An LDPC Code Based Physical Layer Message Authentication Scheme With Prefect Security

Dajiang Chen, Member, IEEE, Ning Zhang, Member, IEEE, Rongxing Lu, Senior Member, IEEE, Xiaojie Fang, Senior Member, IEEE, Kuan Zhang, Member, IEEE, Zhiguang Qin, Member, IEEE, and Xuemin Shen, Fellow, IEEE

Abstract—In this paper, we study physical layer message authentication with perfect security for wireless networks, regardless of the computational power of adversaries. Specifically, we propose an efficient and feasible authentication scheme based on low-density parity-check (LDPC) codes and ϵ -AU₂ hash functions over binary-input wiretap channel. First, a multimessage authentication scheme for noiseless main channel case is presented by leveraging a novel ϵ -AU₂ hash function family and the dual of large-girth LDPC codes. Concretely, the sender Alice first generates a message tag T with message M and key Kby using a lightweight ϵ -AU₂ hash functions; then Alice encodes T to a codeword \overline{X}^n with the dual of large-girth LDPC codes; finally, Alice sends (M, X^n) to the receiver Bob noiselessly. An adversary Eve has infinite computational capacity, and he can obtain M and the output Z^n of the BEC with input X^n . Then, an authentication scheme over binary erasure channel and binary-input wiretapper's channel is further developed, which can reduce the noisy main channel case to noiseless main channel case by leveraging public discussion. We theoretically prove that, the proposed schemes are perfect secure if the number of attacks from Eve is upper bounded by a polynomial times in terms of n. Furthermore, the simulation results are provided to demonstrate that the proposed schemes can achieve high authentication rate with low time latency.

Index Terms— Physical layer security, message authentication, binary-input wiretap channel, LDPC codes.

I. INTRODUCTION

D^{UE} to the broadcast characteristic of wireless medium, wireless communication systems (e.g., 5G networks,

Manuscript received September 12, 2017; revised January 31, 2018; accepted February 16, 2017. Date of publication April 9, 2018; date of current version July 9, 2018. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada, NSFC, under Grant 61502085 and Grant 61520106007 and in part by the China Post-Doctoral Science Foundation under Project 2015M570775. (*Corresponding author: Ning Zhang.*)

D. Chen and Z. Qin are with the School of Information and Software Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: dajiang.chen@uwaterloo.ca; qinzg@uestc.edu.cn).

N. Zhang is with the Department of Computing Science, Texas A&M University at Corpus Christi, Corpus Christi, TX 78412 USA (e-mail: ning.zhang@tamucc.edu).

R. Lu is with the School of Computer Science, University of New Brunswick, Fredericton, NB E3B 5A3, Canada (e-mail: rlu1@unb.ca).

X. Fang is with the Department of Electronics and Information Engineering, Harbin Institute of Technology, Harbin 150001, China (e-mail: fangxiaojie@ hit.edu.cn).

K. Zhang is with the Department of Electrical and Computer Engineering, University of Nebraska-Lincoln, NE 68588 USA (e-mail: kzhang22@unl.edu).

X. Shen is with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: sshen@uwaterloo.ca).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSAC.2018.2825079

vehicular communication networks, and e-health system as discussed in [1] and [2]) face several serious security issues [3], including modification attack, substitution attack, and replay attack. It is a core requirement for wireless systems to provide data integrity and identification with high level of security, which guarantee that the data is not changed in transit and is from the stated sender. Message authentication techniques as proposed in [4] are the typical approaches to ensuring data integrity and identification. In cryptographic systems, Public Key Infrastructure (PKI) based digital signature authentication is usually adopted for message authentication.

However, these traditional crypto-based message authentication schemes have several limitations: 1) due to the dynamic topology of wireless networks (i.e., vehicle networks, and massive IoT) and the energy constrained end devices (i.e., sensors, and mobile phones), frequent key distribution in wireless networks is problematic; and 2) PKI-based schemes heavily rely on the computational hardness of certain mathematical problems and the condition that the adversaries have finite computation power. As the computation power of adversaries keep increasing, those systems become vulnerable. Multiple messages authentication with perfect security (or information-theoretic security/ unconditional security) as proposed in [5] can address these issues. Nevertheless, when performing multiple-message authentication with perfect secrecy over noiseless channel model as derived in [4], it is proved in [5] that the probability for successful attacks is at least $2^{-H(K)/(\ell+1)}$ after *l* times of authentication, which quickly approaches 1 as *l* increases.

Physical layer (PHY-layer) based multiple messages authentication with secret key can achieve perfect security by exploring the characteristics of the bottom layer, e.g., the channel, signal, and hardware, as discussed in [9]. It can provide security even when adversaries have infinite power, and it does not require frequent key exchange. Actually, great research efforts have been devoted on multiple-message authentication achieving perfect security over noisy channel model by using PHY-layer based techniques as proposed in [10], [11], [29], and [30]. However, most of the existing works still focus on theoretical study by using the random coding techniques, and an efficient and feasible authentication solution is urgently needed based on coding scheme with low complexity encoding and decoding algorithms.

In this paper, we study message authentication, where Alice sends multiple messages to Bob in presence of the adversary Eve. We aim to propose an efficient and practical multiple-message authentication scheme by using lightweight

0733-8716 © 2018 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

 ϵ -AU₂ hash functions and Low-desnsity parity-check (LDPC) codes [12], to achieve perfect security with the same secret key *K* over binary-input wiretap channel (BIWC) model. Actually, LDPC codes have been adopted in many wireless systems, especially for 5G systems as discussed in [13]. In this way, the proposed authentication scheme is more suited to 5G. We consider binary-input wiretap channel (BIWC) model, since it is a very general model and many practical wiretap scenarios, such as binary erasure wiretap channel (BEWC), binary symmetric wiretap channel (BSWC), and binary-input Gaussian wiretap channel (BIGWC) belong to such a channel model (details please refer to Sec. III).

We first study multiple-message authentication over noiseless main channel case, where the wiretapper's channel is a binary erasure channel (BEC). It is considered that Eve has infinite computational capability, and can receive (M, Z^n) , where M is the message sent from Alice to Bob, and Z^n is the output of BEC from Alice to Eve given the input X^n . Eve aims to forge a message M and a code X^n , and he is successful if Bob accepts M as a valid message. A multiplemessage scheme is proposed as follows. Alice 1) generates a message tag T with M and K by using ϵ -AU₂ hash functions with high efficiency; 2) leverages the large-girth LDPC codes to encode T to X^n ; and 3) transmits (M, X^n) to Bob though noiseless channel. Upon receiving (M, X^n) , Bob decides to reject or accept the authentication by checking the consistency of (M, X^n) . The conditions/requirements for the proposed scheme achieving perfect security is discussed. The theoretical result shows that, to achieve perfect security, the family of hashing functions and the LDPC code used in this scheme should satisfy certain requirements (please refer to Sec. V-B).

Based on the aforementioned results, we further study the multiple-message authentication for noisy main channel case, where both the main channel and the wiretapper's channel are BECs. A novel solution is proposed for this case by using public discussion. The main idea is to reduce the noisy main channel case to a noiseless main channel case with information interaction over a noiseless but insecure channel (i.e., public discussion). For authentication over binary erasure channel and binary-input wiretapper's channel, it can be generalized from the method for the case above by using stochastically degraded channel technique (please refer to Sec. VI).

For implementation of the proposed schemes, a lightweight ϵ -AU₂ class of hash functions algorithm is proposed by using fast multiplication algorithm in finite field $G(2^{\theta})$. Moreover, the construction of the sequence of large-girth LDPC codes is also discussed to meet the requirement of the proposed schemes. The theoretical analysis shows that, if Alice authenticates a polynomial number of messages and Eve attacks polynomial times in terms of n, then the presented schemes are perfect secure. Furthermore, the simulation results show that the proposed schemes can achieve a low time cost and high authentication rate. To the best of our knowledge, this work is the first to realize message authentication with perfect security by leveraging lightweight ϵ -AU₂ class of hash functions and large-girth LDPC codes.

The remainder of the paper is organized as follows. The related work is reviewed in Sec. II. Sec. III introduces basic concepts and preliminaries that will be used in this paper. Sec. IV introduces the authentication model and adversary model. In Sec. V, we propose the authentication scheme for noiseless main channel case. We present the authentication scheme for noisy main channel case in Sec. VI. In Sec. VII, we discuss the ϵ -AU₂ hashing construction and the large-girth LDPC codes design. Sec. VIII provides the experimental results. The concluding remarks are provided in Sec. IX.

II. RELATED WORK

A. LDPC Codes for PHY-Layer Security

PHY-layer security [5]–[8] has become an emerging technique to improve the security of wireless communication by leveraging the characteristics of the wireless channel, secure channel coding, etc. The related works can be traced back to Wyner's work on wiretap channel model [6], which demonstrates that perfect security can be achieved when wiretap channel was a degraded version of the main channel. Later, Csiszár and Köner [7] generalized Wyner's result using random coding techniques, in which the wiretap channel is not necessary to be a degraded version of the main channel. Since then, a large number of theoretical research on secrecy capacity under different wiretap channel models were conducted [14], [15].

Based on aforementioned results, coding methods to achieve secure transmission over wiretap channel were proposed [18]-[23]. Ozarow and Wyner [18] presented the condition for constructing codes for the modified wiretap channel. As a pioneering work, Thangaraj et al. [19] proposed a coset coding scheme by using the dual of low-density parity-check (LDPC) code to achieve weak secrecy over a binary erasure wiretap channel (BEWC). Suresh et al. [20] leveraged the dual of short-cycle-free LDPC code to achieve the strong secrecy over a BEWC. A coding scheme with strongly secure for binary erasure wiretap channel models by using largegirth LDPC codes was presented in [21]. A linear pre-coder to maximize the average secrecy sum rate was proposed for a multiple-input-multiple-output (MIMO) fading cognitive multiple-access wiretap channel in [22]. Moreover, lowcomplexity MIMO precoding for finite-alphabet signals was discussed in [23] Polar code methods for wiretap channel were studied in [24] and [25]. A secrecy capacity achievable polar code method for general degraded and symmetric wiretap channels was proposed in [24]. In [25], another channel coding scheme with polar codes was presented for binary symmetric wiretap channel models.

Different from the above works on security capacity, in this work, we consider coding method for message authentication over wiretap channel with perfect security. Message authentication with secure polar code has been discussed in [31]. However, as mentioned in [21], the threshold phenomenon of LDPC codes is observed at shorter block-lengths than polarization. Accordingly, there is enough interest in studying message authentication with LDPC codes.

B. PHY-Layer Message Authentication

Even though PHY-layer secure transmission has been extensively investigated [14]-[17], the attention to its sibling PHY-layer message authentication is far from enough. Simmons' work in [4] introduced an authentication model over noiseless channels. Message authentication over noise channel models was studied in [26]-[30]. Korzhik et al. [26] discussed authentication over noise source model with a (noiseless) public discussion channel. Recently, Baracca et al. [27] and Ferrante et al. [28] studied authentication over MIMO fading wiretap channels. More recently, the keyless authentication problem over noise channel model was considered by Jiang [29], [30]. Lai et al. [10] studied multiple message authentication to achieve perfect security over wiretap channel. The authentication rate of Lai's method can be bounded by the capacity of the channel from Alice to the adversary. However, these works are based on random coding techniques and have low authentication efficiency. As a result, such kind of works cannot be efficiently implementable in practice. More recently, Liu et al. [32] and Ren et al. [33] presented a physical layer authentication mechanism by using channel state information. However, the proposed schemes cannot provide perfect security.

Different from existing works, this work focuses on designing a practical multi-message authentication scheme over wiretap channels with LDPC codes to achieve perfect security Specifically, we integrate lightweight ϵ -AU₂ class of hash functions and large-girth LDPC codes in the authentication scheme to achieve perfect security.

