind, λ is half wavelength, $\Delta D = D/\lambda$ is the normalized distance, f_d is the maximum Doppler shift, and τ is the sampling delay. From (9), the range of D can be determined as $(f_d\tau\lambda, \lambda/2)$.

When Eve approaches Bob, the channel coefficients between Alice and Eve are highly correlated with those between Alice and Bob. As a result, the quantized values $\mathbf{q}_{a,e}(T)$ obtained by Eve through channel probing and quantization may exhibit a certain degree of correlation with $\mathbf{q}_{a,b}(T)$. Thus, Eve may have similar channel observation and crack the secret keys generated from the channel. Therefore, the UK-PST secure transmission process may be vulnerable to proximity attack, leading to confidential message leaks and posing significant security risks.

V. OTP Secure Transmission Scheme

In this section, we introduce the VBF-PST scheme and the operational workflow of the VBF, and discussed how the VBF-PST scheme addresses the additional communication overhead caused by information reconciliation and the security threat of information leakage in slowly changing channels. We then provide a theoretical derivation to explain how to select parameters for the VBF to achieve output bit sequences with higher randomness, which can not only resist attacks but also reduce system overhead.

A. VBF-PST Protocol

The key generation process of the VBF-PST scheme involves four steps: channel probing, quantization, Bloom filter mapping, and data transmission. After quantization by both communication parties, the quantized bits are processed through a VBF to obtain partially inconsistent bit sequences, which are then used to encrypt and decrypt the confidential message d. We define $\mathbf{q}_{b,a}(T)$ and $\mathbf{q}_{a,b}(T)$ as the quantized results of the channel probing values of Alice and Bob, with a length of $L_{\mathbf{q}}$. Fig. 4 illustrates the entire process of the VBF-PST scheme.

Step 1: Alice and Bob initialize the same hash table $\mathcal{H}$, with K hash functions. The array of the VBF, with length m,

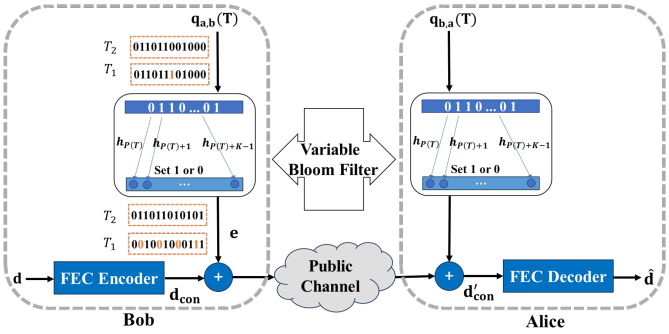


Fig. 4. An illustration of the flowchart of the VBF-PST scheme.

is initialized to all zeros using the $\mathcal{SET}$ method (details are provided in Subsection V-B).

Step 2: Alice and Bob each utilizes at least the first $\lceil \log_2(l_{\mathcal{H}}) \rceil$ odd bits of the quantized values $\mathbf{q}_{b,a}(T)$ and $\mathbf{q}_{a,b}(T)$, along with the current round number C of the key generation process, to compute the hash pointers $P^a(T)$ and $P^b(T)$ for this round, we have

$$P^a(T) = \left(\sum_{i=0}^{n-1} 2^i \mathbf{q}_{b,a}(T)_{2i+1} + C \right) \bmod l_{\mathcal{H}}, n \geq \lceil \log_2(l_{\mathcal{H}}) \rceil, \tag{10}$$

$$P^b(T) = \left(\sum_{i=0}^{n-1} 2^i \mathbf{q}_{a,b}(T)_{2i+1} + C \right) \bmod l_{\mathcal{H}}, n \geq \lceil \log_2(l_{\mathcal{H}}) \rceil. \tag{11}$$

Step 3: Alice and Bob use the quantized values $\mathbf{q}_{b,a}(T)$ and $\mathbf{q}_{a,b}(T)$, which are of length $L_{\mathbf{q}}$, as inputs to the VBF. After the mapping process is finished, they obtain binary quantized bit sequences, represented as $\mathbf{q}_{b,a}^{BF}(T)$ and $\mathbf{q}_{a,b}^{BF}(T)$, both having a length of $L_{\mathbf{q}}^{BF}$. These sequences are not exactly the same, but they are quite similar.

Step 4: Bob enters the secret information $\mathbf{d}$ with a length of $L_{\mathbf{d}}$ into the channel encoder, which produces a parity check sequence with a length of $L_{\mathbf{s}}$, and we can get

$$\mathbf{s} = \mathcal{E}(\mathbf{d}), \tag{12}$$

where $\mathcal{E}(\cdot)$ represents the generation function of $(\mathcal{C}, \mathrm{n}, \mathrm{k}, \mathrm{t})$ ECC, e.g., BCH.

Step 5: Bob concatenates the secret information $\mathbf{q}$ to be transmitted with the output syndrome $\mathbf{s}$ to obtain

$$\mathbf{d}_{\mathrm{con}} = [\mathbf{d}, \mathbf{s}]. \tag{13}$$

Bob encrypts $\mathbf{d}_{\mathrm{con}}$ using the bit sequence $\mathbf{q}_{a,b}^{BF}(T)$ to obtain the ciphertext $\mathbf{e}$. The total length of the ciphertext is $L_{\mathbf{q}}^{BF} = L_{\mathbf{d}} + L_{\mathbf{s}}$. The bits-stream encryption is realized by

$$\mathbf{e} = \mathbf{d}_{\mathrm{con}} \oplus \mathbf{q}_{a,b}^{BF}(T). \tag{14}$$

Step 6: Bob directly transmits the ciphertext $\mathbf{e}$ to Alice over the public transmission channel.

Step 7: Alice decrypts $\mathbf{e}$ with $\mathbf{q}_{b,a}^{BF}(T)$ by

$$\mathbf{d}'_{\mathrm{con}} = \mathbf{e} \oplus \mathbf{q}_{b,a}^{BF}(T) = \mathbf{d}_{\mathrm{con}} \oplus \triangle \mathbf{q}^{BF}, \tag{15}$$

where

$$\triangle \mathbf{q}^{BF} = \mathbf{q}_{b,a}^{BF}(T) \oplus \mathbf{q}_{a,b}^{BF}(T). \tag{16}$$

Step 8: Alice recovers the confidential message $\hat{\mathbf{d}}$ by

$$\hat{\mathbf{d}} = \mathcal{D}(\mathbf{d}'_{con}) = \mathbf{d}, \tag{17}$$

where, $\mathcal{D}(\cdot)$ represents the decoding function of the ECC$(\mathcal{C}, \mathrm{n}, \mathrm{k}, \mathrm{t})$. This paper assumes $\epsilon_{\mathbf{q}}^{BF}$ as the bit disagreement rate between $\mathbf{q}_{b,a}^{BF}(T)$ and $\mathbf{q}_{a,b}^{BF}(T)$, where the number of '1' in $\triangle \mathbf{q}^{BF}$ is $\epsilon_{\mathbf{q}}^{BF} L_{\mathbf{q}}^{BF}$.

This scheme eliminates the information reconciliation step, thus mitigating the additional overhead and risk of information leakage associated with that process. Furthermore, it employs a variable Bloom filter to reorder the bit sequences, effectively resolving the issue of low randomness in the bit sequences in static or quasi-static environments.

In a quasi-static or low-rate environment, to ensure obtaining a bit sequence with better randomness and higher BRR, we process the quantized bit sequences through a VBF. To avoid affecting subsequent data transmission, the VBF must not amplify the original quantization result's discrepancy $\triangle \mathbf{q}$. This principle is essential to further enhance the randomness and BRR of the bit sequences, which will be used as keys in the subsequent secure transmission process.

The VBF can maintain the Jaccard distance between the original input sequence and the output sequence [9]. Assuming no hash collisions, if there are $\epsilon_{\mathbf{q}} L_{\mathbf{q}}$ '1' in the discrepancy $\triangle \mathbf{q}$ between $\mathbf{q}_{a,b}(T)$ and $\mathbf{q}_{b,a}(T)$, then the discrepancy $\triangle \mathbf{q}^{BF}$ between $\mathbf{q}_{a,b}^{BF}(T)$ and $\mathbf{q}_{b,a}^{BF}(T)$ will also have $\epsilon_q L_q$ '1'. Assuming hash collisions exist, then the number of 1-bits in $\triangle \mathbf{q}^{BF}$ between $\mathbf{q}_{a,b}^{BF}(T)$ and $q_{b,a}^{BF}(T)$, denoted as $\epsilon_{\mathbf{q}}^{BF} L_{\mathbf{q}}^{BF}$, will be less than $\epsilon_{\mathbf{q}} L_{\mathbf{q}}$ '1', we have

$$\epsilon_q^{BF} \leq \epsilon_q \frac{L_{\mathbf{q}}}{L_{\mathbf{q}}^{BF}}. \tag{18}$$

In UK-PST, for the ECC$(\mathcal{C}', \mathrm{n}', \mathrm{k}', \mathrm{t}')$, where $\mathrm{k}' = L_{\mathbf{d}}$ and $\mathrm{n}' = L_{\mathbf{q}}$, assuming $\mathcal{C}'$ is the maximum error detection capability of the current ECC, we have

$$\epsilon_{eq} = \frac{\mathrm{t}'}{\mathrm{n}'} = \frac{\mathrm{t}'}{L_{\mathbf{q}}}. \tag{19}$$

In VBF-PST, let us assume $L_{\mathbf{q}} < L_{\mathbf{q}}^{BF}$. Therefore, we have

$$\begin{aligned} \epsilon_{eq}^{BF} &= \epsilon_0 + (1 - 2\epsilon_0)\epsilon_q^{BF} \\ &\leq \epsilon_0 + (1 - 2\epsilon_0)\epsilon_q \frac{L_{\mathbf{q}}}{L_{\mathbf{q}}^{BF}} \\ &\leq \epsilon_0 + (1 - 2\epsilon_0)\epsilon_q = \frac{\mathrm{t}'}{\mathrm{n}'}, \end{aligned} \tag{20}$$

where ϵ_0 represents the probability of errors in the physical channel, ϵ_{eq} represents the error ratio of the equivalent channel when only ECC is used in a cascaded channel, and ϵ_{eq}^{BF} represents the error ratio of the equivalent channel when both ECC and the VBF are used in the channel.

