Enhancing Physical Layer Security in Internet of Things via Feedback: A General Framework

Bin Dai, *Member, IEEE*, Zheng Ma, *Member, IEEE*, Yuan Luo, Xuxun Liu, *Member, IEEE*, Zhuojun Zhuang, and Ming Xiao, *Senior Member, IEEE*

***Abstract*—In this paper, a general framework for enhancing the physical layer security (PLS) in Internet of Things (IoT) systems via channel feedback is established. To be specific, first, we study the compound wiretap channel with feedback, which can be viewed as an ideal model for enhancing the PLS in the down-link transmission of IoT systems via feedback. A novel feedback strategy is proposed and a corresponding lower bound on the secrecy capacity is constructed for this ideal model. Next, we generalize the ideal model (i.e., the compound wiretap channel with feedback) by considering channel states and feedback delay, and this generalized model is called the finite state compound wiretap channel with delayed feedback. Lower bounds on the secrecy capacities of this generalized model with or without delayed channel output feedback are provided, and they are constructed according to variations of the previously proposed feedback scheme for the ideal model. Finally, from a Gaussian fading example, we show that the delayed channel output feedback enhances the achievable secrecy rate of the finite state compound wiretap channel with only delayed state feedback, which implies that feedback helps to enhance the PLS in the down-link transmission of IoT systems.**

***Index Terms*—Compound channel, feedback, secrecy capacity, wiretap channel.**

Manuscript received June 17, 2019; revised September 6, 2019; accepted September 19, 2019. The work of B. Dai was supported by the National Natural Science Foundation of China under Grant 61671391, the China Scholarship Council (file No. 201807005013), and the 111 Project No.111-2-14. The work of Z. Ma was supported by the National Natural Science Foundation of China under Grant U1734209, the Key International Cooperation Project of Sichuan Province under Grant 2017HH0002, the EU Marie Sklodowska-Curie individual Fellowship under Grant 796426, and the NSFC China-Swedish project under Grant 6161101297. The work of Y. Luo was supported by the National Natural Science Foundation of China under Grant 61871264. The work of M. Xiao was supported by Swedish Strategic Research Foundation project "High-reliable Low-latency Industrial Wireless Communications", the EU Marie Sklodowska-Curie Actions Project entitled "High-reliability Low-latency Communications with network coding", and ERA-NET Smart Energy Systems SG+ 2017 Program, "SMART-MLA" with Project number 89029 (and SWEA number 42811-2).

B. Dai is with the School of Information Science and Technology, Southwest JiaoTong University, Chengdu 611756, China, e-mail: daibin@home.swjtu.edu.cn.

Z. Ma is with the School of Electrical Engineering and Computer Science, Royal Institute of Technology (KTH), SE-10044, Stockholm, Sweden, e-mail: zma@home.swjtu.edu.cn.

Y. Luo is with the computer science and engineering department, Shanghai Jiao Tong University, Shanghai 200240, China, email: luoyuan@cs.sjtu.edu.cn.

X. Liu is with the School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510641, China, email: liuxuxun@scut.edu.cn.

Z. Zhuang is with the Innovation Academy for Microsatellites of CAS, Shanghai 201203, China, e-mail: zhuangzhuojun@163.com.

M. Xiao is with the School of Electrical Engineering and the ACCESS Linnaeus Center, Royal Institute of Technology, Sweden, e-mail: mingx@kth.se.