III. PRELIMINARIES

A. Notions

Random variables (RVs) are denoted by X, Y, \cdots , their realizations are denoted by x, y, \cdots , and their domain are denoted by $\mathcal{X}, \mathcal{Y}, \cdots$. Distance between RVs X and X' over \mathcal{X} is $SD(X; X') = \sum_{x \in \mathcal{X}} |P_X(x) - P_{X'}(x)|$. Conditional distance between X and X given Y is defined as

$$\mathrm{SD}(X|Y;X) = \sum_{y \in \mathcal{Y}} P(y) \sum_{x \in \mathcal{X}} |P(x|y) - P(x)|.$$
(1)

Function negl(n) is negligible in n if for any polynomial poly(n), $lim_{n\to\infty}negl(n)poly(n) = 0$. For any postive integer n, [n] denotes the set $\{1, 2, \dots, n\}$. $|\cdot|$ is the cardinality of the set, and $\lfloor \cdot \rfloor$ is the floor function. \oplus and \odot are addition and multiplication operations between two matrices over finite field GF(2), respectively. For any matrices $A_{n\times r}$ and $B_{m\times r}$, matric $A_{n\times r}^T$ denotes the transposition of $A_{n\times r}$, and

$$[A_{n \times r}; B_{m \times r}] = \begin{bmatrix} A_{n \times r} \\ B_{m \times r} \end{bmatrix}.$$
 (2)

Definition 1: Given two finite sets $\mathcal{M} = \{0,1\}^u$ and $\mathcal{T} = \{0,1\}^v$, a family of functions $\{\phi_k : \mathcal{M} \to \mathcal{T}\}_{k \in \mathcal{K}}$ is ϵ -almost strongly universal (ϵ -ASU₂ for short) if: (1) $\Pr_K(\phi_k(m) = t) = \frac{1}{|\mathcal{T}|}$, for any $m \in \mathcal{M}$ and $t \in \mathcal{T}$; and (2) $\Pr_K(\phi_k(m_1) = t_1, \phi_k(m_2) = t_2) \leq \frac{\epsilon}{|\mathcal{T}|}$, for any distinct $m_1, m_2 \in \mathcal{M}$, and any $t_1, t_2 \in \mathcal{T}$. $\{\phi_k\}_k$ is ϵ -almost universal (ϵ -AU₂ for short)

if the second condition is replaced by: (3) $\Pr_K(\phi(m_1) = \phi(m_2)) \leq \epsilon$ for any distinct $m_1, m_2 \in \mathcal{M}$.

A discrete memoryless binary-input channel (i.e, BIC) is defined as a stochastic matrix $W = \{W(y|x) : x \in \{0,1\}, y \in \mathcal{Y}\}$. The channel W is called binary erasure channel with erasure probability ε , denoted by BEC(ε) for short. The channel W is called binary symmetric channel with cross-over probability ε , denoted by BSC(ε) for short. A Gaussian channel with binary input $\{-1, +1\}$ and noise variance σ^2 is called binary input Gaussian channel, denoted by BIGC($0, \delta^2$). A discrete memoryless binary-input wiretap channel, denoted by BIWC, is defined by two BICs W_1 : $\mathcal{X} \to \mathcal{Y}$ and $W_2 : \mathcal{X} \to \mathcal{Z}$, where $\mathcal{X} = \{0, 1\}$ is the input alphabet from the sender Alice, \mathcal{Y} is the output alphabet at the legitimate receiver Bob, and \mathcal{Z} is the output alphabet at the wiretapper Eve.

B. LDPC Codes

Low-density parity-check (LDPC) codes are linear codes which have at least one spare parity-check matrix [34]. Let $C(\lambda(x), \rho(x))$ be an LDPC ensemble, and $C^n(\lambda(x), \rho(x))$ be an LDPC ensemble with *n* variable nodes, where $\lambda(x) =$ $\sum_{i\geq 1} \lambda_i x^i$ is the left degree distribution, and $\rho(x) =$ $\sum_{i\geq 1} \rho_i x^i$ is the right degree distribution. The degree distributions $\lambda(x)$ and $\rho(x)$ are from an edge perspective, that is, λ_i (ρ_i) is the fraction of edges that connect to variable (check) nodes of degree *i*. In other words, λ_i (resp. ρ_i) is the probability that an edge chosen uniformly at random from the graph is connected to a variable node (resp. check node) of degree *i*.

If each variable node and check nodes have the same degree d_v and d_c , respectively, i.e., $\lambda(x) = x^{d_v-1}$ and $\rho(x) = x^{d_c-1}$, the code in ensemble $\mathcal{C}^n(x^{d_v-1}, x^{d_c-1})$ (denoted by $\mathcal{C}(n, d_v, d_c)$ for short) is called *regular LDPC code*. Each LDPC code can correspond to a bipartite graph, named *Tanner graph* [34, Sec. 2.4], as follows. Let \mathcal{C} be an LDPC code and H be a parity-check matrix of \mathcal{C} with dimensions $m \times n$. The tanner graph of \mathcal{C} is graph $\mathcal{G}(P, E)$ with node set $P = \{c_1, \dots, c_m\} \cup \{v_1, \dots, v_n\}$ and edge set $E = \{(c_i, v_j) | \text{if } H_{ij} = 1\}$, where c_i is the check node and v_j is variable node for any $i \in \{1, 2, \dots, m\}$ and $j \in \{1, 2, \dots, n\}$. The tanner graph with n variable nodes is denoted by $\mathcal{G}(n, P, E)$. The girth of a graph is the length of a shortest cycle contained in the graph.

Definition 2: A sequence of Tanner graphs $\{\mathcal{G}(n, P, E)\}_n$ is *large-girth* if its girth increases as $\log n$. A sequence of LDPC codes is called *large-girth LDPC codes* if the sequence of their corresponding Tanner graphs is large-girth.

IV. MESSAGE AUTHENTICATION OVER WIRETAP CHANNEL

In this section, the authentication model is presented, followed by the adversary model and the definition of secure authentication scheme.

A. Authentication Model

Assume that a sender Alice aims to transmit and authenticates multiple messages to a receiver Bob in the presence

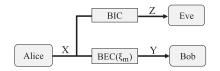


Fig. 1. The authentication channel model.

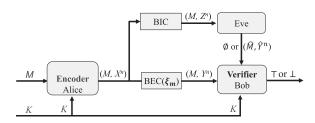


Fig. 2. The authentication model.

of an adversary Eve. As shown in Fig. 1, there is a BEC $W_1 : \mathcal{X} \to \mathcal{Y}$ with erasure probability ξ_m , denoted as $BEC(\xi_m)$, from Alice to Bob, and a BIC $W_2 : \mathcal{X} \to \mathcal{Z}$ from Alice to Eve.

To prevent the attacks from Eve, a secret key K in \mathcal{K} is shared between Alice and Bob before message authentication. Note that, key generation problem over noisy channel has been extensively studied from both theoretical and implementation perspectives, such as [36] and [35]. The authentication key can be obtained with these key generation schemes.

As shown is Fig. 2, if Alice wishes to authenticate a message M in \mathcal{M} , Alice first generates a message authentication code (MAC) X^n with input K, M, and a randomness R, i.e., $X^n = f(M, K, R)$ for an encoding function f; and then, Alice sends M with a error-correcting code over wiretap channel (W_1, W_2) ; finally, Alice transmits X^n to Bob over (W_1, W_2) . After receiving (M, Y^n) , Bob computes $Ver = g(M, K, Y^n)$ with a verify function g, where $Ver \in \{\bot, \top\}$. If $Ver = \top$, Bob accepts M and sends a decision bit 1 to Alice; otherwise, he rejects it and sends a decision bit 0 to Alice. It is assumed that Eve can decode M, and can view the output Z^n of W_2 .

B. Adversary Model

In this work, the adversary model is given as follows. (1) Eve's computing power is considered to be infinite. Eve knows the whole authentication scheme and the values of the parameters, except the secret key shared by Alice and Bob; (2) Eve is allowed to view the output of the channel $W_2: \mathcal{X} \to \mathcal{Z}$, to know the messages authenticated by Alice, and to learn the decision bit for each authentication; and (3) There is a noiseless channel from Eve to Bob, and Eve can send any information to Bob noiselessly. Such an assumption gives more advantages to the adversary. Eve's goal is to forge a message \hat{M} and a MAC \hat{Y}^n such that, Bob accepts \hat{M} as a legal message when he receives (\hat{M}, \hat{Y}^n) over the noiseless channel from Eve to Bob.

We desire to ensure that even if Eve has *adaptively* attacked for polynomial times in terms of n, he still cannot cheat Bob to accept a false authentication. The formal attack model (including Type I and Type II attacks) is as follows. Let M_i ($i = 1, 2, \cdots$) be the sequence of messages authenticated by Alice, and X_i^n be the codeword of M_i . From the adversary model, Eve can receive the message M_i and the output of W_2 , i.e., Z_i^n ; Eve also can learn the decision bit b_i .

Type I: Eve can launch an attack by substituting M'_i for M_i when Alice sends (M_i, X_i^n) to Bob. Here, the forged message M'_i is based on M_i , Eve's local random source R and the information collected previously: $\{(M_j, Z_j^n)\}_{j=1}^{i-1}$ and decision bits $\{b_j\}_{j=1}^{i-1}$ in stage I; as well as $\{(\hat{M}_t, \hat{Z}_t^n)\}_t$ and decision bits $\{\hat{b}_t\}_t$ in stage II below.

Type II: Eve can adaptively send (\hat{M}_t, \hat{X}_t^n) to Bob noiselessly. Eve will learn Bob's decision bit \hat{b}_t . Here (\hat{M}_t, \hat{X}_t^n) is sent according to R and the information is collected previously: $\{(M_j, Z_j^n, b_j)\}_j$ in stage I; and $\{(\hat{M}_j, \hat{Z}_j^n), \hat{b}_j\}_{j=1}^{t-1}$ in stage II.

In this model, we allow that Eve can arbitrarily interleave Type I attacks and Type II attacks. We use *succ* to denote the event that Eve succeeds in a Type I or Type II attack. The model to allow Eve to learn the verification result has been considered in [41]. It is practical as the receiver's action following rejecting or accepting could be visible.

C. Secure Authentication Scheme

A cryptographic scheme is perfect secure if it cannot be broken even if the adversary had unlimited computing power. For the remainder of this article, unless otherwise specified, "secure" means "perfect secure".

Definition 3: A cryptographic scheme Π_n for a wiretap channel $W_1 : \mathcal{X} \to \mathcal{Y}, W_2 : \mathcal{X} \to \mathcal{Z}$ is a secure authentication scheme if the following holds (keep notions in the model).

- 1. Completeness: When the wiretapper Eve does not present, there exists $\alpha > 0$ such that $Pr(D = \bot) \le exp(-n\alpha)$, where n is the number of use of the wiretap channel (W_1, W_2) .
- 2. Authentication: For any wiretapper Eve, the probability of success Pr(Succ(Eve)) is negligible in n after attacking polynomial times in terms of n.

In addition to the security requirement, we also define authentication rate as the efficiency metric. The authentication rate ρ_{auth} is the ratio of the source message length to the codeword length, i.e., $\rho_{auth} = \frac{1}{n} log|\mathcal{M}|$, where $|\mathcal{M}|$ is the cardinality of message space \mathcal{M} .

V. AUTHENTICATION FOR NOISELESS MAIN CHANNEL CASE

In this section, we study authentication over wiretap channel model where the main channel is noiseless. A novel authentication scheme for binary erasure wiretapper's channel is presented. Then, the requirements for the proposed scheme achieving perfect security are discussed.

A. Authentication for Noiseless Main Channel Case

We consider message authentication over binary erasure wiretap channel in which the main channel is noiseless and the wiretap channel is a binary erasure channel BEC(ζ), where ζ is the probability of erasure in the wiretapper's channel.

Setup: Let $\Phi = {\phi_k}_{k \in \mathcal{K}}$ be a collection of ϵ -AU₂ hashing functions from $\mathcal{M} = {0,1}^u$ to $\mathcal{T} = {0,1}^v$. Let \mathcal{C} be a (n,l)linear code with $v \leq n - l$, and G be a generator matrix for \mathcal{C} with rows $\mathbf{g}_1, \mathbf{g}_2, \cdots, \mathbf{g}_l$. Out of the 2^{n-l} cosets of code \mathcal{C} , Alice chooses v linearly independent vectors $\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_v$ from ${0,1}^n/\mathcal{C}$, and announces them to Bob.

Authentication: Assume that Alice and Bob share a secret key $k \in \mathcal{K}$. If Alice desires to authenticate a messages $\mathbf{m} \in \mathcal{M}$ to Bob, they interact as follows.