Thus, when $L_{\mathbf{q}} < L_{\mathbf{q}}^{BF}$ under the same transmission conditions, the VBF-PST will not surpass the error detection capacity limit of the ECC. This ensures that the secret information may be successfully recovered without any issues.

Therefore, we have successfully shown the fundamental principles outlined in this section.

In this scheme, the role of the VBF is to process the suboptimal bit sequences obtained from quantization in quasi-static or low-speed environments. It essentially functions as the privacy amplification step in PKG. Since the parameters of the VBF are adjustable, it not only enhances the randomness of the bit sequences but also improves the BRR of successive bit sequences, effectively addressing the security vulnerabilities of the UK-PST scheme. Compared to the IK-PST scheme [7], which is based on identical keys for physical-layer secure transmission, the VBF-PST scheme eliminates the information reconciliation step and does not require information exchange over a public channel. This reduces communication overhead and lowers the risk of information leakage. Sections V-B and V-C provide detailed explanations of the operational workflow and parameter design of the VBF.

B. Variable Bloom Filter

A VBF consists of K hash functions and a bitmap with a size of M. The VBF, denoted as $BF\{\mathcal{H}, K, P(T), M, \mathcal{SET}\}$. In the VBF, $\mathcal{H} = \{h_1, h_2, ..., h_{l_\mathcal{H}}\}$ represents the sequence of $l_\mathcal{H}$ hash functions stored, K indicates the number of hash functions used for bit mapping in the VBF, with $K \in [1, l_\mathcal{H}]$. $P(T)$ represents the position of the first hash function used in this round of key generation in the hash book $\mathcal{H}$, with $P(T) \in [1, l_\mathcal{H}]$. When the filter performs bit mapping, it does not use all the hash functions in the hash library but selects K consecutive hash functions starting from the index $P(T)$. Define J as the set of bit indices in the VBF, denoted by $J = \{1, 2, \ldots, M\}$, where M represents the size of the bitmap included within the VBF.

The VBF can be considered as a universal set U containing all the elements of a set $A = \{x_1, x_2, \ldots, x_N\}$, where N represents the size of set A. For an element x_i in A, $h_l(x_i)$ denotes the index value obtained by hashing the element x_i using hash function $h_l(\cdot)$, and $h_l(x_i) \in J$, where $l = 1, ..., l_\mathcal{H}$ and $i = 1, ..., N$. In the VBF, $\mathcal{SET}$ represents the operation of setting bits after bit mapping. Let $v(h_l(x_i))$ denote the value of the bit array at index $h_l(x_i)$, where $\mathcal{SET} = 1$ indicates that the VBF sets the value of bit array $v(h_l(x_i))$ to 1, and $\mathcal{SET} = 2$ indicates that the VBF sets the value of bit array $v(h_l(x_i))$ to either 1 or 0.

$$\mathcal{SET} = 1 \rightarrow v\big(h_l(x_i)\big) = 1, \tag{21}$$

$$\mathcal{SET} = 2 \rightarrow v\big(h_l(x_i)\big) = \begin{cases} 1, & l \bmod 2 = 1; \\ 0, & l \bmod 2 = 0. \end{cases} \tag{22}$$

Unlike traditional BFs, the variability in our approach allows users to select the number of hash functions K and the bit-setting method $\mathcal{SET}$ based on the input and output sequence lengths prior to key generation. This flexibility leads to a higher degree of randomness in the bit sequences, which can then be used for secure OTP transmission.

In this scheme, we define the quantized bit sequence as $\mathbf{X}$ with a length of N, serving as the input to the filter. We define the output quantized bit sequence of the filter as $\mathbf{Y}$ with a length of M. The core function of the VBF is to enhance the randomness and consistency of the sequence by dynamically adjusting its parameters. The main processing steps include initialization, appending bits, obtaining the mapped index, and setting the bits. The dynamic process of the VBF is as follows: During initialization, the starting index $P(T)$ of the hash functions used in this round of key generation is determined based on the first three bits of each quantized bit sequence. Subsequently, each bit is processed through the hash function indexed by $P(T)$. After processing each bit, the starting position $P(T)$ is updated to introduce different combinations of hash functions until the entire sequence is processed. Due to the dynamic nature of $P(T)$, even if the input sequences are correlated within adjacent time intervals, the randomness of the output sequence is significantly enhanced. A VBF is shown in Fig. 5 with several elements inserted on the filter.

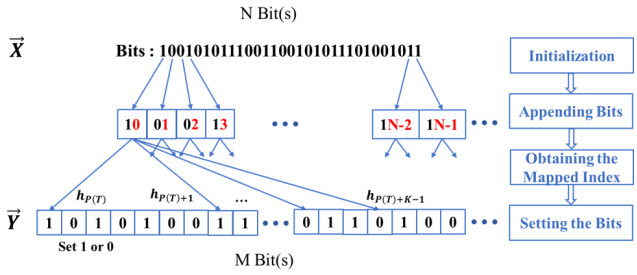


Fig. 5. Flowchart of Variable Bloom Filter.

The VBF first initializes a vector $\mathbf{Y}$ of length M filled with zeros. Then, it calculates the length N of the input sequence $\mathbf{X}$. For each bit X_i in the input sequence $\mathbf{X} = [X_1, X_2, ..., X_N]$, it appends a position bit, resulting in the vector $\mathbf{X}' = [X_1 0, X_2 1, ..., X_N(N-1)]$. Next, it calculates the index for each bit X_i in $\mathbf{X}'$ using K hash functions: $index = h_l(X_i(i-1)) \bmod M+1$, where $l \in [P(T), (P(T)+K-1) \bmod l_\mathcal{H}+1]$. Finally, it sets the bits of $\mathbf{Y}$ according to the setting method $\mathcal{SET}$ and the index of the bit array. This process continues until all bits in $\mathbf{X}$ have been processed. The specific algorithm is outlined in **Algorithm 1**.

Since $P(T)$ changes with each round of bit sequence processing, the VBF applies different hash functions to each bit. As a result, even if there are issues such as an uneven distribution of 0's and 1's in the bit sequence or low BRR between adjacent frames, the filter can still produce a highly random bit sequence. This ensures that potential security threats due to information leakage in slow-varying channels are effectively addressed.

C. Scheme Parameter Design

Different from traditional BFs, the VBF $BF\{\mathcal{H}, K, P(T), M, \mathcal{SET}\}$ automatically changes $P(T)$ during the processing of each round of quantized bit sequences, while the hash functions set $\mathcal{H}$, the number of hash functions used K, and the setting method $\mathcal{SET}$ are shared by both parties. The hash library $\mathcal{H}$ and the output bit sequence length M are common and known to both parties. There are only three parameters that users can choose: the number of hash functions K, the initial value of $P(T)$,

Algorithm 1: VBF Workflow Algorithm

Input: Quantized bit sequence $\mathbf{X} = [X_1, X_2, ..., X_N], X_i \in \{0, 1\}, 1 \leq i \leq N$, $K = k_{\text{con}}, P(T) = p_{\text{con}}, \mathcal{SET} = set_{\text{con}}$

Output: Output bit sequence $\mathbf{Y} = [Y_1, Y_2, ..., Y_M]$

initialization: Filtered Output Key $\mathbf{Y} = [Y_1, Y_2, ..., Y_M], Y_j = 0, 1 \leq j \leq M$

for $i \leftarrow 1, N$ **do**
  for $z \leftarrow P(T), P(T) + K - 1$ **do**
    $l = z \bmod l_{\mathcal{H}} + 1$;
    $\text{index} = h_l(X_i(i-1)) \bmod M + 1$;
    if $\mathcal{SET} == 2$ **then**
      if $l \bmod 2 == 1$ **then**
        $Y_{index} = 1$;
      else
        $Y_{index} = 0$;
      end
    end
    else if $\mathcal{SET} == 1$ **then**
      $Y_{index} = 1$
    end
  end
end

and the setting method $\mathcal{SET}$. Subsection V-B provides a detailed explanation of the selectable ranges for these three parameters. The choice of parameters greatly affects the randomness and BRR of the resulting bit sequences. In this section, we calculate the entropy of the bit sequences to determine the parameters that theoretically yield the highest bit sequence entropy.