I. Introduction

INTERNET of Things (IoT) is taking the centre-stage of the upcoming 5G as the devices are expected to be a major component of 5G network. Due to the broadcasting nature of wireless communication, signals in the IoT systems are more vulnerable to eavesdropping, and hence the secure communication over the IoT systems is one of the most pressing problems needed to be solved. The study of the secure transmission over communication systems started from Wyner in his groundbreaking work on the wiretap channel (WTC) [1], where a transmitter broadcasts its message W over N channel uses to a legitimate receiver and an eavesdropper via a degraded broadcast channel (BC), and the perfect secrecy is guaranteed if the information leakage rate $\frac{1}{N}I(W;Z^N)$, where Z^N denotes the received signal at the eavesdropper, vanishes as the codeword length N tends to infinity. [1] The secrecy capacity, defined as the channel capacity with perfect secrecy constraint, was established in [1]. Subsequently, [3] generalized the model studied in [1] by considering a general (not degraded) BC and the transmission of a common message which can be decoded by both the legitimate receiver and the eavesdropper. The follow-up studies of [1]-[3] include the Gaussian wiretap channel [4], the BC with two secret messages [5]-[6], and one transmitter broadcasts a secret message to multiple legitimate receivers and one eavesdropper (wiretap broadcast channel) [7]-[8].

Here note that in [1], Wyner further pointed out that the secrecy capacity is positive if the legitimate receiver's channel is less noisy than the eavesdropper's. Then it is natural to ask the following two questions:

- 1) How to achieve positive secrecy capacity when the eavesdropper's channel is less noisy than the legitimate receiver's?
- 2) When the legitimate receiver's channel is less noisy than the eavesdropper's, how to further enhance the secrecy capacity?

The answer to both questions is artificial noise (AN) [9]-[16] and channel feedback (CF). However, we should notice that since the IoT devices (e.g., sensors and actuators) have significant energy constraint [17]-[18], *AN may not be suitable for IoT systems*, and hence CF is of particular interest for enhancing the physical layer security (PLS) in IoT systems.

[1]Here the perfect secrecy defined in [1] is in fact weak perfect secrecy. Another definition of the perfect secrecy is strong perfect secrecy [2], which is defined as the information leakage $I(W;Z^N)$ at the eavesdropper vanishes as N tends to infinity.

The study of the effect of CF on the PLS of communication channels started from [19], where the pioneering work [1] has been re-visited by considering the situation that the legitimate receiver's received channel output is sent back to the transmitter through an additional noiseless feedback channel which is not known by the eavesdropper. Since the legitimate receiver's channel output is perfectly known by the transmitter and completely not known by the eavesdropper, we can generate secret key from it, and this key helps to protect the transmitted message. Combining the above idea of generating secret key from the CF with the random binning scheme for the WTC [1], [19] proposed a coding scheme which splits the transmitted message into two parts, where one sub-message is encoded in the same way as that of the WTC [1], and the other is encrypted by the secret key. Compared with the secrecy capacity of the WTC [1], it is easy to see that encrypting part of the message by the key leads to the fact that the channel output feedback enhances the secrecy capacity of the WTC.

Here note that in [19], the feedback channel is only used to send the legitimate receiver's channel output, and what happens when the feedback channel can transmit anything as the legal receiver wishes? [20] studied this case and pointed out that for the legal receiver, the best way is to transmit randomly generated sequences (served as secret keys) over the feedback channel. Assume that the transmission rate of the feedback channel is up to R_f, using a coding scheme in the same way as that in [19], [20] showed that sending pure secret key is better than sending legal receiver's channel output if R_f is larger than the key rate in [19], and vice versa. The work of [19] and [20] indicates that there is no difference between sending pure secret key and sending legitimate receiver's channel output, and the main purpose of the feedback is to allow the legitimate receiver and the transmitter to share the secret key. In recent years, the above secret key based feedback coding scheme has been widely used in communication systems with feedback channel. To be specific, for the communication channels with legal receiver's channel output feedback, [23]-[24] studied PLS of the channels with memory or memoryless states and legitimate receiver's channel output feedback, and proposed variations of the secret key based feedback coding scheme in [19]. For the communication systems with feedback channels directly transmitting pure secret keys, [21] extended the work of [20] to a broadcast situation, where two legitimate receivers of the broadcast channel independently send their secret keys to the transmitter via two noiseless feedback channels, and these keys help to increase the achievable secrecy rate region of the broadcast wiretap channel [7]. [22] introduced memoryless channel state into the work of [20], and showed that the transmitted message can be protected by two keys, where one is from the feedback channel, and the other is generated by the channel state. Very recently, [25] showed that for the general WTC with CF, a better choice of the transmitter is to produce not only secret key but also auxiliary message from CF, where the auxiliary message is used to improve the legal receiver's decoding performance. [25] proved that for the wiretap channel with channel output feedback, this new feedback scheme performs better than the widely used secret key based feedback scheme. Moreover, [26] and [27] showed that the classical Schalkwijk-Kailath (SK) feedback scheme for the Gaussian channel [28] achieves the secrecy capacity of the Gaussian wiretap channel with channel output feedback, and it equals the capacity of the same channel model without the secrecy constraint. However, we should notice that the results of [26] and [27] only work in the Gaussian case.