- 1. Alice first computes $\mathbf{t} = \phi_k(\mathbf{m}) = [t_1, t_2, \cdots, t_v]$ (which is called the *message tag*), and then, encodes \mathbf{t} to x^n using the following steps:
 - selects a vector $\mathbf{s} = [s_1, s_2, \cdots, s_l]$ uniformly at random in vector space $\{0, 1\}^l$;
 - computes x^n by

$$x^{n} = \mathbf{t} \odot [\mathbf{r}_{1}; \mathbf{r}_{2}; \cdots; \mathbf{r}_{v}] \oplus \mathbf{s} \odot [\mathbf{g}_{1}; \mathbf{g}_{2}; \cdots; \mathbf{g}_{l}].$$
(3)

Finally, Alice sends **m** and x^n over wiretap channel W_1, W_2).

2. Assume Bob receives \mathbf{m}' and y^n and Eve receives \mathbf{m} and z^n , respectively, where the value of \mathbf{m}' and y^n depends on whether the attack happens or not (i.e, if the attack does not happen, m' = m and y^n is the output of channel W_1 ; otherwise, $m'(\neq m)$ is the forged message that Eve wants to send to Bob, and y^n can be equal or unequal to the output of channel W_1 , which depends on the strategy of the adversary). Upon \mathbf{m}' and y^n , Bob first computes $\mathbf{t}' = \phi_k(\mathbf{m}')$, and then, verifies if

$$a^n \stackrel{\scriptscriptstyle \Delta}{=} \mathbf{t}' \odot [\mathbf{r}_1; \mathbf{r}_2; \cdots; \mathbf{r}_v] \oplus y^n \in \mathcal{C}.$$
(4)

If $a^n \in C$, Bob accepts it; otherwise, he rejects it.

Note that, if H is the parity check matrix of C, then $a^n \in C$ if and only if $H \cdot [a^n]^T = [0^n]^T$, where $[a^n] = [a_1, a_2, \cdots, a_n]$. For details of the proposed scheme for noiseless main channel case with binary erasure wiretapper's channel (NL-BE case), please refer to Alg. 1.

Algorithm 1 Authentication for NL-BE Case

Encoding: If Alice wants to authenticate a message **m**, she

- 1: computes message tag $\mathbf{t} = \phi_k(\mathbf{m}) = [t_1, t_2, \cdots, t_v];$
- 2: selects a vector $\mathbf{s} = [s_1, s_2, \cdots, s_l]$ uniformly at random in $\{0, 1\}^l$;
- 3: computes message authentication code x^n by $x^n = \mathbf{t} \odot$ $[\mathbf{r}_1; \mathbf{r}_2; \cdots; \mathbf{r}_v] \oplus \mathbf{s} \odot [\mathbf{g}_1; \mathbf{g}_2; \cdots; \mathbf{g}_l];$

4: sends
$$(\mathbf{m}, x^n)$$
 to Bob over wiretap channel (W_1, W_2) .

Verifying: After received (\mathbf{m}', y^n) , Bob

1: computes $a^n = [a_1, \cdots, a_n] = \phi_k(\mathbf{m}') \odot [\mathbf{r}_1; \mathbf{r}_2; \cdots; \mathbf{r}_v] \oplus y^n;$

2: calculates
$$H \cdot [a^n]^T$$
, and verifies if $H \cdot [a^n]^T = [0^n]^T$

3: if $H \cdot [a^n]^T = [0^n]^T$, he accepts **m**' and sends the decision bit 1 to Alice; otherwise, he rejects it, and sends 0 to Alice.

B. The Conditions for Perfect Security

In what follows, we give the authentication theorem regarding the conditions for Alg. 1 to be perfect secure.

Theorem 1: Given a wiretap channel (W_1, W_2) with noiseless main channel W_1 and noise wiretapper's channel $W_2 = BEC(\zeta)$, let $P_e^{(n)}(\xi)$ be the probability of block error for code from $C^n(\lambda, \rho)$ over BEC(ξ). If the following conditions hold:

- (1) the family of ϵ -AU₂ hashing functions { $\phi_k : \mathcal{M} \mapsto \mathcal{T}$ } $\mathcal{T}_{k \in \mathcal{K}}$ satisfies the requirements that ϵ and $\frac{|\mathcal{T}|}{|\mathcal{K}|}$ are negligible in n;
- (2) there exists $\eta \in [0, 1]$ such that, for any $\xi < \eta$,

$$P_e^{(n)}(\xi) < \exp^{-\alpha n}$$
 (for some constant $\alpha > 0$),

the ϵ -AU₂ hashing functions and the dual of a code from $C^n(\lambda, \rho)$ used in Alg.1 can achieve perfect secrecy over (W_1, W_2) for $\zeta > 1 - \eta$.

The detailed proof of this theorem will be provided in Appendix B. With this theorem, we only have to construct a family of hashing function satisfying condition (1) and an LDPC code satisfying condition (2) to ensure the security of the proposed authentication scheme. We will discuss how to design the computationally efficient ϵ -AU₂ class of hash functions and an LDPC code to meet these requirements in Sec. VII.

VI. AUTHENTICATION FOR NOISY MAIN CHANNEL CASE

In this section, we study message authentication for noisy main channel case. Firstly, a novel authentication scheme over BEWC is proposed with public discussion. Then, an authentication scheme over binary erasure main channel and binary-input wiretapper's channel is presented by leveraging stochastically degraded channel technique.

A. The Proposed Authentication Scheme Over BEWC

We first consider message authentication over BEWC (W_1, W_2) , where $W_1 = \text{BEC}(\xi_m)$ and $W_2 = \text{BEC}(\xi_w)$. In this case, an intuitive method to achieve perfect secure authentication is to design an LDPC code C with generator matrix G, and then select a set of vectors $\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_v$ in $\{0, 1\}^n / C$ with the following properties: (1) *Security*: the probability of block error of the dual of C over $\text{BEC}(\xi_w)$ decreases exponentially with n; and (2) *Reliability*: $\overline{G} = [G; D]$ is a generator matrix of an LDPC code \overline{C} such that the probability of block error over $\text{BEC}(\xi_m)$ is small enough, where $D = [\mathbf{r}_1; \mathbf{r}_2; \cdots; \mathbf{r}_v]$. However, as mentioned in [19], it is difficult to construct this code.

One promising solution is to authenticate message with public discussion. As shown is Fig. 3, the authentication channel model includes a wiretap channel (BEC(ξ_m),BEC(ξ_w)). Moreover, to simplify the explanation, we assume that there is an insecure and noise-free channel between Alice and Bob, named public channel, which is fully controlled by Eve. Note that, 1) the public channel can be considered as a noise channel with an error-correcting code; and 2) as the information transmitted over public channel has been encoded with some

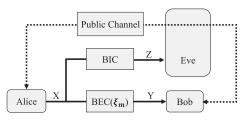


Fig. 3. The authentication channel model for noisy main channel case.

error-correcting code, we cannot guarantee that Eve cannot decode it. Therefore, the assumption that the public channel is fully controlled by Eve is reasonable and even strengthens Eve's capability.

We then present a novel authentication scheme for binary erasure main channel case. The main idea is to reduce the noisy main channel case to a noiseless main channel case through public discussion, which is shown as follows. When Bob receives the output y^n of BEC(ξ_m) with input x^n from Alice, he sends the index set EP_B of the the erased positions (i.e., $EP_B = \{i : y_i = ?\}$) to Alice over noiseless public channel. After obtaining the index set EP_B , Alice transmits $x^n(EP_B) = \{x_i : i \in EP_B\}$ to Bob over the public channel. Accordingly, Bob can obtain x^n when he received $x^n(EP_B)$.

For security, we need to determine which wiretapper's channel is considered in Alg. 1 to generate x^n . Based on the assumption, Eve can observe the output z^n from BEC(ξ_w) and obtain $x^n(EP_B)$ from public channel. We denote $EP_E = \{i : z_i = ?\}$). From the law of large number, if n is large enough, the cardinality of EP_B is less than $n(\xi_m + \frac{1}{2}\sigma)$, and that of EP_E is larger than $n(\xi_w - \frac{1}{2}\sigma)$, where $\sigma = \sigma(n)$, and $\sigma(n) \to 0$ when $n \to \infty$. The worse case for legitimate users is $EP_B \subseteq EP_E$, which means Eve can learn the maximum number of x_i $(i = 1, \dots, n)$ from z^n and $x^n(EP_B)$. Thus, we have

$$|EP_E| - |EP_B| \ge n(\xi_w - \xi_m - \sigma) \tag{5}$$

for large *n*. In the worse case, i.e., $|EP_E| - |EP_B| = n(\xi_w - \xi_m - \sigma)$, it can provide the security if Alg. 1 is used to generate x^n by taking wiretapper's channel as $BEC(\xi_w - \xi_m - \sigma)$, where $0 < \sigma < min\{\xi_w - \xi_m, \frac{1}{2}\xi_w\}$.

It is worth pointing out that $|EP_B| < n(\xi_w + \frac{1}{2}\sigma)$ in the general case. Let $b = \lfloor (\xi_m + \frac{1}{2}\sigma)n - |EP_B| \rfloor$. We have that b > 0in the general case. Based on the discussion above, we improve our method during public discussion stage to further prevent Eve from tampering with the response $x^n(EP_B)$. Specifically, Bob first chooses an index set $RP_B = \{j_1, \dots, j_b\}$ from $[n]/EP_B$ randomly, and sends $EP_B \cup RP_B$ (instead of EP_B) to Alice over the public channel. After receiving $EP_B \cup RP_B$, Alice responses Bob by transmitting $x^n(EP_B \cup RP_B)$ over the public channel. As Bob know $x^n(RP_B)$, he can check if the response $x^n(EP_B \cup RP_B)$ is tampered. Since Eve cannot distinguish RP_B from $EP_B \cup RP_B$, it is difficult to forge a response $\hat{x}^n(EP_B \cup RP_B)$ such that the response can pass Bob's check (i.e., $\hat{x}^n(RP_B) = x^n(RP_B)$).

For details of the proposed authentication scheme for binary erasure main channel case with binary erasure wiretapper's channel (BE-BE case), please refer to Alg. 2.

Algorithm 2 Authentication for BE-BE Case

- 1. Let BEWC (W_1, W_2) be a wiretap channel with main channel $W_1 = \text{BEC}(\xi_m)$ and wiretapper's channel $W_2 = \text{BEC}(\xi_w)$.
- 2. Alice executes Step 1-3 of encoding process in Alg.1 with BEWC(W'_1, W'_2), in which, W'_1 is a noiseless channel, and $W'_2 = \text{BEC}(\xi_w - \xi_m - \sigma)$ for some constant σ satisfying $0 < \sigma < \min\{\xi_w - \xi_m, \frac{1}{2}\xi_w\}$.
- 3. Alice sends **m** with an error-correcting code and x^n to Bob over wiretap channel (W_1, W_2) .
- 4. Let (\mathbf{m}, y^n) be the received information by Bob, and $EP_B = \{i : y_i = ?\} = \{i_1, \dots, i_a\}$ be the erased positions. Bob first computes $b = \lfloor (\xi_m + \frac{1}{2}\sigma)n - |EP_B| \rfloor$. And then, Bob chooses an index set $RP_B = \{j_1, \dots, j_b\}$ from $\{1, \dots, n\}/EP_B$ randomly. Finally, Bob transmits $EP_B \cup RP_B$ to Alice over the public channel.
- 5. Alice checks if $|EP_B \cup RP_B| \le n(\xi_w + \frac{1}{2}\sigma)$. If so, Alice sends $x^n(EP_B \cup RP_B)$ to Bob over the public channel; if not, Alice returns to step 2. Here $x^n(EP_B \cup RP_B)$ is the elements of x^n corresponding to the index set $EP_B \cup RP_B$;
- 6. Bob first checks if $x^n(RP_B) = y^n(PR_B)$. If so, Bob executes the verifying process in Alg.1 with BEWC(W'_1, W'_2); if not, Bob rejects m.

B. The Conditions for Perfect Security

Based the discussion above, we have the following result regarding the conditions for Alg. 2 to be perfect secure.

Theorem 2: Let (W_1, W_2) be a wiretap channel, where $W_1 = \text{BEC}(\xi_m)$ and $W_2 = \text{BEC}(\xi_w)$ with $\xi_m < \xi_w$. If hashing functions and the dual of a LDPC code satisfying Condition (1) and (2) in Theorem 1 with (W'_1, W'_2) , respectively, used in Alg. 2 over (W_1, W_2) , where σ a constant satisfying $0 < \sigma < \min\{\xi_w - \xi_m, \frac{1}{2}\xi_w\}, W'_1$ is a noiseless channel, and $W'_2 = \text{BEC}(\xi_w - \xi_m - \sigma)$. Then, when *n* is large enough, it can provide the perfect security of Alg. 2 for $\zeta > 1 - \eta$.