We assume that the indexes generated by each hash function in the hash set are uniformly distributed relative to the bitmap inside the filter (i.e., the probability density of hash indexes is uniform), given by $\Pr(\{index = j\}) = 1/M$, where Pr represents probability and $j = 1, ..., M$. We define $\Pr_j(0)$ as the probability that the bit at index j in the bitmap is eventually set to '0', and $\Pr_j(1)$ as the probability that the bit at index j in the bitmap is eventually set to '1'. According to the information entropy formula, we can obtain the information entropy H of the output bit sequence $\vec{Y}$ as follows:

$$H = -\Pr_j(1)\log_2\left(\Pr_j(1)\right) - \Pr_j(0)\log_2\left(\Pr_j(0)\right). \quad (23)$$

We conduct theoretical analysis on the VBF. For each hash function, the probability of setting any position to 1 in $\mathbf{Y}$ is $p_1 = \frac{1}{M}$, and the probability of not setting it to '1' (keeping it as '0') is $p_0 = 1 - \frac{1}{M}$, where M represents the output length of the BF.

When the VBF is set to $\mathcal{SET} = 1$, the probability of $Y_j = 0$ is $\Pr_j(0) = (p_0)^{KN} = \left(1 - \frac{1}{M}\right)^{KN}$. Thus, the probability of $Y_j = 1$ is $\Pr_j(1) = 1 - \Pr_j(0) = 1 - \left(1 - \frac{1}{M}\right)^{KN}$. The partial derivative of the entropy H with respect to K is calculated as follows:

$$\frac{\partial H}{\partial K} = N\Pr_j(0)\ln p_0 \log_2(\Pr_j(0)^{-1} - 1). \quad (24)$$

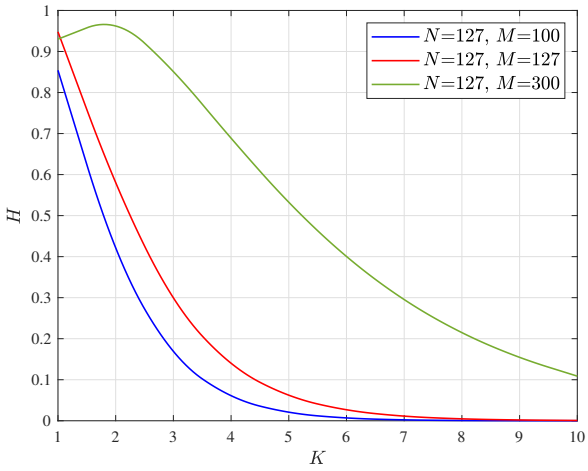


Fig. 6. Entropy H when $\mathcal{SET} = 1$.

The maximum entropy value achieved by different values of K varies with different values of N and M. The distribution of entropy H with respect to K is illustrated in Fig. 6.

As observed from Fig. 6, when $K < -\frac{1}{N\log_2\left(1-\frac{1}{M}\right)}$, H increases with the increase of K; when $K > -\frac{1}{N\log_2\left(1-\frac{1}{M}\right)}$, H decreases with the increase of K. Therefore, selecting the nearest positive integer to $\frac{1}{N\log_2\left(1-\frac{1}{M}\right)}$ as the value of K in $\mathcal{SET} = 1$ results in the highest entropy.

When the VBF is set to $\mathcal{SET} = 2$, the probability of $Y_j = 1$ and $Y_j = 0$ will vary depending on the parity of K and $P(T)$. Therefore, the discussion on the entropy variation needs to be divided into four cases. a is represented as $K = 2a + 1$ or $K = 2a$ based on the parity of K

Proposition 1: When $\mathcal{SET} = 2$, the optimal value of a is given by

$$a \in \left(0, -\frac{\ln(2M)}{2N\ln(1-\frac{1}{M})}\right). \quad (25)$$

Proof: See Appendix.

Fig. 7 provides a visual comparison of the specific values of entropy H under different parameters. For the same value of a, the entropy H is higher when K is even compared to when K is odd. Therefore, when the setting mode $\mathcal{SET} = 2$, the entropy H is higher within the range $a \in (0, -\frac{\ln(2M)}{2N\ln(1-\frac{1}{M})})$ for even values of K. Additionally, when K is even, the entropy H is slightly higher when the hash pointer $P(T)$ is even compared to when it is odd, though the difference is not significant. Conversely, when K is odd, the entropy H is slightly higher when the hash pointer $P(T)$ is odd, and the difference gradually decreases as K increases, eventually approaching zero.

Overall, before using the variable filter, we need to first determine the value of $\mathcal{SET}$. Based on the aforementioned theoretical derivations and simulation results, the entropy of sequences with $\mathcal{SET} = 2$ is always higher than that with $\mathcal{SET} = 1$, so we prioritize choosing $\mathcal{SET} = 2$. Next, we calculate a based on the input length N and output length

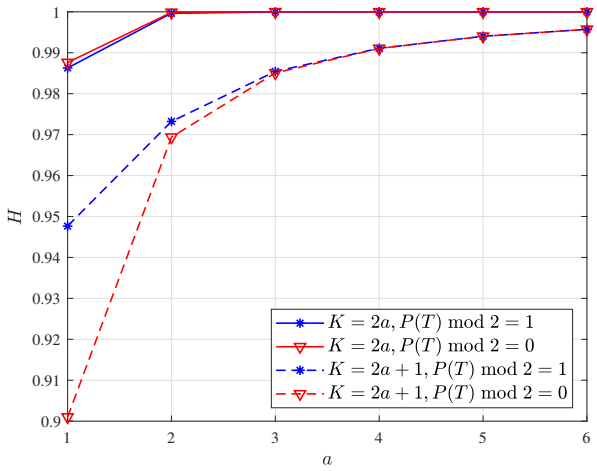


Fig. 7. Entropy H when $\mathcal{SET} = 2$.

M of the generated key, resulting in $a \in (0, -\frac{\ln(2M)}{2N\ln(1-\frac{1}{M})})$. Since the entropy of the sequence is higher when K is even, we prioritize choosing K as an even number and calculate $K = 2a$. If K is odd, we compute $K = 2a + 1$. The value of $P(T)$ is determined by the first $\lceil \log_2(l_{\mathcal{H}}) \rceil$ odd bits of the quantized sequence and the current round number C of the bit sequence generation process. (details are provided in Subsection V-A step 2)

Specifically, when $N = M = 127$ bits, for $\mathcal{SET} = 1$, the maximum entropy value is approximately 0.95 when $K = 1$ (based on the formula 23). For $\mathcal{SET} = 2$, the calculation yields $a \in (0, 2.75)$, with the maximum entropy value of approximately 1 being achieved when $K = 2$ and $K = 4$. In this case, selecting $\mathcal{SET} = 2$ and choosing $K = 2$ or $K = 4$ results in the highest randomness for the generated bit sequences.

VI. Performance Analysis of VBF-PST

In this section, we primarily introduce the theoretical basis for VBF-PST scheme's resistance against differential attacks and proximity attacks, and analyze the communication overhead of the scheme.

A. Differential Attack

In VBF-PST, according to Step 6, Eve can intercept the ciphertext transmitted in two consecutive secure transmissions over the public channel. The ciphertexts adhere to the following equation:

$$\mathbf{e}'(T_1) = \mathbf{d}'_{\text{con}}(T_1) \oplus \mathbf{q}^{BF}_{a,b}(T_1), \tag{26}$$

$$\mathbf{e}'(T_2) = \mathbf{d}'_{\text{con}}(T_2) \oplus \mathbf{q}^{BF}_{a,b}(T_2). \tag{27}$$

By observing $\mathbf{e}'(T_1)$ and $\mathbf{e}'(T_2)$, the eavesdropper Eve can obtain the following equation:

$$\begin{aligned}\mathbf{e}'_{attack} &= \mathbf{e}'(T_1) \oplus \mathbf{e}'(T_2) \\ &= (\mathbf{d}'_{con}(T_1) \oplus \mathbf{q}^{BF}_{a,b}(T_1)) \oplus (d'_{con}(T_2) \oplus \mathbf{q}^{BF}_{a,b}(T_2)) \\ &= (\mathbf{d}'_{con}(T_1) \oplus \mathbf{d}'_{con}(T_2)) \oplus (\mathbf{q}^{BF}_{a,b}(T_1) \oplus q^{BF}_{a,b}(T_2)) \\ &= (\mathbf{d}'_{con}(T_1) \oplus \mathbf{d}'_{con}(T_2)) \oplus \triangle\mathbf{q}^{BF}_{ab}(T_1, T_2)),\end{aligned} \tag{28}$$

where $\triangle\mathbf{q}^{BF}_{ab}(T_1, T_2)$ represents the dissimilarity of $\mathbf{q}^{BF}_{a,b}(T_1)$ and $\mathbf{q}^{BF}_{a,b}(T_2)$ in the quantized bit sequence.