In IoT systems, the up-link transmission is from sensors to controllers, and the down-link transmission is from controllers to actuators. A classical scenario for the PLS in the down-link transmission of the IoT systems is depicted in Figure 1, where a controller tries to send a secret message to several actuators in the presence of an eavesdropper. An ideal model characterizing this classical scenario is called the compound wiretap channel, where the channels for all actuators and the eavesdropper are independent of one another. Achievable secrecy rates (lower bounds on the secrecy capacities) of various compound wiretap channels were provided in [29]-[31], and it is shown that if the eavesdropper's channel is less noisy than all actuators' (legitimate receivers') channels, the achievable secrecy rate equals zero. As we mentioned before, AN is not suitable for the scenario in Figure 1. Moreover, we should notice that *the already existing feedback schemes in [19]-[25] can not be applied to the communication scenario in Figure 1 due to the reason that each actuator does not know others' CF, and hence the feedback of all channels cannot be used to generate a common secret key shared between the transmitter and all the actuators.*

In this paper, we try to answer the following two fundamental questions:

- 1) How to increase the achievable secrecy rate of the compound wiretap channel model by using CF?
- 2) Is there a more practical model for the scenario shown in Figure 1? If so, can we apply the feedback scheme of 1) to this more practical model?

This paper provides the comprehensive answers to the aforementioned questions. Our main contributions are summarized as follows:

- 1) We study the compound wiretap channel with feedback, see Figure 2. An achievable secrecy rate, which is constructed according to a novel feedback strategy, is provided for the model of Figure 2.
- 2) A more practical model for the scenario in Figure 1 is provided (see Figure 3), and an achievable secrecy rate for this new model is obtained according to a modified feedback scheme of the model of Figure 2. From a Gaussian fading example, we show that the achievable secrecy rate of the new model is larger than that of the same model without channel output feedback, which implies that the proposed feedback scheme enhances the PLS in the down-link transmission of the IoT systems.

In the remainder of this paper, random variable (RV), value and alphabet are denoted by uppercase letter, lowercase letter and calligraphic letter, respectively. The random vector and its value are written in a similar way. For example, suppose that X_1 is a RV, and x_1 is a real value in the alphabet $\mathcal{X}_1$. Similarly, Suppose that $X_{1,1}^N$ is a random vector $(X_{1,1}, ..., X_{1,N})$, and

$x_{1,1}^N = (x_{1,1}, ..., x_{1,N})$ is a vector value in $\mathcal{X}_1^N$ (the N-th Cartesian power of $\mathcal{X}_1$). Moreover, for simplicity, the probability $Pr\{X = x\}$ is denoted by $P(x)$, and in the remainder of this paper, the base of the log function is 2.

The outline of this paper is organized as follows. Section II investigates the compound wiretap channel with feedback (see Figure 2), and provides bounds on the secrecy capacity of this ideal model. Section III studies a generalized version of the model of Figure 2, called the finite state compound wiretap channel with delayed feedback (see Figure 3), and also provides bounds on the secrecy capacity of this generalized model. In Section IV, the capacity results on the model of Figure 3 are further illustrated by a Gaussian fading example. Final conclusion is given in Section V.