The detailed proof of this theorem will be given in Appendix C. This theorem shows that perfect secure authentication over the binary erasure wiretap channel can be achieved by leveraging public discussion method. To further illustrate the proposed scheme, we have the following example. Consider a BIWC(W_1, W_2) in which W_1 is BEC(0.1) and $W_2 = \text{BEC}(0.617)$. Taking $\delta = 0.017$, Alice can execute Alg. 1 with BEWC(W'_1, W'_2) to generate a code word, where W'_1 is a noiseless main channel, and $W'_2 = \text{BEC}(0.5)$ is a noise wiretapper's channel. Then, Alice and Bob follow steps 3-6 of Alg. 2 for perfect secure authentication.

To further illustrate the main idea of Alg. 2, we use a simple example to show how it works without considering the security and authentication efficiency. In this example, Hamming (15,4) code with generation matrix G is used, and the linearly independent vectors $[\mathbf{r}_1; \cdots; \mathbf{r}_4]$ is randomly chosen from $\{0, 1\}^n/G$. As shown in Fig. 4, Alice first computes $\mathbf{t} = \phi_k(\mathbf{m})$ (details will be provided

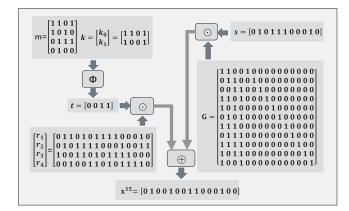


Fig. 4. A simple example of encoding process.

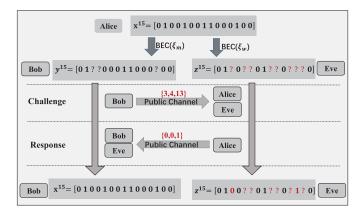


Fig. 5. A simple example of public discussion.

in Section VII-B); and then, computes $x^{15} = \mathbf{t} \odot [\mathbf{r}_1; \cdots; \mathbf{r}_4] \oplus \mathbf{s} \odot G$, where **s** is randomly selected from $\{0, 1\}^{11}$.

As shown in Fig. 5, when Alice transmits x^{15} over $(BEC(\xi_m), BEC(\xi_w))$ and **m** over noiseless channel, Bob and Eve receive (\mathbf{m}', y^n) and (\mathbf{m}, z^n) , respectively, where \mathbf{m}' depends on Eve's attack. Then, Bob checks the erased positions $EP_B = \{3, 4, 13\}$ of y^{15} , and sends it to Alice though the public channel. The next, Alice responses Bob by sending $x^{15}(EP_B) = \{0, 0, 1\}$ to Bob over the public channel. After obtaining $x^{15}(EP_B)$, Bob recovers x^{15} and computes $H \odot [y^{15} \oplus \phi_k(\mathbf{m}')]^T$, where H is the parity check matrix of Hamming (15, 4) code. If the result is $[0^n]^T$, Bob accepts \mathbf{m} ; otherwise, Bob rejects it. At the same time, Eve can recovers the erased bits which are in index set EP_E .

C. Authentication for Binary-Input Wiretapper's Channel

Now we consider message authentication over binary erasure main channel and binary-input wiretapper's channel. First of all, the basic relationship between two memoryless noisy channel is introduced as follows.

Definition 4: A channel $W_1 : \mathcal{X} \to \mathcal{Z}$ is stochastically degraded with respect to channel $W_2 : \mathcal{X} \to \mathcal{Y}$ if there exists a channel $W_3 : \mathcal{Y} \to \mathcal{Z}$ such that $W_1(z|x) = \sum_{y \in \mathcal{Y}} W_2(z|y) W_3(y|x)$ for any $(x, z) \in \mathcal{X} \times \mathcal{Z}$.

From [37, Prop. 6.4], any binary-input channel is stochastically degraded with respect to a binary erasure channel. Algorithm 3 Authentication for BE-BI Case

- 1. Given a BIWC (W_1, W_3) with a binary erasure main channel $W_1 = BEC(\xi_m)$ and a binary-input wiretapper's channel W_3 .
- Create a binary erasure channel W₂ =BEC(ξ_w) by computing ζ with Eq. (6), where W in Eq. (6) is W₃.
 Execute Alg. 2 with BEWC (W₁, W₂).

Specifically, if $W : \{0,1\} \to \mathbb{Z}$ is a binary-input channel, channel W is stochastically degraded with respect to $BEC(\xi_w)$, where

$$\xi_w = \int_z \min_{u \in \{0,1\}} W(z|u) dz.$$
 (6)

The authentication scheme for binary erasure main channel case with binary-input wiretapper's channel (BE-BI case) can be obtained as following algorithm (e.g., Alg. 3).

We generalize the theoretical result of proposed authentication scheme over BEWC (i.e. Theorem 2) to binary-input wiretapper's channel case.

Theorem 3: Let (W_1, W_3) be a wiretap channel, where $W_1 = \text{BEC}(\xi_m)$, and W_3 is binary-input wiretapper's channel. Suppose that W_3 is stochastically degraded with respect to an erasure channel $W_2 = BEC(\xi_w)$. If the hashing functions and the dual of an LDPC code satisfying Condition (1) and (2) in Theorem 2 with wiretap channel (W_1, W_2) , respectively, used in Alg. 3 over (W_1, W_3) , it can provide the perfect security of Alg. 3 for $\zeta > 1 - \eta$.

The detailed proof will be provided in Appendix D. With this theorem, we can achieve perfect secure authentication by using stochastically degraded channel technique even when the wiretapper's channel is BIC.

We illustrate the scheme with the following example. Consider the a BIWC(W_1, W_3) in which $W_1 = \text{BEC}(\xi_m)$ with $\xi_m = 0.1$, and W_3 is binary-input Gaussian channel with input $\{-1, +1\}$ and noise variance 1. By Eq. 6, we can create a binary erasure channel $W_2 = \text{BEC}(\xi_w)$, where

$$\xi_w = \int_z \min_{u \in \{-1,1\}} W(z|u) dz \tag{7}$$

$$= 2 \int_{1}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 0.317$$
 (8)

Thus, Alice and Bob can execute Alg. 2 with $BEWC(W_1, W_2)$ to achieve secure authentication.

VII. IMPLEMENTATION OF THE PROPOSED SCHEMES

In this section, we discuss the implementation of the proposed authentication schemes. The key step is to design a computationally efficient ϵ -AU₂ class of Hash functions and a sequence of large-girth regular LDPC codes to meet the security requirements of the secrecy theorems (i.e., Theorem 1, 2 and 3).

A. Finite Field $GF(2^{\theta})$ Generation

For implementation, it is necessary to propose a ϵ -AU₂ hashing scheme such that the requirements of Theorem 1

TABLE I Some Integer θ With $\Phi(x)$ Irreducible in GF(2)[x]

100	148	180	210	268	292
660	700	786	820	1018	1186

is satisfied with high computing efficiency. The first step of which is to generate a finite field $GF(2^{\theta})$. Let θ be the degree of a polynomial

$$\Psi(x) = x^{\theta} + x^{\theta-1} + \dots + x + 1$$
(9)

with the following properties: (1) $\theta + 1$ is prime; and (2) 2 is a primitive root modulo $\theta + 1$, i.e., $2^{\theta/p} \neq 1 \mod (\theta + 1)$ for any prime p dividing θ . As mentioned in [40], there are many positive integers that satisfy these properties, and some of them is given in Table.1.

From the basic knowledge of finite theory, $\Psi(x)$ is a irreducible in GF(2)[x], and the quotient $GF(2)[x]/(\Psi(x))$ can be used to describe $GF(2^{\theta})$, where GF(2)[x] is the polynomial rings over GF(2), and $(\Psi(x))$ is the idea of GF(2)[x] generated by $\Psi(x)$. In this case, any element $(\alpha_0, ..., \alpha_{s-1})$ in $GF(2^{\theta})$ can be expressed as $\alpha(x) \triangleq \alpha_0 + \alpha_1 x + \cdots + \alpha_{\theta-1} x^{\theta-1}$.

B. Lightweight ϵ -AU₂ Scheme

In [39], a family of hash functions is proposed by den Boer, which is given as follows.

Definition 5: Let q be a prime power, τ be a positive integer, and GF(q) be a finite feild. Set $\mathcal{M} = (GF(q))^{\tau}$, $\mathcal{K} = (GF(q))^2$, and $\mathcal{T} = GF(q)$. For any key $k = (k_0, k_1) \in \mathcal{K}$ such that $k_0, k_1 \in GF(q)$, define $\phi_k : \mathcal{M} \to \mathcal{T}$ as

$$\phi_k(a_1, \cdots, a_\tau) = k_0 + a_1 k_1 + \dots + a_\tau k_1^\tau, \qquad (10)$$

for each $(a_1, \cdots, a_\tau) \in \mathcal{M}$.

From [39], the hashing family $\{\phi_k\}_k$ is ϵ -AU₂ with $|\mathcal{M}| = q^{\tau}$, $|\mathcal{K}| = q^2$, $|\mathcal{T}| = q$ and $\epsilon = \tau/q$. Clearly, by taking $q = 2^{\theta}$ and $\tau = ploy(\theta)$, the $ploy(\theta)/2^{\theta}$ -AU₂ (and also $ploy(\theta)/2^{\theta}$ -AU₂) class of hash functions constructed from Definition 5 satisfies the requirements of Theorem 1.

To make our hashing computation efficiently, we need to consider a lightweight algorithm for multiplication in $GF(2^{\theta})$. Fortunately, Silverman proposed an algorithm for multiplication in $GF(2^{\theta})$ with complexity $\theta + 1$ [40]. For convenience, the operation performed by Silverman's multiplication algorithm is denoted by $Mul(\cdot)$.

Based on the discussion above, we propose a lightweight hashing algorithm as follows. (1) We select a positive integer θ , and generate a finite field $GF(2^{\theta})$ using the method in Sec.VII-A; and select two positive integers τ_0 and τ such that, $\tau = ploy(\theta)$ for a polynomial function $poly(\cdot)$ and τ is divisible by τ_0 (denoting $\pi = \tau/\tau_0$). (2) If (k_0, k_1) be the secure key, we pre-compute τ_0 -vector $[k_1^1, k_1^2, \cdots, k_1^{\tau_0}]$, π -vector $[k_1^{\tau_0}, k_1^{2\tau_0}, \cdots, k_1^{\pi\tau_0}]$, and then save them. (3) For any $m = (a_1, \cdots, a_{\tau})$, we can compute $\phi_k(m)$ by

$$\phi_k(m) = k_0 + \sum_{i=0}^{\pi-1} k_1^{i\tau_0} \sum_{j=1}^{\tau_0} a_{i\tau_0+j} k_1^j.$$
(11)

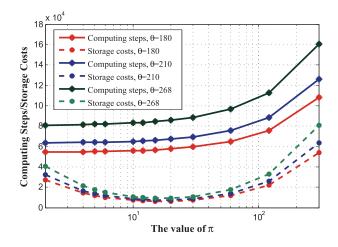


Fig. 6. Computing steps versus storage costs under different values of π .

It is worth pointing out that the value of π is the tradeoff between the time complexity and storage space cost. Actually, the time complexity of Alg. 5 is $(\tau + \pi)\theta$, and the storage cost is $(\tau_0 + \pi)\theta$ bits, where $\pi\tau_0 = \tau$. Fig. 6 shows the computing steps and storage costs for different values of π . For instance, if $\theta = 268$, $\tau = 300$, and $\tau_0 = 15$, the message length is about 80KB, the key length is 536 bits, the tag length is 268 bits, $\varepsilon \approx 2^{-260}$, the number of computing steps is about 8.4×10^4 , and the storage costs are about 9 KB. However, if $\theta = 268$, $\tau = 300$, and $\tau_0 = 50$, the number of computing steps is about 7.3×10^4 , and the storage costs are about 15 KB.

C. Large-Girth Regular LDPC Codes Construction

We show how to find a large-girth regular LDPC code to meet the security requirements of the proposed authentication scheme.