In VBF-PST, both communicating parties undergo processing by a VBF before engaging in secure transmission at times T_1 and T_2. Moreover, the filter employs time-varying parameter $P(T)$. Therefore, even if the input sequences at T_1 and T_2 exhibit strong correlation, the variation in the random mapping functions used at T_1 and T_2 weakens the correlation between $\mathbf{q}^{BF}_{a,b}(T_1)$ and $\mathbf{q}^{BF}_{a,b}(T_2)$, resulting in a higher BRR $\aleph^{BF}_{\mathbf{q}}(T_1, T_2)$ compared to $\aleph_{\mathbf{q}}(T_1, T_2)$. In the worst-case scenario where $\mathbf{q}_{a,b}(T_1)$ and $\mathbf{q}_{a,b}(T_2)$ are identical, the BRR between $\mathbf{q}^{BF}_{a,b}(T_1)$ and $\mathbf{q}^{BF}_{a,b}(T_2)$ obtained after processing by the VBF is certainly greater than 0. Consequently, in the same static environment, VBF-PST is expected to offer better resistance against differential attacks compared to previous secure transmission schemes based on incomplete identical bit sequences.

B. Proximity Attack

In VBF-PST, The eavesdropper Eve can utilize the quantized value $\mathbf{q}_{a,e}(T)$ to obtain $\mathbf{q}^{BF}_{a,e}(T)$ through a BF. Let ϵ_e denote the bit disagreement rate between $\mathbf{q}_{a,e}(T)$ and $\mathbf{q}_{a,b}(T)$, and ϵ^{BF}_e denote the bit disagreement rate between $\mathbf{q}^{BF}_{a,e}(T)$ and $\mathbf{q}^{BF}_{a,b}(T)$.

In Subsection V-A, it has been demonstrated that the VBF-PST scheme does not amplify the differences in the original quantized bits, which also implies that VBF-PST does not increase Eve's decoding difficulty. If hash collisions exist, the VBF-PST scheme may even reduce the disparity between $\mathbf{q}^{BF}_{a,b}(T)$ and $\mathbf{q}^{BF}_{b,a}(T)$. This means that if ECC with strong error correction capabilities is used in FEC, Eve may potentially exploit ECC to recover the confidential message.

In this paper, we select Bose–Chaudhuri–Hocquenghem (BCH) codes $(\mathrm{n}, \mathrm{k}, \mathrm{t})$ as the ECC due to their lightweight and low complexity [38], making them widely used [39], [40]. To better resist proximity attacks, we propose an adaptive error correction parameter selection scheme in this paper. This scheme determines the appropriate BCH$(\mathrm{n}, \mathrm{k}, \mathrm{t})$ used during communication based on the position of the eavesdropper Eve and the length of the output sequence $L^{BF}_{\mathbf{q}}$. This ensures that legitimate communication parties, Alice and Bob, can recover the confidential message without errors while the eavesdropper Eve cannot, thus resisting proximity attacks. The parameters n and t are then determined based on the calculated results $L^{BF}_{\mathbf{q}}$ and ϵ^{BF}_q of the BF. **Algorithm 2** summarizes the adaptive error correction parameter selection algorithm.

This section assumes an actual transmission channel error ratio of $\epsilon_0 = 0.01$ and $L_{\mathbf{q}} = L^{BF}_{\mathbf{q}}$. Specifically, assuming Alice and Bob have a maximum sampling delay τ of 0.01s,

Algorithm 2: Adaptive ECC Parameter Selection Scheme

Input: The confidential message $\mathbf{d}'_{\rm con}, \tau, D, L_{\mathbf{q}}^{BF}$
Output: The parity check bits $\mathbf{s}'$

1 If $L_{\mathbf{q}}^{BF} = 2^m - 1$, where $m \in \{3, 4, \ldots, 16\}$, then $\mathrm{n} = L_{\mathbf{q}}^{BF}$.
2 If $L_{\mathbf{q}}^{BF} = 2^m - 1$, where $m \in \{3, 4, \ldots, 16\}$, then $\mathrm{n} = 2^m - 1$.
3 Calculate the autocorrelation coefficient ρ_{ab} and the eavesdropper's channel correlation coefficient ρ_{ae} based on τ and D.
4 Compute the statistical error bits ϵ_q and ϵ_e based on ρ_{ab} and ρ_{ae}.
5 Calculate the range of t and select t such that: $\mathrm{t} \geq \epsilon_{eq}^{BF} L_{\mathbf{q}}^{BF} = (1-2\epsilon_0)\epsilon_q^{BF} L_{\mathbf{q}}^{BF} + \epsilon_0 L_{\mathbf{q}}^{BF} \geq (1-2\epsilon_0)\epsilon_q L_{\mathbf{q}} + \epsilon_0 L_{\mathbf{q}}^{BF}$ $\mathrm{t} < (1-2\epsilon_0)\epsilon_e^{BF} L_{\mathbf{q}}^{BF} + \epsilon_0 L_{\mathbf{q}}^{BF}$
6 Assuming no hash collisions in the BF, t satisfies: $(1-2\epsilon_0)\epsilon_q L_{\mathbf{q}} + \epsilon_0 L_{\mathbf{q}}^{BF} \leq \mathrm{t} < (1-2\epsilon_0)\epsilon_e L_{\mathbf{q}} + \epsilon_0 L_{\mathbf{q}}^{BF}$, We choose t as:$\mathrm{t} = (1-2\epsilon_0)\big(\epsilon_e + \epsilon_q\big) L_{\mathbf{q}}/2 + \epsilon_0 L_{\mathbf{q}}^{BF}$.
7 Based on n and t, determine the appropriate parameters for the BCH, where $\mathrm{k} \leq \mathrm{n} - 2\mathrm{t} - 1$. Encode the confidential message $\mathbf{d}'_{\rm con}$ using the $\mathrm{BCH}(\mathrm{n}, \mathrm{k}, \mathrm{t})$ code to obtain the parity check bits $\mathbf{s}'$.

a maximum Doppler frequency f_d of 5 Hz, and a signal-to-noise ratio of 25 dB, the range of D values is $(0.05\lambda$,$0.5\lambda)$. The auto-correlation coefficient ρ_{ab} between Alice and Bob is assumed to be $\geq J_0(2\pi f_d \tau) = 0.9755$. Consequently, $\epsilon_q \approx 0.05$.

Suppose Eve is at a distance of $D = 0.1\lambda$ from Bob, then $\rho_{ae} = J_0(2\pi \triangle D) = 0.9$, and $\epsilon_e \approx 0.17$. When $L_{\mathbf{q}}^{BF} = 128$, the reference range for t is $[7, 22)$ and the calculated values are $\mathrm{n} = 127$ and $\mathrm{t} = 15$. When $L_{\mathbf{q}}^{BF} = 255$, the reference range for t is $[13, 45)$ and the calculated values are $\mathrm{n} = 255$ and $\mathrm{t} = 30$.

Suppose Eve is at a distance of $D = 0.18\lambda$ from Bob, then $\rho_{ae} = J_0(2\pi \triangle D) = 0.7$, and $\epsilon_e \approx 0.23$. When $L_{\mathbf{q}}^{BF} = 128$, the reference range for t is $[7, 30)$ and the calculated values are $\mathrm{n} = 127$ and $\mathrm{t} = 19$. When $L_{\mathbf{q}}^{BF} = 255$, the reference range for t is $[13, 60)$ and the calculated values are $\mathrm{n} = 255$ and $\mathrm{t} = 37$.

C. Communication Overhead

VBF-PST reduces communication overhead by omitting the information reconciliation step, while addressing security concerns through the application of a variable Bloom filter to the quantized bit sequence, followed by the use of unidentical keys for OTP secure communication. Therefore, it is essential to compare the communication overhead of the VBF-PST scheme with that of traditional OTP secure schemes that utilize identical keys. This paper specifically considers a comparison of communication overhead between the VBF-PST scheme and the identical key-based physical-layer secure transmission (IK-PST) scheme, which represents a straightforward method of directly cascading key generation and OTP. In contrast to the UK-PST scheme, the IK-PST scheme includes an additional information reconciliation and privacy amplification step.

We consider the communication overhead caused by the information transmissions from Alice to Bob. IK-PST needs two times of the information transmission while VBF-PST only needs one information transmission. We measure the communication overhead by the interaction delay T_{delay}. The delays for both schemes are calculated as

$$\begin{aligned} T_{delay}^{IK} &= \left(\frac{L_{\mathbf{s}} + L_0}{B} + \frac{\mathrm{dist}}{c}\right) + \left(\frac{L_{\mathbf{d}} + L_0}{B} + \frac{\mathrm{dist}}{c}\right) \\ &= \frac{L_{\mathbf{q}}}{B} + 2\left(T_0 + \frac{\mathrm{dist}}{c}\right) \end{aligned} \tag{29}$$

and

$$T_{dalay}^{BF} = \frac{L_{\mathbf{q}} + L_0}{B} + \frac{\mathrm{dist}}{c} = \frac{L_{\mathbf{q}}}{B} + T_0 + \frac{\mathrm{dist}}{c}, \tag{30}$$

respectively, where B is the system bandwidth, $dist$ is the transmission distance, c is the velocity of light, and L_0 is the indispensable overhead in a frame, e.g., the synchronization header, PHY header, and frame payload and $T_0 = L_0/B$ is the time cost by transmitting these bits. As observed from (29) and (30)

$$T_{dalay}^{BF} < T_{delay}^{IK}. \tag{31}$$

D. Computation Complexity

Computation complexity is of critical importance, especially for systems with limited resources. The VBF-PST scheme eliminates the need for a complex reconciliation phase and instead utilizes dissimilar binary sequences as the keys for encryption and decryption. The discrepancies between these keys introduce errors during the recovery process of d. These errors are analogous to transmission errors caused by channel distortions. Therefore, we may regard these errors as part of the equivalent channel errors and integrate the error correction task with the existing channel coding of the system. The computation complexity is primarily determined by the complexity of the decoder and the VBF.

In this paper, we select the BCH code $\mathcal{C}(n, k, t)$ as the error-correcting code, with the upper and lower bounds of the decoder complexity [41] given by:

$$\zeta^{\rm UB}(\mathcal{C}) = \left(45k^2 + 4k\right) n^2 (\log n)^2, \tag{32}$$

$$\zeta^{\rm LB}(\mathcal{C}) = 45k^2 n^2 (\log n)^2. \tag{33}$$

The IK-PST scheme uses two BCH codes, $\mathcal{C}_0(n_0, k_0, t_0)$ for the physical channel with error probability ϵ_0 and $\mathcal{C}(n, k, t)$ for the information reconciliation with ϵ_q. Therefore, we use

$$\zeta_{\rm IK}^{\rm UB} = \zeta^{\rm UB}(\mathcal{C}_0) + \zeta^{\rm UB}(\mathcal{C}), \tag{34}$$

$$\zeta_{\rm IK}^{\rm LB} = \zeta^{\rm LB}(\mathcal{C}_0) + \zeta^{\rm LB}(\mathcal{C}) \tag{35}$$

for approximate calculation of complexity of IK-PST.

In the VBF-PST scheme, a BCH code $\mathcal{C}'(n', k', t')$ and a VBF $BF\{\mathcal{H}, K, P(T), M, \mathcal{SET}\}$ are employed. Assuming the actual communication channel error rate is ϵ_0, ϵ_q is the BDR of the sequence before the filter input, and ϵ_q^{BF} is the

BDR of the sequence after the filter output, the equivalent channel error rate [36] is given by:

$$\epsilon_{eq} = \epsilon_0 + \epsilon_q^{BF} - 2\epsilon_0\epsilon_q^{BF}. \tag{36}$$

When $\epsilon_0 = 0$, we have $\epsilon_{eq} = \epsilon_q^{BF}$. When $\epsilon_0 > 0$ and the length of the sequence input to the filter is N, the computational complexity of the BF is given by $\zeta_{\mathrm{BF}} = KN$. The upper and lower bounds of the computation complexity for the VBF-PST scheme are approximately:

$$\zeta_{\mathrm{BF}}^{\mathrm{UB}} = \zeta^{\mathrm{UB}}(\mathcal{C}') + \zeta_{\mathrm{BF}}, \tag{37}$$

$$\zeta_{\mathrm{BF}}^{\mathrm{LB}} = \zeta^{\mathrm{LB}}(\mathcal{C}') + \zeta_{\mathrm{BF}}. \tag{38}$$

Based on the research context of this paper, the reasonable ranges for the parameters are as follows: K is within $[1, 5]$, N is within $[128, 256]$, k' is within $[36, 100]$, and n' is within $[128, 256]$. Calculations yield $\zeta_{\mathrm{BF}} \in [128, 1280]$, $\zeta^{\mathrm{UB}}(\mathcal{C}) \in [4.699\times10^{11}, 1.887\times10^{12}]$, and $\zeta^{\mathrm{LB}}(\mathcal{C}) \in [4.699\times10^{11}, 1.887\times10^{12}]$. Therefore, regardless of the specific values of K, N, k' and n' within their respective ranges, $\zeta^{\mathrm{UB}}(\mathcal{C})$ and $\zeta^{\mathrm{LB}}(\mathcal{C})$ are consistently significantly larger than ζ_{BF}. It can be derived that $\zeta^{\mathrm{UB}}(\mathcal{C}) > \zeta_{\mathrm{BF}}$ and $\zeta^{\mathrm{LB}}(\mathcal{C}) > \zeta_{\mathrm{BF}}$. Therefore, the conclusions are drawn that

$$\zeta_{\mathrm{BF}}^{\mathrm{UB}} < \zeta_{\mathrm{IK}}^{\mathrm{UB}}, \tag{39}$$

$$\zeta_{\mathrm{BF}}^{\mathrm{LB}} < \zeta_{\mathrm{IK}}^{\mathrm{LB}}. \tag{40}$$

VII. Simulation Results

In this section, we evaluate the performance of the VBF-PST scheme from the perspectives of consistency, randomness, and security through simulation. We adopt four evaluation metrics: BDR, BRR, random tests and the probability of successfully recovering secret information in the scheme.

A. Simulation Parameters

The LTE EPA (Extended Pedestrian A) model, a pedestrian channel model extension in the 3GPP protocol [42], is employed in this paper to generate uplink and downlink channels and is chosen to emulate a pedestrian scenario with low mobility, characterized by a maximum Doppler shift of 5 Hz. The channel bandwidth is set to 10 MHz. The simulation employs a TDD mode, where uplink and downlink transmissions share the same frequency band but are separated in time. The reference signals used for both uplink and downlink are the Demodulation Reference Signals (DMRS), which are essential for channel estimation and synchronization. The channel estimation process utilizes cubic spline interpolation, which is known for its smoothness and accuracy in estimating channel responses between pilot symbols. The frequency and time window sizes for interpolation are set to 41 and 27, respectively, to balance between smoothing out noise and capturing the rapid variations in the channel. The amplitude of the channel coefficients is used to generate bit sequences. In the simulations, we applied equiprobable two-bit quantization to the channel probing results to generate the key sequences, repeating the simulation 1000 times. Simulation results are compared against the UK-PST scheme described in [7]. The configuration of simulation parameters is outlined in Table I.

TABLE I
Simulation Parameters and Values

Parameter	Value
Channel Model	EPA
Channel Bandwidth	10 MHz
Maximum Doppler Shift	5 Hz
Uplink Reference Signal	DMRS
Downlink Reference Signal	DMRS
Number of Simulations	1000
Quantization Method	Dual-bit Quantization
Signal-to-Noise Ratio	25 dB
Carrier Frequency	2.4 GHz
Eve Channel Correlation	0.7
Duplex Mode	TDD
Channel Estimation Method	Cubic Spline Interpolation
Carrier Number	600

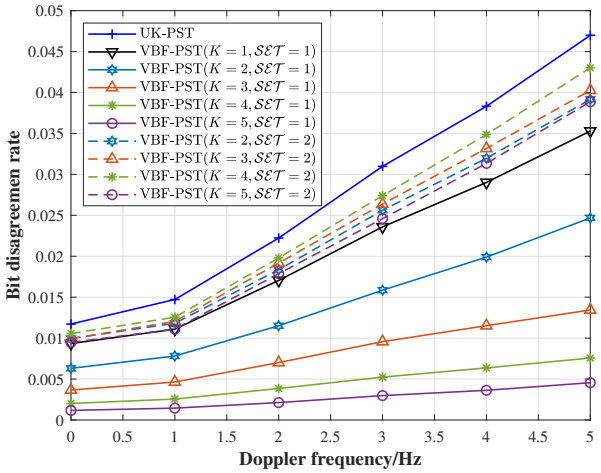


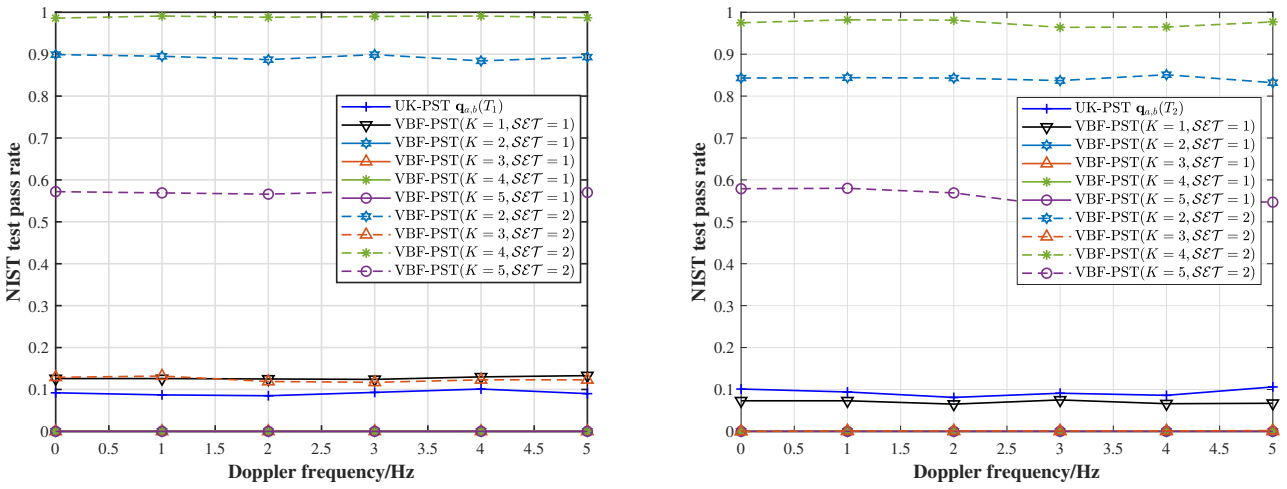
Fig. 8. Comparison of Bit Disagreement Rate between UK-PST and VBF-PST Schemes.