From [34, Sec. 2.5], the design rate of $C(n, d_v, d_c)$ can be expressed by

$$r(n, d_v, d_c) = 1 - \frac{\int_0^1 \rho(x)}{\int_0^1 \lambda(x)} = 1 - \frac{d_c}{d_v}.$$
 (12)

Since any real number r in [0,1] can be approximated infinitely by $1 - \frac{d_c}{d_v}$, where d_c and d_v can be any positive integer. We only need to realize the proposed scheme by leveraging regular LDPC codes to satisfy the requirement of code rate in different wiretapper's channels. Here, we adopt the coding scheme with large-girth regular LDPC codes which can achieve block error probability with double exponential decrease on binary erasure channel [21]. For details of the coding scheme, please refer to [21]. From [21, Th. 9], we have the following result.

Theorem 4: Let $\{C_n\}_n$ be the sequence of large-girth (d_v, d_c) -regular LDPC codes by using [21, Algorithm 4]. If $\varepsilon < \varepsilon_{th}$, we have

$$P_e(\mathcal{C}_n,\varepsilon) \le \mathcal{O}(\exp^{-\beta n^{\alpha \log(d_v-1)}})$$
(13)

for some postive constants α and β .

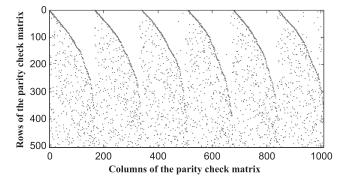


Fig. 7. The sparsity pattern of the parity check matrix for code C(1008, 504).

Proof: Let $P_B^{MP}(\mathcal{C}_n, \varepsilon)$ be the block-error probability under MAP decoding of code \mathcal{C}_n over channel BEC(ε). Then, we have

$$P_e(\mathcal{C}_n,\varepsilon) \le P_B^{MP}(\mathcal{C}_n,\varepsilon) \le nP_b^{MP}(\mathcal{C}_n,\varepsilon)$$
(14)

$$\leq \mathcal{O}(\exp^{-\beta n^{\alpha} \log(a_v - 1)}). \tag{15}$$

VIII. PERFORMANCE EVALUATION

In this section, we evaluate the performance of the proposed schemes by MATLAB R2012a using a desktop computer with a 2.50 GHz Intel CPU, 8GB RAM, and windows 10 OS.

A. Overhead of the Proposed Schemes

We consider the time cost of ϵ -AU₂ hash functions, encoding at Alice, and verifying at Bob. Following the method mentioned in Sec. VII-C, we generate three LDPC codes in ensemble $C(x^3, x^6)$: C(504, 252), C(1008, 504), and C(10080, 5040). Fig.7 shows the sparsity pattern of the parity check matrix for code C(1008, 504). In our experiments, their dual codes will be used in the proposed authentication schemes. It is worth noting that Alg. 2 and Alg. 3 are based on Alg. 1, and the time complexity of the first two algorithms depends crucially on that of the last one. Thus, it is only necessary to consider the overhead of Alg. 1.

Taking $\tau_0 = 50$, Fig. 8 shows the time cost for hashing, encoding and verifying against τ under different values of θ , where the dual code of C(10080, 5040) is leveraged Alg. 1. It can be seen that (1) the time cost for both of them exhibits a near-liner dependence on the value of τ ; (2) The larger the value of θ , the greater the slope of the corresponding line; and (3) The time cost for hashing occupies almost 99.5% of the whole cost for encoding and verifying.

To further decrease the time cost, one possible solution is to use a parallel approach for hash computation. Note that a parallel algorithm using π processors with time complexity $(\tau_0 + 1)\theta + 1$ can be revised from the proposed hash function scheme (in Sec. VII-B) directly. Another possible solution is to use the existing lightweight keying hash functions, such as MD5, SHA1, SHA2 and SHA3. The negative impact of the method is that it would weaken the security of the proposed schemes.

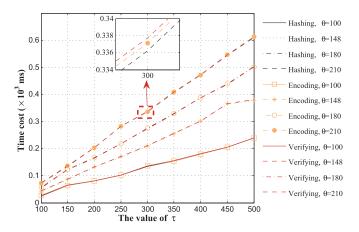


Fig. 8. Time cost for hashing, encoding and verifying.

TABLE II Performance Under Different Parameter Settings

θ	τ	LDPC code	$log \mathcal{M} $	Rate
100	300	C(504, 252)	30 KB	60
180	400	C(504, 252)	72 KB	143
268	500	C(504, 252)	143 KB	266
100	300	C(1008, 504)	30 KB	30
180	400	C(1008, 504)	72 KB	72
268	500	C(1008, 504)	143 KB	133
100	300	C(10080, 5040)	30 KB	3
180	400	C(10080, 5040)	72 KB	7
268	500	C(10080, 5040)	143 KB	13

B. Efficiency of the Proposed Schemes

Table II shows the authentication rate of Alg. 1 under different parameter settings, where $\tau_0 = 50$. Since the density evolution threshold of $C(x^3, x^6)$ is 0.429, i.e., $\varepsilon_{th} = 0.429$. According to Theorem 4, the probability of block error for the sequence of large-girth (3, 6)-regular LDPC codes, which is generated by using [21, Alg. 4], over BEC(ε) can be exponentially small (in terms of n) when $\varepsilon < 0.429$. From Theorem 3, if the wiretapper's channel (i.e., the channel from Alice to Carol) can be stochastically degraded with respect to an erasure channel $W_2 = BEC(\zeta)$, where $\zeta > 1 - \varepsilon_{th} = 0.571$, the proposed authentication scheme is secure with authentication rate in Table II under different parameter settings, where $log|\mathcal{M}|$ is the length of the message.

From Table II, we find that the authentication rate decreases with the length of LDPC code increases. To further improve the authentication rate, it just needs to make minor changes to our schemes. Let the LDPC code be C(n, r). The maximum value of v (in Alg.1) is n - r, i.e., there are at most n - rlinearly independent vectors from $\{0,1\}^n/C(n,r)$. Let \tilde{v} be a positive integer satisfying $\tilde{v} \leq \lfloor \frac{n-r}{\theta} \rfloor$. Denote $v = \tilde{v}\theta$, and let $\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_v$ be the linearly independent vectors from $\{0,1\}^n/C(n,r)$. A new hashing family is defined as

$$\widetilde{\phi}_{\widetilde{\mathbf{k}}} = \phi_{\mathbf{k}_1} \times \dots \times \phi_{\mathbf{k}_{\widetilde{v}}} \tag{16}$$

for any $\widetilde{\mathbf{k}} = (\mathbf{k}_1, \cdots, \mathbf{k}_{\widetilde{v}}) \in \mathcal{K}^{\widetilde{v}}$, where $\mathcal{M} = \{0, 1\}^{\tau\theta}$, $\mathcal{T} = \{0, 1\}^{\theta}$, $\mathcal{K} = \{0, 1\}^{2\theta}$, and $\phi_{\mathbf{k}} : \mathcal{M} \to \mathcal{T}$. For any

TABLE III Performance of the Revised Scheme

θ	τ	v	LDPC code	$log \mathcal{M} $	Rate
100	300	2	C(504,252)	60 KB	119
180	400	1	C(504, 252)	72 KB	143
268	500	1	C(504, 252)	143 KB	266
100	300	5	C(1008, 504)	150 KB	149
180	400	3	C(1008, 504)	216 KB	212
268	500	1	C(1008, 504)	143 KB	133
100	300	30	C(10080, 5040)	0.9GB	90
180	400	28	C(10080, 5040)	2 GB	202
268	500	15	C(10080, 5040)	2.2 GB	213

$$\widetilde{\mathbf{m}} = (\mathbf{m}_1, \cdots, \mathbf{m}_{\widetilde{v}}) \in \mathcal{M}^{\widetilde{v}}$$
, the tag can be calculated **t** by

$$\mathbf{t} = (t_1, \cdots, t_v) = \mathbf{k}(\widetilde{\mathbf{m}}). \tag{17}$$

It is clear that $\{\phi_{\tilde{k}}\}_{\tilde{k}}$ is ϵ -AU₂ if $\{\phi_k\}_k$ is ϵ -AU₂. The revised scheme follows the same steps of encoding process and verifying process of Alg.1 by replacing **m** with $\tilde{\mathbf{m}}$, **k** with $\tilde{\mathbf{k}}$, and $\{\phi_k\}_k$ with $\{\phi_{\tilde{k}}\}_{\tilde{k}}$. The performance of the revised scheme is given in Table III. The result shows that the authentication rate of the revised scheme increases dramatically, especially for large code length scenarios, compared with that of the old one.

IX. CONCLUSION

In this paper, we have proposed efficient and practical multiple messages authentication scheme for wireless communication to provision data integrity and identification. The proposed scheme can achieve perfect security with the same key, by leveraging the lightweight ε -AU₂ hash functions and the dual of the large-girth LDPC codes. Theoretical analysis has verified that the proposed scheme is perfectly secure given that Eve can obtain the information through reviewing a polynomial number of messages authenticated by Alice, and launch a polynomial number of attacks adaptively. Simulation results have also been provided to demonstrate that the proposed schemes can achieve high authentication rate with low time latency. For the future work, we will develop a computationally efficient scheme for Gaussian binary-input wiretap channel model.

APPENDIX

In this Appendix, we will prove the authentication theorems, i.e., Theorem 1, 2 and 3. Before that, we will introduce the result proved by Ozarow and Wyner [18], which connects the equivocation of the eavesdropper to algebraic properties of the generator matrix.

Theorem 5 [18]: Let C(n,l) be an LDPC code with generator matric $G = [g_1, \dots, g_n]$, where g_i represents the *i*-th column of G. Let T is an uniformly distributed RV in $\mathcal{T} = \{0, 1\}^v$. If Z^n is the eavesdropper's observation with μ unerased positions given by $\{i : z_i \neq ?\} = \{i_1, \dots, i_\mu\}$ when the codeword X^n of T is transmitted over $BEC(\xi)$, where X^n is obtained by following Step 2 and 3 of encoding process in Alg. 1. Then, $H(T|Z^n) = H(T)$ iff $G_{\mu}(\triangleq [g_{i_i}, \dots, g_{i_{\mu}}])$ has rank μ .

A. Useful Lemmas

Now we present some lemmas that will be used to prove Theorem 1. Lemma 1 shows that Eve obtains no significant amount of information about secret key K and tag T, after eavesdropping 1 time of authentication that gives Eve information Z^n and M. Let K be uniformly distributed over \mathcal{K} , M be arbitrary message in \mathcal{M} authenticated by Alice, and $T = \phi_K(M)$ be the message tag generated from an ϵ -AU₂ hashing functions with key K and input M. Let G_{μ} be the sub-matrix of G corresponding to the unerased positions in Z^n , and E be a RV in $\{0,1\}$ such that E = 0 if G_{μ} is not full rank; otherwise, E = 1. Then, we have the Lemma as follows.

Lemma 1: If the probability $Pr(E = 0) < exp^{-\alpha n}$ for some $\alpha > 0$, there exists $\beta_1 > 0$ such that, when n is sufficiently large, the following inequalities hold

$$I(T; Z^n, M) \le \exp^{-\beta_1 n} \tag{18}$$

$$I(K; Z^n, M) \le \exp^{-\beta_1 n}.$$
(19)

Proof: Since $T = \Phi_K(M)$, K is an uniformly distributed RV, and $\{\Phi_k\}_k$ is ε -AU₂ hashing functions. We obtain that T is an uniformly distributed RV, and I(T; M) = 0. Then, by the fact that $M \to T \to Z^n$ forms a Markov chain, we can upper bound $I(T; Z^n, M)$ as

$$I(T; Z^{n}, M) = I(T; Z^{n}|M) + I(T; M)$$
(20)

$$= H(Z^{n}|M) - H(Z^{n}|T, M) \quad (\text{as } I(T; M) = 0) \quad (21)$$

$$= H(Z^{n}|M) - H(Z^{n}|T) = I(T;Z^{n}) - I(M;Z^{n})$$
(22)

$$\leq I(T; Z^n) = H(T) - H(T|Z^n).$$
 (23)

Moreover, from Theorem 5, we can lower bound $H(T|Z^n)$ as

$$H(T|Z^n) \ge H(T|Z^n, E) \tag{24}$$

$$\geq H(T|Z^n, E=1) \operatorname{Pr}(E=1) \tag{25}$$

$$= H(T)(1 - \exp^{-\alpha n}).$$
(26)

So, from the above two bounds, we have

$$I(T; Z^n, M) \le H(T)(\exp^{-\alpha n}) \le n \cdot \exp^{-\alpha n}.$$
 (27)

By the fact that RVs K and M are independent, and hence, I(M; K) = 0; and $MK \to T \to Z^n$ forms a Markov chain, we can upper bound $I(K; Z^n, M)$ as

$$I(K; Z^{n}, M) = I(K; Z^{n} | M) + I(K; M)$$
(28)

$$= I(K, M; Z^{n}) - I(M; Z^{n})$$
(29)

$$\leq I(T;Z^n) - I(M;Z^n) \tag{30}$$

$$\leq I(T;Z^n) \leq n \cdot \exp^{-\alpha n}.$$
 (31)

The second Lemma is a generalization of Lemma 1, which shows that Eve also obtains no significant amount of information about key K and tags T_1, \dots, T_J after eavesdropping J times of authentication that gives Eve information $M_1Z_1^n, \dots, M_JZ_J^n$.