B. Consistency Analysis

The BDR is defined as the number of different bits between Alice and Bob divided by the total number of bits of the raw key, which is calculated using the raw bits before information reconciliation. We calculate the BDR of $\mathbf{q}_{b,a}(T)$ and $\mathbf{q}_{a,b}(T)$ in 1000 simulations. Fig. 8 illustrates the bit disagreement rate of the proposed scheme. The results indicate that under different parameters, the VBF-PST scheme achieves lower bit disagreement rate compared to the UK-PST scheme. Consequently, the number of inconsistent bits output by the filter will not exceed the error detection capability limit of the error-correcting code, thereby preventing the scheme from failing to accurately recover the secret information. This simulation demonstrates that the proposed scheme adheres to the fundamental principle of not amplifying the original quantization result differences (details are provided in Subsection V-A).

C. Randomness Analysis

We use the NIST test suite [43] to evaluate the randomness of our generated key sequences. The idea is to test the NIST pass rates of $\mathbf{q}_{a,b}(T_1)$ and $\mathbf{q}_{a,b}(T_2)$ separately and then plot them. Fig. 9a and Fig. 9b depict the NIST randomness test pass ratio of different schemes at times T_1 and T_2, respectively.



(a) Time T_1 (b) Time T_2

Fig. 9. Comparison of NIST randomness test pass ratio between UK-PST and VBF-PST schemes.

The results indicate that before employing the BF, the NIST randomness test pass ratio of $\mathbf{q}_{a,b}(T)$ in adjacent times T_1 and T_2 for the UK-PST scheme is only 0.1. This implies that out of 10 sets of ciphertexts transmitted securely, only one set of keys $\mathbf{q}_{a,b}(T)$ can pass all NIST randomness tests. In contrast, for the VBF-PST scheme, the NIST randomness test pass ratio are close to 1 for $K = 2$ with $\mathcal{SET} = 2$ and $K = 4$ with $\mathcal{SET} = 2$, indicating that in the VBF-PST scheme, selecting $\mathcal{SET} = 2$ and even values of N_{hash} within the range of $(0, \ln(254))$ yields the best randomness for the obtained keys. This confirms the theoretical findings in Subsection V-C.

For the case of $K = 1, \mathcal{SET} = 1$, where the entropy value H is maximized, the NIST pass ratio is still not high. The primary reason lies in the fundamental differences in how the BF processes the input sequence under these two modes. Under $\mathcal{SET} = 1$, the BF maps each element of the input sequence to the output sequence via hash functions and sets the corresponding positions to 1. Although this method retains some randomness from the input sequence, the high density of 1's in the output sequence, compromises the overall randomness distribution. In contrast, under $\mathcal{SET} = 2$ the BF alternately sets positions in the output sequence to 1 or 0 based on the parity of the hash function indices. Such an approach distributes 1's and 0's bits more evenly, thereby significantly enhancing the randomness of the output sequence. Moreover, the more complex influence of hash function indices on the output sequence under $\mathcal{SET} = 2$ further reduces the clustering of 1's and 0's, thus improving randomness. Therefore, selecting appropriate parameters for the VBF can effectively improve the randomness of the bit sequences.

D. Security Analysis

For the BRR defined in Subsection IV-B, it is noteworthy that for binary bits, a higher BRR does not necessarily mean better resistance against differential attacks. If $\aleph_{\mathbf{q}}(T_1, T_2) = 1$, it implies that if Eve guesses $\mathbf{q}_{a,b}(T_1)$, she can obtain

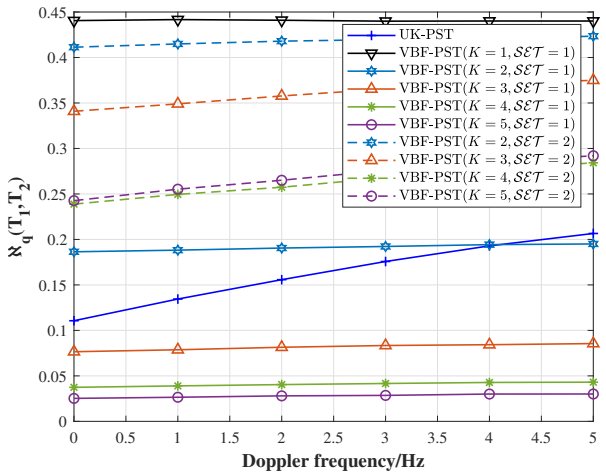
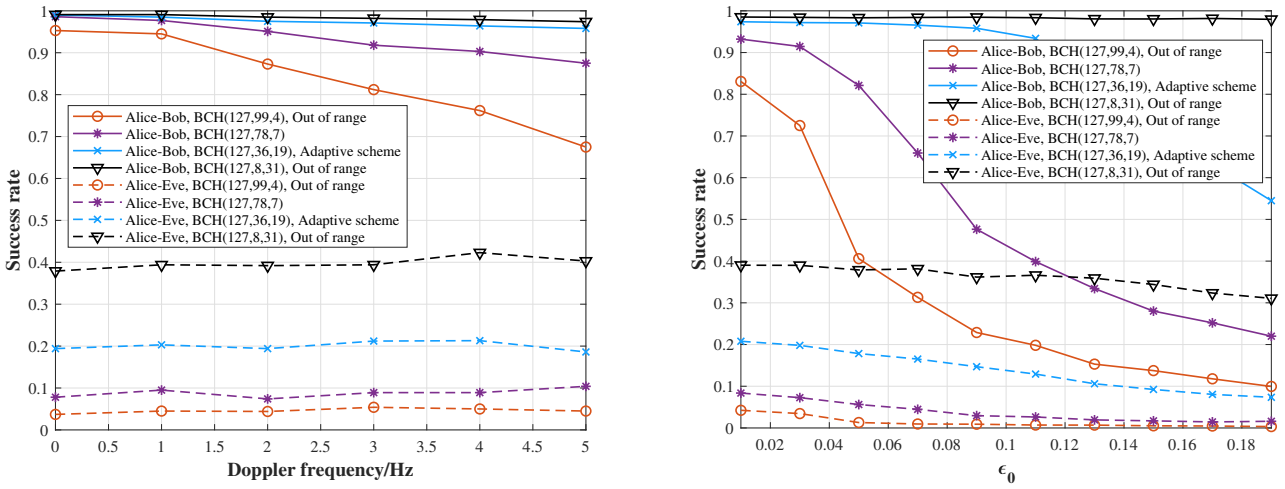


Fig. 10. Comparison of Bit Replacement Ratio between UK-PST and VBF-PST Schemes.

$\mathbf{q}_{a,b}(T_2)$ by simply flipping each bit of $\mathbf{q}_{a,b}(T_1)$. Therefore, this section considers $\aleph_{\mathbf{q}}(T_1, T_2) = 0.5$ as a better defense against differential attacks. We investigated the impact of different filter parameter configurations on the BRR. During the simulation, the Doppler frequency in the experimental environment varied from 0 Hz to 5 Hz in increments of 0.5 Hz, with each frequency point tested 1,000 times to ensure the statistical reliability of the results. Fig. 10 illustrates the BRR of different schemes at adjacent times T_1 and T_2. The results show that $K = 2, \mathcal{SET} = 2$ can maintain the BRR stably around 0.45, while $K = 4, \mathcal{SET} = 2$ achieves an BRR around 0.25, which is superior to that of the UK-PST scheme. The simulation results verify the conclusion in Subsection VI-A that the correlation between $\mathbf{q}_{a,b}^{BF}(T_1)$ and $\mathbf{q}_{a,b}^{BF}(T_2)$ weakens. Therefore, this scheme can enhance the resistance of the quantized bit sequences against differential attacks.

In the simulation of near-end attacks, it is challenging to



(a) Under different Doppler frequencies.

(b) Under different channel errors conditions.