Lemma 2: Let M_1, \dots, M_J be J arbitrary messages in \mathcal{M} authenticated by Alice, and $T_i = \phi_K(M_i)$ be the

tags generated from an ϵ -AU₂ hashing functions with key K and input M_j , where $j = 1, \dots, J$. If the probability $\Pr(E = 0) < \exp^{-\alpha n}$ for some $\alpha > 0$, there exists $\beta_1 > 0$ such that, when n is large enough, the following inequalities hold:

$$I(T_j; M_1 Z_1^n \cdots M_J Z_J^n) \le 2^{-\beta_1 n}$$
 (for $j = 1, \cdots, J$); (32)

$$I(K; M_1 Z_1^n \cdots M_J Z_J^n) \le J \cdot 2^{-\beta_1 n}.$$
(33)

Proof: For any i < j, define $M_i^j = M_i \cdots M_j$, and $T_i^j = T_i \cdots T_j$. Given $M^J = m^J$, for any $j \in \{1, \dots, J\}$, as T is determined by (K, m^J) , X_j^n is determined by T_j and the randomness of sampling X_j^n , and Z_j^n is determined by X_j^n and the noise in channel W_2 , the following Markov chain holds:

$$Z_j^n \to T_j \to K \to T_1^{j-1} T_{j+1}^J \to (Z_1^n, \cdots, Z_{j-1}^n Z_{j+1}^n, \cdots, Z_J^n).$$
(34)

Thus, $Z_j^n \to T_j \to (Z_1^n, \dots, Z_{j-1}^n, Z_{j+1}^n, \dots, Z_J^n)$ forms a Markov chain under condition that $M^J = m^J$. Hence, by data processing inequality, we have $I(T_j; Z_1^n, \dots, Z_J^n | m^J) \leq I(T_j, Z_j^n | m^J)$. Averaging over m^J ,

$$I(T_j; Z_1^n, \cdots, Z_J^n | M^J) \le I(T_j, Z_j^n | M^J) = I(T_j, Z_j^n | M_j).$$
(35)

So, by Lemma 1, we have

$$I(T_j; Z_1^n, \cdots, Z_J^n M^J) = I(T_j; Z_1^n, \cdots, Z_J^n | M^J) + I(T_j, M^J)$$
(36)

$$\leq I(T_j M_j; Z_j^n | M_j) + I(T_j, M_j) \tag{37}$$

$$= I(T_{i}; Z_{i}^{n} M_{i}) \le 2^{-\beta_{1} n}$$
(38)

for any j in $\{1, \dots, J\}$.

Further, for any $j \in \{1, \dots, J\}$ and when $M^J = m^J$, $Z_1^n \cdots Z_{j-1}^n \to K \to Z_j^n$ forms a Markov chain as X_j^n is determined by (K, m^J) and the randomness of sampling X_j^n , and Z_j^n is determined by X_j^n and the noise in channel W_2 . Hence,

$$I(K; Z_j^n | Z_1^n \cdots Z_{j-1}^n, M^J = m^J) \le I(K; Z_j^n | M^J = m^J).$$
(39)

Averaging over m^J , we have

$$I(K; Z_j^n | Z_1^n \cdots Z_{j-1}^n M^J) \le I(K; Z_j^n | M^J) = I(K; Z_j^n | M_j).$$
(40)

Therefore, by chain rule of mutual information,

$$I(K; Z_1^n \cdots Z_J^n M^J)$$

= $I(K; M^J) + I(K; Z_1^n \cdots Z_J^n | M^J)$ (41

$$= I(K; Z_1^n \cdots Z_I^n | M^J), \quad (K \text{ is independent of } M^J)$$
(42)

$$\leq \sum_{j} I(K; Z_j^n | M_j) \leq J 2^{-n\beta_1}. \quad (By \text{ Lemma 1})$$
(43)

Lemma 3: If the probability of block error $P_e^{(n)}(\xi)$ for code from $C^n(\lambda, \rho)$ over BEC(ξ) is exponential, i.e., $P_e^{(n)}(\xi) < \exp^{-\alpha n}$ for some constant $\alpha > 0$, when $\zeta < \eta$ for some $\eta \in [0, 1]$. Then, when the dual of a code from $C^n(\lambda, \rho)$ used in the proposed scheme over wiretap channel $(BEC(0), BEC(\zeta))$ with $\zeta > 1 - \eta$, there exists $\beta_1 > 0$ such that the inequalities (32) and (33) hold with sufficiently large n.

Proof: From the important interpretation of $P_e^{(n)}(\xi)$ in [20]: for a parity check matrix H with degree distribution (λ, ρ) , the probability that erased columns of H over a BEC (ξ) from a full-rank submatrix can be lower bounded by $1 - P_e^{(n)}(\xi)$. Thus, if the dual of a code from $C^n(\lambda, \rho)$ used in the proposed scheme over $(BEC(0), BEC(\zeta))$ with $\zeta > 1 - \eta$, we have $Pr(E = 0) < \exp^{-\alpha n}$ for some $\alpha > 0$. By Lemma 2, the lemma follows.

B. Proof of Theorem 1

Proof: When the wiretapper does not present, if Alice wants to authenticate **m**, she first generates x^n by following the proposed scheme, and then sends (\mathbf{m}, x^n) to Bob. At Bob side, he receives (\mathbf{m}, x^n) as the channel between them is noiseless. Clearly, Bob will accept **m** since $a^n = \mathbf{t} \odot [\mathbf{r}_1; \mathbf{r}_2; \cdots; \mathbf{r}_v] \oplus x^n = \mathbf{s} \odot [\mathbf{g}_1; \mathbf{g}_2; \cdots; \mathbf{g}_l] \in \mathcal{C}$. The completeness of the proposed scheme holds. Next, we focus on the authentication property.

Let $M^J = M_1 \cdots M_J$ be the sequence of messages authenticated by Alice, and X_j^n, Z_j^n be the input and output over channel W_2 when Alice authenticates M_j . Let R_E be Eve's random tape.

Because M^J is chosen by Alice according to distribution P_{M^J} , which is independent of R_E . In addition, X_j^n is determined by (K, M_j) together with the randomness of sampling X_j^n (i.e., **RV S** corresponding to the random bits **s** in the proposed scheme), and Z_j^n is determined by X_j^n together with the noise in channel W_2 . Hence, for any $j \in$ $\{1, \dots, J\}, (M^J, K, X_1^n Z_1^n \cdots X_j^n Z_j^n)$ is independent of R_E . By Lemma 2, we have

$$I(K; R_E M^j Z_1^n \cdots Z_j^n) = I(K; M^j Z_1^n \cdots Z_j^n) \le j 2^{-n\beta_1}$$

$$\tag{44}$$

for a constant $\beta_1 > 0$ and any $j \leq J$. Let $V_j = RM^j Z_1^n \cdots Z_j^n$. Lemma 1 in [38] shows the relationship between conditional distance and mutual information, i.e., $SD(X|Y;X) \leq \sqrt{2 \ln 2 I(X;Y)}$ for any RVs X and Y. By [38, Lemma 1], we obtain that

$$\mathsf{SD}(K|V_j;K) \le \sqrt{2j\ln 2} \cdot 2^{-n\beta_1/2}.$$
(45)

According to the adversary model, Eve can adaptively launch the following attacks. (1) Type I attack: when Alice Authenticates M_j and sends out (M_j, X_j^n) , Eve can revise M_j to $M'_j (\neq M_j)$. He succeeds if Bob accepts (M'_j, X_j^n) . (2) Type II attack: at any time, Eve can send a pair (\hat{M}, \hat{X}^n) to Bob over a noiseless channel. He succeeds if Bob accepts (\hat{M}, \hat{X}^n) .

Let b_{ℓ} be the result of the ℓ th attack, where $b_{\ell} = 1$ if Eve succeeds in this attack, and $M^{j_{\ell}-1}$ be the authenticated messages before Eve launches the ℓ th attack. Then, Eve's view before the ℓ th attack can be denoted by $U_{\ell} :=$ $(V_{j_{\ell}-1}, b_1, \cdots, b_{\ell-1}).$

If the ℓ th attack is Type I, $b_{\ell} = 1$ iff $\mathbf{T}'_{j_{\ell}} \odot [\mathbf{r}_1; \mathbf{r}_2; \cdots; \mathbf{r}_v] \oplus X^n_{j_{\ell}} \in \mathcal{C}$, where $X^n_{j_{\ell}} = \mathbf{T}_{j_{\ell}} \odot [\mathbf{r}_1; \mathbf{r}_2; \cdots; \mathbf{r}_v] \oplus \mathbf{S}_{j_{\ell}} \odot [\mathbf{g}_1; \mathbf{g}_2; \cdots; \mathbf{g}_l]$, $\mathbf{T}_{j_{\ell}} = \phi_K(M_{j_{\ell}})$, and $\mathbf{T}'_{j_{\ell}} = \phi_K(M'_{j_{\ell}})$.

Thus, $b_{\ell} = 1$ iff $\mathbf{T}'_{j_{\ell}} = \mathbf{T}_{j_{\ell}}$. In other words, $b_{\ell} = 1$ iff Eve chooses a $M'_{j_{\ell}}$ such that $\phi_K(M'_{j_{\ell}}) = \phi_K(M_{j_{\ell}})$.

If the ℓ th attack is Type II, $b_{\ell} = 1$ iff $\hat{\mathbf{T}}_{j_{\ell}} \odot [\mathbf{r}_1; \mathbf{r}_2; \cdots; \mathbf{r}_v] \oplus \hat{X}_{j_{\ell}}^n \in \mathcal{C}$, where $(\hat{M}_{j_{\ell}}, \hat{X}_{j_{\ell}}^n)$ is Eve's output in this attack, $\mathbf{T}_{j_{\ell}} = \phi_K(M_{j_{\ell}})$, and $\hat{\mathbf{T}}_{j_{\ell}} = \phi_K(\hat{M}_{j_{\ell}})$. $\hat{X}_{j_{\ell}}^n$ can be uniquely rewritten as

$$\hat{X}_{j_{\ell}}^{n} = T_{0} \odot [\mathbf{r}_{1}; \mathbf{r}_{2}; \cdots; \mathbf{r}_{v}] \oplus \mathbf{S}_{0} \odot [\mathbf{g}_{1}; \mathbf{g}_{2}; \cdots; \mathbf{g}_{l}].$$
(46)

So, $b_{\ell} = 1$ iff $\hat{\mathbf{T}}_{j_{\ell}} = \mathbf{T}_0$, i.e., $b_{\ell} = 1$ iff Eve chooses a pair of $\langle \hat{M}_{j_{\ell}}, T_0 \rangle$ such that $\phi_K(\hat{M}_{j_{\ell}}) = \mathbf{T}_0$.