Fig. 11. The probability of successfully recovering secret information in the VBF-PST scheme.

control the correlation between the eavesdropper Eve's channel and the legitimate channel. In this section, noise is added to the legitimate channel to simulate the eavesdropper channel, with a consideration of the correlation coefficient ρ_{ae}. Assuming $BF\{\mathcal{H}, K = 2, q(T), M = 127, \mathcal{SET} = 2\}$ is selected as the VBF in the simulation, and the actual transmission channel error $\epsilon_0 = 0.01$. According to the adaptive ECC parameter selection scheme, the selection range of t is $[7, 22)$. In this section, four BCH codes, namely BCH$(127, 99, 4)$, BCH$(127, 78, 7)$, BCH$(127, 36, 19)$ and BCH$(127, 8, 31)$ are selected as ECCs for 1000 simulations to compare the impact of different ECCs on secret information recovery. Fig. 11 illustrates the probability of successfully recovering secret information under different Doppler frequencies and different channel errors conditions.

Based on Fig. 11a, when $\mathrm{t} = 19$, the legitimate channel successfully recovers the secret information with an average ratio consistently above 0.9, whereas the eavesdropping channel's success ratio remains consistently below 0.2. For BCH coding schemes that exceed the recommended t range in this approach, when $\mathrm{t} = 4$, although the eavesdropping channel's success ratio in recovering the key is only 0.05, the legitimate channel's success ratio rapidly declines to around 0.65 with varying Doppler shifts, indicating poor robustness of the VBF-PST scheme. When $\mathrm{t} = 31$, while the legitimate channel's success ratio in recovering the key improves significantly, the eavesdropping channel's success ratio increases to around 0.4, suggesting that Eve can recover nearly half of the key, which is considered insecure.

Based on Fig. 11b, for the legitimate channel, as t decreases within the recommended range, its robustness against different channel errors diminishes. When $\epsilon_0 = 0.19$ and $\mathrm{t} = 7$, the probability of successfully recovering the key is much lower compared to this approach, where the success ratio of key recovery is maintained above 0.5. Therefore, employing an adaptive scheme to calculate appropriate ECC usage ranges and providing coding parameters t that balance key distribution robustness and resistance to near-end attacks is essential. Simulation results validate that the adaptive error correction code scheme can effectively withstand near-end attacks while ensuring robust key distribution.

VIII. Conclusions

This work is devoted to enhancing the wireless transmission security by exploiting the diversity and the uncertainty of channel characteristics in multipath environments. We have first reviewed the basic principles of existing secure transmission schemes and identified their limitations and security issues in static or slowly-varying environments. Motivated by these, we have proposed a secure transmission scheme based on variable bloom filter-based physical-layer secure transmission, called VBF-PST, to enhance the randomness and bit replacement rate of bit sequences by judiciously constructing the BF. We have also derived a theoretical guideline for selecting optimal parameters for the BF, demonstrating its effectiveness in improving randomness and resisting differential and proximity attacks. Additionally, we have introduced a new method to determine optimal ECC parameters based on channel conditions. Simulation results have shown the proposed VBF-PST effectively improves the randomness of bit sequences, leading to a NIST randomness test pass ratio approaching 1, while maintaining a bit replacement rate of around 0.45, thus possessing resistance against differential attacks.

As a future work, we plan to extend the VBF-PST framework to accommodate more complex and dynamic wireless environments, such as high-mobility or heterogeneous networks. Besides, it is interesting to explore lightweight cryptographic enhancements to further reduce computational overhead and develop real-time implementations of the proposed scheme in resource-constrained IoT systems.

APPENDIX
PROOF OF THE OPTIMAL RANGE OF VALUES FOR a

Let $m = 1 - \frac{1}{M}$. When $K = 2a, P(T) (\mathrm{mod}\ 2) = 1$, the probability of $Y_j = 1$ is $Pr_j(1) = \frac{\mathrm{m}(1-\mathrm{m}^{2aN})}{1+\mathrm{m}}$, and the probability of $Y_j = 0$ is $Pr_j(0) = 1 - Pr_j(1)$. The partial derivative of the entropy H of $\mathbf{Y}$ with respect to a is calculated as follows:

$$\frac{\partial H}{\partial a} = \frac{2N \ln m}{m+1} m^{2aN+1} (\log_2(\mathrm{m}p_1(a)) - \log_2(p_2(a))), \quad (41)$$

where

$$p_1(a) = 1 - \mathrm{m}^{2aN}, \quad (42)$$

$$p_2(a) = 1 + \mathrm{m}^{2aN+1}. \quad (43)$$

When $K = 2a, P(T) (\mathrm{mod}\ 2) = 0$, the probability of $Y_j = 1$ is $Pr_j(1) = \frac{1-\mathrm{m}^{2aN}}{1+\mathrm{m}}$, and the probability of $Y_j = 0$ is $Pr_j(0) = 1 - Pr_j(1)$. The partial derivative of the entropy H of $\mathbf{Y}$ with respect to a is calculated as follows:

$$\frac{\partial H}{\partial a} = \frac{2N \ln m}{m+1} m^{2aN} (\log_2(p_1(a)) - \log_2(p_3(a))), \quad (44)$$

where

$$p_3(a) = \mathrm{m} + \mathrm{m}^{2aN}. \quad (45)$$

When $K = 2a+1, P(T) (\mathrm{mod}\ 2) = 1$, the probability of $Y_j = 1$ is $Pr_j(1) = \frac{(1-\mathrm{m}^{(2a+1)N})(1-\mathrm{m}^{2(a+1)})}{(1-\mathrm{m}^{2a+1})(1+\mathrm{m})}$, and the probability of $Y_j = 0$ is $Pr_j(0) = 1 - Pr_j(1)$. The partial derivative of the entropy H of $\mathbf{Y}$ with respect to a is calculated as follows:

$$\frac{\partial H}{\partial a} = \frac{(p_4(a) + p_5(a)m) 2 \ln m}{p_6(a)(1-\mathrm{m}^{2a+1})} \log_2(\frac{p_6(a)}{p_7(a)} - 1), \quad (46)$$

where

$$p_4(a) = (\mathrm{m}^{2a+1} - 1)(1 - \mathrm{m}^{2a+2}) \mathrm{m}^{(2a+1)N} N, \quad (47)$$

$$p_5(a) = \left(1 - \mathrm{m}^{(2a+1)N}\right)(1-m)\mathrm{m}^{2a}, \quad (48)$$

$$p_6(a) = (1 - \mathrm{m}^{2a+1})(1+m), \quad (49)$$

$$p_7(a) = (1 - \mathrm{m}^{(2a+1)N}). \quad (50)$$

When $K = 2a+1, P(T) (\mathrm{mod}\ 2) = 0$, the probability of $Y_j = 1$ is $Pr_j(1) = \frac{(1-\mathrm{m}^{(2a+1)N})(1-\mathrm{m}^{2a})}{(1-\mathrm{m}^{2a+1})(1+\mathrm{m})}$, and the probability of $Y_j = 0$ is $Pr_j(0) = 1 - Pr_j(1)$. The partial derivative of the entropy H of $\mathbf{Y}$ with respect to a is calculated as follows:

$$\frac{\partial H}{\partial a} = \frac{(p_8(a) - p_5(a)) 2 \ln m}{p_6(a)(1-\mathrm{m}^{2a+1})} \log_2(\frac{p_6(a)}{p_9(a)} - 1), \quad (51)$$

where

$$p_8(a) = (\mathrm{m}^{2a} - 1)(1 - \mathrm{m}^{2a+1}) \mathrm{m}^{(2a+1)N} N, \quad (52)$$

$$p_9(a) = \left(1 - \mathrm{m}^{(2a+1)N}\right)(1 - \mathrm{m}^{2a}). \quad (53)$$

The entropy H is monotonically increasing with respect to a. However, as a increases, K also increases, leading to higher filter complexity. Therefore, we focus on the extremum of the entropy increment to determine the optimal parameter. By setting the partial derivatives of the entropy H for the four cases to zero, we find the extremum of the entropy increment. The conclusion is that when $a \in (0, -\frac{\ln(2M)}{2N \ln(1-\frac{1}{M})})$, the entropy H increment is monotonically increasing. Beyond this range, the entropy H increment decreases, and the filter complexity increases. Therefore, $a \in (0, -\frac{\ln(2M)}{2N \ln(1-\frac{1}{M})})$ is the optimal range for a.

REFERENCES

[1] C. Cheng, R. Lu, A. Petzoldt, and T. Takagi, "Securing the Internet of Things in a quantum world," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 116–120, Feb. 2017.

[2] G. S. Vernam, "Secret signaling system," U.S. Patent 1310719, Jul. 1919.

[3] C. E. Shannon, "Communication theory of secrecy systems," *Bell Labs Tech. J.*, vol. 28, no. 4, pp. 656–715, 1949.

[4] G. Zheng, G. Fang, R. Shankaran, and M. A. Orgun, "Encryption for implantable medical devices using modified one-time pads," *IEEE Access*, vol. 3, pp. 825–836, Jun 2015.