Let L = poly(n) be the upper bound on the number of attacks, where $poly(\cdot)$ is any polynomial function. Then, Eve's success probability can be expressed as $\Pr\left(\bigvee_{\ell=1}^{L} b_{\ell} = 1\right)$. As every successful attacker must experience the first successful attack, without loss of generality, we consider the scenario that an attacker who will stop after the first successful attack. In this case, $b_{\ell} = 1$ implies $b_1 = \cdots = b_{\ell-1} = 0$. Defining $\overline{U}_{\ell} = (V_{j_{\ell}}, b_1, \cdots, b_{\ell-1})$, and $\overline{U}_{\ell}^0 = \{\overline{U}_{\ell} : b_1, \cdots, b_{\ell-1} = 0^{\ell-1}\}$, we have,

$$P(b_{\ell} = 1) = \sum_{u_{\ell} \in \bar{\mathcal{U}}_{\ell}^{0}} P(b_{\ell} = 1, \bar{U}_{\ell} = u_{\ell})$$
(47)

$$= \sum_{u_{\ell} \in \bar{\mathcal{U}}_{\ell}^{0}} P(\bar{U}_{\ell} = u_{\ell}) P(b_{\ell} = 1 | \bar{U}_{\ell} = u_{\ell}).$$
(48)

For given V = v, let $u_{\ell} = v|0^{\ell-1}$ for each $\ell \in \{1, \dots, L\}$. In Type I attack, denoting $\mathcal{E}_{u_{\ell}} = \{k \in \mathcal{K} : \Phi_k(M'_{j_{\ell}}) \neq \Phi_{k_0}(M_{j_{\ell}})\}$, since $M'_{j_{\ell}}, M_{j_{\ell}}$ are deterministic in Eve's view \bar{U}_{ℓ} , the set $\mathcal{E}_{u_{\ell}}$ is completely determined by $\bar{U}_{\ell} = u_{\ell}$. So, by [11, Lemma 5],

$$\Pr(b_{\ell} = 1 | \bar{U}_{\ell} = u_{\ell}) = P_{K | \bar{U}_{\ell} = u_{\ell}} (\mathcal{E}_{u_{\ell}}^{c})$$

$$\leq P_{K} (\mathcal{E}_{u_{\ell}}^{c}) + \frac{1}{2} \mathsf{SD}(P_{K | \bar{U}_{\ell} = u_{\ell}}; P_{K})$$
(49)

$$\leq \epsilon + \frac{1}{2} \mathsf{SD}(P_{K|\bar{U}_{\ell}=u_{\ell}}; P_K).$$
(50)

Averaging over \overline{U}_{ℓ} , we have

$$\Pr(b_{\ell} = 1) \le \epsilon + \frac{1}{2} \mathsf{SD}(P_{K|U_{\ell}}; P_K).$$
(51)

In Type II attack, given $\bar{U}_{\ell} = u_{\ell}$, since Eve's view U_{ℓ} is part of \bar{U}_{ℓ} , it follows that $(\hat{M}_{j_{\ell}}, \hat{X}_{j_{\ell}}^n)$ is deterministic in u_{ℓ} . Thus, $(\hat{M}_{j_{\ell}}, T_0)$ is deterministic in u_{ℓ} . Let

$$\mathcal{E}_{u_{\ell}} = \{ k \in \mathcal{K} : \Phi_k(\hat{M}_{j_{\ell}}) \neq T_0 \}.$$
(52)

Then, by [11, Lemma 5],

$$\Pr(b_{\ell} = 1 | \bar{U}_{\ell} = u_{\ell}) = P_{K | \bar{U}_{\ell} = u_{\ell}} (\mathcal{E}_{u_{\ell}}^{c})$$

$$\leq P_{K} (\mathcal{E}_{u_{\ell}}^{c}) + \frac{1}{2} \operatorname{SD}(P_{K | \bar{U}_{\ell} = u_{\ell}}; P_{K})$$
(53)

$$\leq \frac{|\mathcal{I}|}{|\mathcal{K}|} + \frac{1}{2} \mathsf{SD}(P_{K|\bar{U}_{\ell}=u_{\ell}}; P_K) \text{ (by the def. of } \varepsilon\text{-}AU_2)$$
(54)

Averaging over \overline{U}_{ℓ} , we have

$$\Pr(b_{\ell}=1) \le \frac{|\mathcal{T}|}{|\mathcal{K}|} + \frac{1}{2} \mathsf{SD}(P_{K|\bar{U}_{\ell}}; P_K).$$
(55)

Now we bound $SD(P_{K|\bar{U}_{\ell}}; P_K)$. We first show that (b_1, \dots, b_{ℓ}) is deterministic in (K, V). In Type I attack,

 b_{ℓ} is determined by $(K_0, M'_{j_{\ell}}, M_{j_{\ell}})$, which is further determined by $(K_0, V_{j_{\ell}}, b_1, \dots, b_{\ell-1})$. In Type II attack, b_{ℓ} is determined by $(K, \hat{M}_{\ell}, \hat{X}^n_{j_{\ell}})$, which is further determined by $(K, V_{j_{\ell}}, b_1, \dots, b_{\ell-1})$. Likewise, in both two attacks, $b_{\ell-1}$ is determined by $(K, V_{j_{\ell}}, b_1, \dots, b_{\ell-2})$. So, $(b_{\ell-1}, b_{\ell})$ is determined by $(K, V_{j_{\ell}}, b_1, \dots, b_{\ell-2})$. Following the same discuss as above, we have (b_1, \dots, b_{ℓ}) is deterministic in (K, V). Here, we denote $V_{j_{\ell}}$ by V for simplicity.

Given
$$\bar{U}_{\ell} = u_{\ell} = v0^{\ell-1}$$
, let $\mathcal{K}_{v}^{\ell} \stackrel{ue_{J}}{=} \bigcap_{i=1}^{\ell-1} \mathcal{E}_{u_{i}}$. Since $\bar{U}_{\ell} = (V, b_{1}, \cdots, b_{\ell-1})$. From rule $P_{AB} = P_{A}P_{B|A}$, we obtain that $P_{K\bar{U}_{\ell}}(k, u_{\ell}) = P_{KV}(k, v)P(b_{1}, \cdots, b_{\ell}|k, v) = P_{KV}(k, v)$
(56)

if $(b_1, \dots, b_{\ell-1})$ is determined by (k, v); 0 otherwise. Note that \mathcal{K}_v^{ℓ} is the set of all possible k such that $(b_1, \dots, b_{\ell-1})$ is determined by (k, v). Thus,

$$P_{\bar{U}_{\ell}}(u_{\ell}) = \sum_{k \in \mathcal{K}_v^{\ell}} P_{KV}(k, v) = P_{KV}(\mathcal{K}_v^{\ell}, v).$$
(57)

Hence, from the two equations above, we have the bound of $SD(P_{K|\bar{U}_\ell};P_K)$ as follows.

$$SD(P_{K|\bar{U}_{\ell}}; P_{K})$$

$$= \sum_{u_{\ell}} \sum_{k \in \mathcal{K}} P_{\bar{U}_{\ell}}(u_{\ell}) |P_{KV}(k|u_{\ell}) - P_{K}(k)| \qquad (58)$$

$$= \sum_{v} \sum_{k \in \mathcal{K}_{v}^{\ell}} |P_{KV}(k, v) - P_{KV}(\mathcal{K}_{v}^{\ell}, v)P_{K}(k)| + \sum_{v} \sum_{k \notin \mathcal{K}_{v}^{\ell}} |P_{KV}(\mathcal{K}_{v}^{\ell}, v)P_{K}(k)| \qquad (59)$$

$$\leq \mathsf{SD}(K|V;K) + 2\sum_{v} P_{KV}(\mathcal{K} \setminus \mathcal{K}_{v}^{\ell}, v)$$
(60)

$$\leq 2\mathsf{SD}(K|V;K) + 2\sum_{v} P_{K}(\mathcal{K} \setminus \mathcal{K}_{v}^{\ell}) P_{V}(v)$$
(by [11, Lemma 5] (61)

$$\leq 2\mathsf{SD}(K|V;K) + 2(\ell-1)\epsilon',\tag{62}$$

where $\epsilon' = \max(\epsilon, \frac{|\mathcal{T}|}{|\mathcal{K}|}).$

Since Eve's success probability Pr(Succ(Eve)) can be expressed as $Pr(\bigvee_{\ell=1}^{L} b_{\ell} = 1)$. So,

$$\Pr(\operatorname{Succ}(\operatorname{Eve})) = \Pr\left(\vee_{\ell=1}^{L} b_{\ell} = 1\right) \leq \sum_{\ell} P(b_{\ell} = 1)$$
(63)

$$\leq \sum_{\ell} \left[\epsilon' + \frac{1}{2} \mathsf{SD}(P_{K|\bar{U}_{\ell}}; P_K) \right] \text{ (by Eq. (51), (55))} \quad (64)$$

$$\leq \sum_{\ell} \left[\epsilon' + \mathsf{SD}(K|V;K) + (\ell - 1)\epsilon' \right] \tag{65}$$

$$\leq \sum_{\ell} \left[\mathsf{SD}(K|V;K) + \ell\epsilon' \right] \qquad (\text{where } V = V_{j_{\ell}}) \quad (66)$$

$$\leq \sum_{\ell} \left[\sqrt{2j_{\ell} \ln 2} \cdot 2^{-n\beta_1/2} + \ell \epsilon' \right]$$
 (by Eq. (45)) (67)

$$\leq \sum_{\ell=1}^{L} \left[\sqrt{2L \ln 2} \cdot 2^{-n\beta_1/2} + \ell \epsilon' \right] \tag{68}$$

$$\leq L\sqrt{2L\ln 2} \cdot 2^{-n\beta_1/2} + L^2\epsilon' \tag{69}$$

This is negligible as L is polynomial in n and ϵ' is negligible. This completes our theorem.

C. Proof of Theorem 2

Proof: Bob can view $x^n([n]/EP_B)$ from the output y^n of channel W_1 , and he also can obtain $x^n(EP_B)$ from Alice's response $x^n(EP_B \cup RP_B)$ over the public channel. Thus, Bob can obtain x^n as $x^n = x^n(EP_B) \cup x^n(\{1, \dots, n\}/EP_B)$. The completeness of the presented scheme holds. Now we prove the authentication property.

We denote the authentication game with Alg. 2 as Γ . Consider a new authentication game Γ' that Alice authenticates **m** to Bob with Alg. 1 over wiretap channel (W'_1, W'_2) in the presence of the adversary Oscar. We now claim that, Eve's view in game Γ can be considered as Oscar's view in game Γ' . In fact, by the law of large number, $|EP_B| \leq n(\xi_m + \frac{1}{2}\sigma)$, and $|EP_E| \geq n(\xi_w - \frac{1}{2}\sigma)$ for a large n, where $\sigma = \sigma(n)$, and $\sigma(n) \to 0$ when $n \to \infty$. Eve can observe $x^n([n]/EP_E)$ from the wiretapper's channel W_2 and $x^n(EP_B \cup ER_B)$ from the public channel. Hence, the erasure positions at Eve are $EP_E - (EP_B \cup ER_B)$. By the definition of ER_B , we have $|EP_B \cup ER_B| \leq n(\xi_m + \frac{1}{2}\sigma)$. So, for large n, the number of the erasure positions can be bounded by

$$|EP_E - (EP_B \cup ER_B)| \ge |EP_E| - |EP_B \cup ER_B|$$
$$\ge n(\xi_w - \xi_m - \sigma). \tag{70}$$

Thus, we have $Pr(Succ(Eve)) \leq Pr(Succ(Oscar))$. From Theorem 1, Pr(Succ(Oscar)) is negligible in terms of *n*. Namely, Pr(Succ(Eve)) is negligible in terms of *n*.

D. Proof of Theorem 3

Proof: Assume that there is an adversary Oscar who can observe the output of channel W_2 . From Theorem 2, $\Pr(Succ(Oscar))$ is negligible in terms of n. Since the output of channel W_3 is the degraded version of the output of channel W_2 , we have $\Pr(Succ(Eve)) \leq \Pr(Succ(Oscar))$. Thus, $\Pr(Succ(Eve))$ is negligible. The rigorous mathematical proof can be obtained by combining the data-processing inequality and the proof of Theorem 2. We omit it due to the space limitation.

REFERENCES

- J. G. Andrews *et al.*, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.
- [2] N. Zhang *et al.*, "Software defined networking enabled wireless network virtualization: Challenges and solutions," *IEEE Netw.*, vol. 31, no. 5, pp. 42–49, May 2017.
- [3] K. Zhang, K. Yang, X. Liang, Z. Su, X. Shen, and H. H. Luo, "SecuriSecurity and privacy for mobile healthcare networks: From a quality of protection perspective," *IEEE Wireless Commun.*, vol. 22, no. 4, pp. 104–112, Aug. 2015.
- [4] G. J. Simmons, "Authentication theory/coding theory," in Advances in Cryptology-CRYPTO. New York, NY, USA: Springer-Verlag, 1985, pp. 411–431.
- [5] U. M. Maurer, "Authentication theory and hypothesis testing," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1350–1356, Jul. 2000.
- [6] A. D. Wyner, "The wire-tap channel," Bell Syst. Tech. J., vol. 54, pp. 1355–1387, Oct. 1975.
- [7] I. Csiszár and J. Köner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 3, pp. 339–348, May 1978.