[5] L. Peng, G. Li, J. Zhang, and A. Hu, "Securing M2M transmissions using nonreconciled secret keys generated from wireless channels," in *Proc. IEEE GLOBECOM Workshop Trusted Commun. Phys. Layer Security (TCPLS)*, Abu Dhabi, UAE, Dec. 2018, pp. 1–6.

[6] G. Li, L. Hu, and A. Hu, "Lightweight group secret key generation leveraging non-reconciled received signal strength in mobile wireless networks," in *Proc. IEEE ICC Workshops WPLS*, Shanghai, China, May 2019, pp. 1–6.

[7] G. Li, Z. Zhang, J. Zhang, and A. Hu, "Encrypting wireless communications on the fly using one-time pad and key generation," *IEEE Internet Things J.*, vol. 8, no. 1, pp. 357–369, Jan. 2021.

[8] B. H. Bloom, "Space/time trade-offs in hash coding with allowable errors," *' Commun. ACM*, vol. 13, no. 7, pp. 422–426, Jul. 1970.

[9] W. Xue, D. Vatsalan, W. Hu, and A. Seneviratne, "Sequence data matching and beyond: New privacy-preserving primitives based on Bloom filters," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 2973–2987, 2020.

[10] Ú. Erlingsson, V. Pihur, and A. Korolova, "Rappor: Randomized aggregatable privacy-preserving ordinal response," in *Proc. CCS*, 2014, pp. 1054–1067.

[11] W. Xu, Z. Li, W. Xue, X. Yu, B. Wei, J. Wang, C. Luo, W. Li, and A. Y. Zomaya, "Inaudiblekey: Generic inaudible acoustic signal based key agreement protocol for mobile devices," in *ACM/IEEE IPSN*, 2021, pp. 106–118.

[12] S. N. Premnath, S. Jana, J. Croft, P. L. Gowda, M. Clark, S. K. Kasera, N. Patwari, and S. V. Krishnamurthy, "Secret key extraction from wireless signal strength in real environments," *IEEE Trans. Mobile Comput.*, vol. 12, no. 5, pp. 917–930, May. 2013.

[13] H. Liu, Y. Wang, J. Yang, and Y. Chen, "Fast and practical secret key extraction by exploiting channel response," in *Proc. IEEE INFOCOM*, 2013, pp. 3048–3056.

[14] W. Xi, X.-Y. Li, C. Qian, J. Han, S. Tang, J. Zhao, and K. Zhao, "KEEP: Fast secret key extraction protocol for D2D communication," in *Proc. IEEE 22nd Int. Symp. Qual. Service (IWQoS)*, May 2014, pp. 350–359.

[15] Y. Liu, S. C. Draper, and A. M. Sayeed, "Exploiting channel diversity in secret key generation from multipath fading randomness," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 5, pp. 1484–1497, Oct 2012.

[16] J. Zhang, A. Marshall, R. Woods, and T. Q. Duong, "Secure key generation from OFDM subcarriers' channel responses," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec 2014, pp. 1302–1307.

[17] N. Aldaghri and H. Mahdavifar, "Fast secret key generation in static environments using induced randomness," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2018, pp. 1–6.

[18] Aldaghri, Nasser and Mahdavifar, Hessam, "Physical layer secret key generation in static environments," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 2692–2705, Feb. 2020.

[19] H. Chen, Z. Wang, Y. Li, R. Yang, Y. Zhao, R. Zhou, and K. Zheng, "Deep learning-based bloom filter for efficient multi-key membership testing," *Data Science and Engineering*, vol. 8, no. 3, pp. 234–246, 2023.

[20] H. Quan, B. Wang, M. Li, and I. Leontiadis, "FastReach: A system for privacy-preserving reachability queries over location data," *Computers & Security*, vol. 135, p. 103513, 2023.

[21] H. Dai, M. Shahzad, A. X. Liu, and Y. Zhong, "Finding persistent items in data streams," *Proc. VLDB Endowment*, vol. 10, no. 4, pp. 289–300, 2016.

[22] Y. Peng, J. Guo, F. Li, W. Qian, and A. Zhou, "Persistent bloom filter: Membership testing for the entire history," in *Proc. Int. Conf. Manage. Data*, 2018, pp. 1037–1052.

[23] J. Li, Z. Li, Y. Xu, S. Jiang, T. Yang, B. Cui, Y. Dai, and G. Zhang, "Wavingsketch: An unbiased and generic sketch for finding top-k items in data streams," in *Proc. 26th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2020, pp. 1574–1584.
[24] A. Papadopoulos and D. Katsaros, "A-tree: Distributed indexing of multidimensional data for cloud computing environments," in *Proc. IEEE 3rd Int. Conf. Cloud Comput. Technol. Sci.*, 2011, pp. 407–414.
[25] S. Xiong, Y. Yao, S. Li, Q. Cao, T. He, H. Qi, L. Tolbert, and Y. Liu, "kBF: Towards approximate and bloom filter based key-value storage for cloud computing systems," *IEEE Trans. Cloud Comput.*, vol. 5, no. 1, pp. 85–98, Jan. 2017.
[26] M. Mitzenmacher, "A model for learned bloom filters and optimizing by sandwiching," *Proc. 32nd Int. Conf. Neural Inf. Process. Syst*, vol. 31, pp. 462–471, 2018.
[27] Q. Liu, L. Zheng, Y. Shen, and L. Chen, "Stable learned bloom filters for data streams," *Proc. VLDB Endowment*, vol. 13, no. 12, pp. 2355–2367, Aug. 2020.
[28] J. Anderson, Q. Huang, W. Krichene, S. Rendle, and L. Zhang, "Superbloom: Bloom filter meets transformer," *CoRR*, 2020. [Online]. Available: https://arxiv.org/abs/2002.04723
[29] R. Patgiri, A. Biswas, and S. Nayak, "deepBF: Malicious URL detection using learned bloom filter and evolutionary deep learning," *Computer Communications*, vol. 200, pp. 30–41, 2023.
[30] N. Dayan, M. Athanassoulis, and S. Idreos, "Optimal bloom filters and adaptive merging for LSM-trees," *ACM Trans. Database Syst.*, vol. 43, no. 4, p. 16:1–16:48, 2018.
[31] S. Luo, S. Chatterjee, R. Ketsetsidis, N. Dayan, W. Qin, and S. Idreos, "Rosetta: A robust space-time optimized range filter for key-value stores," in *Proceedings of SIGMOD*, online conference Portland, OR, USA, June 14-19, 2020, pp. 2071–2086.
[32] Y. Chai, Y. Chai, X. Wang, H. Wei, and Y. Wang, "Adaptive lower-level driven compaction to optimize LSM-tree key-value stores," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 6, pp. 2595–2609, Jun. 2020.
[33] R. Schnell and C. Borgs, "Randomized response and balanced bloom filters for privacy preserving record linkage," in *Proc. IEEE 16th Int. Conf. Data Mining Workshops (ICDMW)*, Dec. 2016, pp. 218–224.
[34] R. Schnell, T. Bachteler, and J. Reiher, "Privacy-preserving record linkage using Bloom filters," *BMC Med. Informat. Decis. Making*, vol. 9, no. 1, pp. 1–11, Dec. 2009.
[35] C. Sun, H. Yang, and G. Li, "AmpRmdr: Remainder-based secret key generation using wireless channel amplitude," *IEEE Access*, vol. 8, pp. 228 178–228 187, 2020.
[36] T. M. Cover, *Elements of information theory*. Hoboken, NJ, USA: John Wiley & Sons, 1999.
[37] X. He, H. Dai, W. Shen, P. Ning, and R. Dutta, "Toward proper guard zones for link signature," *IEEE Trans. Wireless Commun.*, vol. 15, no. 3, pp. 2104–2117, Mar. 2016.
[38] J. Massey, "Shift-register synthesis and BCH decoding," *IEEE Trans. Inf. Theory*, vol. 15, no. 1, pp. 122–127, Jan. 1969.
[39] H. Mani and S. Hemati, "Symbol-level stochastic Chase decoding of Reed-Solomon and BCH codes," *IEEE Trans. Commun.*, vol. 67, no. 8, pp. 5241–5252, Aug. 2019.
[40] L. Deng, Z. Liu, Y. L. Guan, X. Liu, C. A. Aslam, X. Yu, and Z. Shi, "Perturbed adaptive belief propagation decoding for high-density parity-check codes," *IEEE Trans. Commun.*, vol. 69, no. 4, pp. 2065–2079, Apr. 2021.
[41] B. G. Bajoga and W. Walbesser, "Decoder complexity for BCH codes," in *Proc. Inst. Elect. Eng.*, vol. 120, no. 4, 1973, pp. 429–431.
[42] 3GPP, "Proposal for LTE channel models," 3rd Generation Partnership Project (3GPP), meeting, 5 2007. [Online]. Available: https://portal.3gpp.org/ngppapp/CreateTdoc.aspx?mode=view&contributionId=137066
[43] A. Rukhin, J. Soto, and Nechvatal, *A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications, Rev. 1a,*. NIST Standard SP 800-22, Apr. 2010.