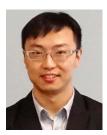
- [8] D. Chen *et al.*, "S2M: A lightweight acoustic fingerprints-based wireless device authentication protocol," *IEEE Internet Things J.*, vol. 4, no. 1, pp. 88–100, Feb. 2017.
- [9] Y. Zhu, L. Wang, K.-K. Wong, and R. W. Heath, "Physical layer security in large-scale millimeter wave ad hoc networks," in *Proc. IEEE Globcom*, Washington, DC, USA, Dec. 2016, pp. 4–8.
- [10] L. Lai, H. El Gamal, and H. V. Poor, "Authentication over noisy channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 906–916, Feb. 2009.
- [11] D. Chen, S. Jiang, and Z. Qin, "Message authentication code over a wiretap channel," in *Proc. ISIT*, 2015, pp. 2301–2305.
- [12] J. Bao, X. Gao, X. G. Xu, R. Guo, and B. Jiang, "Low rate QC LDPC codes with reconfigurable structures for space information networks," *J. Commun.*, vol. 11, no. 3, pp. 255–262, Mar. 2016.
- [13] R. G. Maunder, "A vision for 5G channel coding," Acceler-Comm, Hampshire, U.K., White Paper, 2016. [Online]. Available: https://www.accelercomm.com/download/pdf/5G-Channel-Coding.pdf
- [14] U. Maurer and S. Wolf, "Secret-key agreement over unauthenticated public channels—Part I: Definitions and a completeness result," *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 822–831, Apr. 2003.
- [15] R. Ahlswede and I. Csiszar, "Common randomness in information theory and cryptography. II. CR capacity," *IEEE Trans. Inf. Theory*, vol. 44, no. 1, pp. 225–240, Jan. 1998.
- [16] N. Zhang, N. Cheng, N. Lu, X. Zhang, J. W. Mark, and X. Shen, "Partner selection and incentive mechanism for physical layer security," *IEEE Trans. Wireless Commun.*, vol. 14, no. 8, pp. 4265–4276, Aug. 2015.
- [17] X. Fang, N. Zhang, S. Zhang, D. Chen, X. Sha, and X. Shen, "On physical layer security: Weighted fractional Fourier transform based user cooperation," *IEEE Trans. Wireless Commun.*, vol. 16, no. 8, pp. 5498–5510, Aug. 2017.
- [18] L. H. Ozarow and A. D. Wyner, "Wire-tap channel II," ATT Bell Lab. Tech. J., vol. 63, no. 10, pp. 2135–2157, 1984.
- [19] A. Thangaraj, S. Dihidar, A. R. Calderbank, S. W. McLaughlin, and J.-M. Merolla, "Applications of LDPC codes to the wiretap channel," *IEEE Trans. Inf. Theory*, vol. 53, no. 8, pp. 2933–2945, Aug. 2007.
- [20] A. T. Suresh, A. Subramanian, A. Thangaraj, M. Bloch, and S. W. McLaughlin, "Strong secrecy for erasure wiretap channels," in *Proc. IEEE Inf. Theory Workshop*, Dublin, Ireland, Sep. 2010, pp. 1–5.
- [21] A. Subramanian, A. Thangaraj, M. Bloch, and S. W. McLaughlin, "Strong secrecy on the binary erasure wiretap channel using largegirth LDPC codes," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 585–594, Sep. 2011.
- [22] J. Jin, C. Xiao, M. Tao, and W. Chen, "Linear precoding for fading cognitive multiple-access wiretap channel with finite-alphabet inputs," *IEEE Trans. Veh. Technol.*, vol. 66, no. 4, pp. 3059–3070, Apr. 2017.
- [23] Y. Wu et al., "Low-complexity MIMO precoding for finite-alphabet signals," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4571–4584, May 2017.
- [24] E. Hof and S. Shamai (Shitz), "Secrecy-achieving polar-coding," in Proc. IEEE Inf. Theory Workshop, Dublin, Ireland, Sep. 2010, pp. 1–5.
- [25] H. Mahdavifar and A. Vardy, "Achieving the secrecy capacity of wiretap channels using polar codes," *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6428–6443, Oct. 2011.
- [26] V. Korzhik, V. Yakovlev, G. Morales-Luna, and R. Chesnokov, "Performance evaluation of keyless authentication based on noisy channel," in *Proc. MMM-ACNS*, Heidelberg, Germany, 2007, pp. 115–126.
- [27] P. Baracca, N. Laurenti, and S. Tomasin, "Physical layer authentication over MIMO fading wiretap channels," *IEEE Trans. Wireless Commun.*, vol. 11, no. 7, pp. 2564–2573, Jul. 2012.
- [28] A. Ferrante, N. Laurenti, C. Masiero, M. Pavon, and S. Tomasin. (2013). "On the achievable error region of physical layer authentication techniques over Rayleigh fading channels." [Online]. Available: https://arxiv.org/abs/1303.0707
- [29] S. Jiang, "Keyless authentication in a noisy model," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 6, pp. 1024–1033, Apr. 2014.
- [30] S. Jiang, "On the Optimality of Keyless Authentication in a Noisy Model," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 6, pp. 1250–1261, Jun. 2015.
- [31] D. Chen, N. Cheng, N. Zhang, K. Zhang, Z. Qin, and X. Shen, "Multimessage authentication over noisy channel with polar codes," in *Proc. IEEE MASS*, Orlando, FL, USA, Oct. 2017, pp. 46–54.
- [32] S. Fang, Y. Liu, and P. Ning, "Mimicry attacks against wireless link signature and new defense using time-synched link signature," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 7, pp. 1515–1527, Jul. 2016.
- [33] W. Wang, Z. Sun, S. Piao, B. Zhu, and K. Ren, "Wireless physicallayer identification: Modeling and validation," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 9, pp. 2091–2106, Sep. 2016.

- [34] T. Richardson and R. Urbanke, *Modern Coding Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2008.
- [35] D. Chen, Z. Qin, X. Mao, P. Yang, Z. Qin, and R. Wang, "SmokeGrenade: An efficient key generation protocol with artificial interference," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 11, pp. 1731–1745, Nov. 2013.
- [36] Q. Wang, K. Xu, and K. Ren, "Cooperative secret key generation from phase estimation in narrowband fading channels," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 9, pp. 1666–1674, Oct. 2012.
- [37] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge, U.K.: Cambridge Univ. Press, 2011.
- [38] I. Csiszár, "Almost independence and secrecy capacity," Problems Inf. Transmiss., vol. 32, no. 1, pp. 40–47, 1996.
- [39] B. den Boer, "A simple and key-economical unconditional authentication scheme," J. Comput. Security, vol. 2, no. 1, pp. 65–71, 1993.
- [40] J. H. Silverman, "Fast multiplication in finite field $GF(2^n)$," in *Proc. CHES*, 1999, pp. 122–134.
- [41] R. Safavi-Naini and P. R. Wild, "Information theoretic bounds on authentication systems in query model," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2426–2436, Jun. 2008.



Dajiang Chen (M'15) received the B.Sc. degree from Neijiang Normal University in 2005 and the M.Sc. degree Sichuan University in 2009, and the Ph.D. degree in information and communication engineering from the University of Electronic Science and Technology of China (UESTC) in 2014. He is currently an Assistant Professor with the School of Information and Software Engineering, UESTC. He was a Post-Doctoral Fellow with the University of Waterloo, Waterloo, ON, Canada, from 2015 to 2017, and the School of Information and

Software Engineering, UESTC, from 2014 to 2017. His current research interest includes information theory, secure channel coding, and their applications in wireless network security, wireless communications and other related areas. He serves/served as a Technical Program Committee Member for the IEEE Globecom, the IEEE VTC, the IEEE WPMC, and the IEEE WF-5G.



Ning Zhang (M'15) received the B.Sc. degree from Beijing Jiaotong University, Beijing, China, in 2007, the M.Sc. degree from the Beijing University of Posts and Telecommunications, Beijing, in 2010, and the Ph.D. degree from the University of Waterloo, Waterloo, ON, Canada, in 2015. He was a Post-Doctoral Fellow with the University of Waterloo and University of Toronto, respectively. He is currently an Assistant Professor at Texas A&M University at Corpus Christi, Corpus Christi, TX, USA. His current research interests include physical

layer security, dynamic spectrum access, 5G, and vehicular networks.



Rongxing Lu (S'09–M'11–SM'15) received the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Waterloo, Canada, in 2012. He was an Assistant Professor with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, from 2013 to 2016. He has been an Assistant Professor with the Faculty of Computer Science (FCS), University of New Brunswick (UNB), Canada, since 2016. He was a Post-Doctoral Fellow with the University of Waterloo from 2012 to 2013. He received

the most prestigious Governor General's Gold Medal, when he received his Ph.D. degree, the 8th IEEE Communications Society (ComSoc) Asia Pacific Outstanding Young Researcher Award in 2013. He is currently a Senior Member of the IEEE Communications Society. His research interests include applied cryptography, privacy enhancing technologies, and IoT-big data security and privacy. He has published extensively in his areas of expertise (with citation over 11,100 and H-index 51 from the Google Scholar in 2018). He was a recipient of eight best (student) paper awards from some reputable journals and conferences. He currently serves as the Vice-Chair (Publication) of the IEEE ComSoc CIS-TC (Communications and Information Security Technical Committee). He is the Winner of 2016–2017 Excellence in Teaching Award, FCS, UNB.





Xiaojie Fang (S'14) received the B.Sc. and M.Sc. degrees from the Department of Electronics and Information Engineering, Harbin Institute of Technology, in 2010 and 2012, respectively, where he is currently pursuing the Ph.D. degree. He was a Visiting Scholar with the Broadband Communications Research Group, Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada. His current research interests include physical layer security and coding and modulation theory.



Professor with the Department of Electrical and Computer Engineering, University of Nebraska-Lincoln, USA, since 2017. His research interests include security and privacy for mobile social networks, e-healthcare systems, cloud/edge computing, and cyber-physical systems.



Zhiguang Qin (S'95–A'96–M'14) was the Dean of the School of Software of University of Electronic Science and Technology of China (UESTC). He is the Director of the Key Laboratory of New Computer Application Technology and the Director of UESTC-IBM Technology Center. His research interests include wireless sensor networks, mobile social networks, information security, applied cryptography, information management, intelligent traffic, electronic commerce, and distribution and middleware. He served as the General Co-Chair for WASA 2011, Bigcom 2017.



Xuemin (Sherman) Shen (M'97–SM'02–F'09) received the B.Sc. degree in electrical engineering from Dalian Maritime University, China, 1982, and the M.Sc. and Ph.D. degrees from Rutgers University, Brunswick, NJ, USA, 1987 and 1990, respectively. He is a University Professor and the Associate Chair for Graduate Studies, Department of Electrical and Computer Engineering, University of Waterloo, Canada. His research interests include wireless resource management, wireless network security, social networks, smart grid, and vehicular

ad hoc and sensor networks. He is the elected IEEE ComSoc VP Publication, was a member of the IEEE ComSoc Board of Governor, and the Chair of Distinguished Lecturers Selection Committee. He is a registered Professional Engineer of Ontario, Canada, an Engineering Institute of Canada Fellow, a Canadian Academy of Engineering Fellow, a Royal Society of Canada Fellow, and a Distinguished Lecturer of the IEEE Vehicular Technology Society and Communications Society. He received the IEEE ComSoc Education Award, the Joseph LoCicero Award for Exemplary Service to Publications, the Excellent Graduate Supervision Award in 2006, and the Premiers Research Excellence Award from the Province of Ontario, Canada, in 2003. He served as the Technical Program Committee Chair/Co-Chair for the IEEE Globecom16, Infocom14, IEEE VTC10 Fall, and Globecom07, the Symposia Chair for IEEE ICC10, the Tutorial Chair for IEEE VTC'11 Spring and IEEE ICC08, the General Co-Chair for ACM Mobihoc15, Chinacom07 and QShine06, the Chair for the IEEE Communications Society Technical Committee on Wireless Communications, and P2P Communications and Networking. He serves/served as the Editor-in-Chief for the IEEE INTERNET OF THINGS JOURNAL, the IEEE NETWORK, the Peer-to-Peer Networking and Application, and the IET Communications, a Founding Area Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and an Associate Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY and the **IEEE WIRELESS COMMUNICATIONS.**