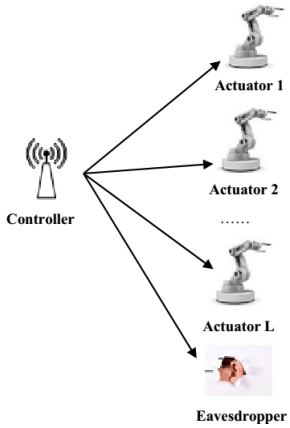

Fig. 1: The PLS in the down-link transmission of the IoT systems

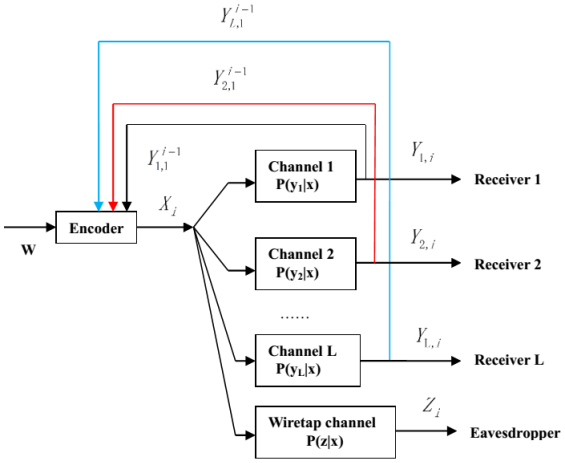

Fig. 2: The compound wiretap channel with feedback

II. The Compound Wiretap Channel with Feedback

The discrete memoryless compound wiretap channel with feedback is shown in Figure 2, where a transmitter wishes to broadcast his/her secret message to L legitimate receivers and an eavesdropper attempts to eavesdrop this secret message via

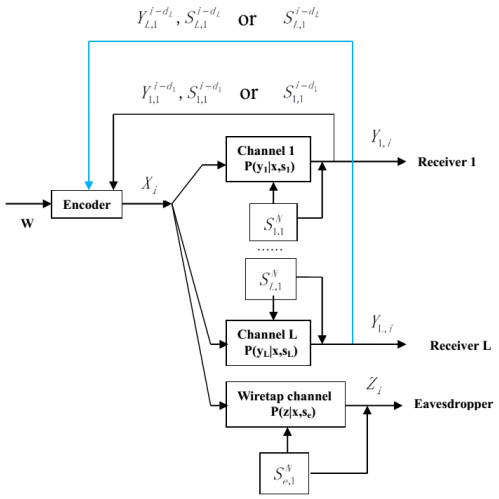

Fig. 3: The finite state compound wiretap channel with delayed feedback

a wiretap channel. The overall channel transition probability of the model of Figure 2 is given by

$$\begin{aligned} &P(y_{1,1}^N, y_{2,1}^N, ..., y_{L,1}^N, z^N | x^N) \\ &= P(z^N|x^N) \prod_{j=1}^{L} P(y_{j,1}^N | x^N) \\ &= \prod_{i=1}^{N} \left(P(z_i|x_i) \prod_{j=1}^{L} P(y_{j,i}|x_i) \right), \end{aligned} \tag{1}$$

where $x_i \in \mathcal{X}$, $y_{j,i} \in \mathcal{Y}_j$ and $z_i \in \mathcal{Z}$. Here note that z^N (x^N) is an abbreviation of z_1^N (x_1^N), and a similar convention is applied to Z_1^N (X_1^N).

The transmitted message W is uniformly distributed over $\mathcal{W} = \{1, 2, ..., |\mathcal{W}|\}$. Since all legitimate receivers send their received channel outputs back to the transmitter via feedback channels, the i-th ($i \in \{1, 2, ..., N\}$) channel input X_i is given by

$$X_i = f_i(W, Y_{1,1}^{i-1}, Y_{2,1}^{i-1}, ..., Y_{L,1}^{i-1}), \tag{2}$$

where f_i is a stochastic encoding function, and $Y_{j,1}^{i-1}$ ($j \in \{1, 2, ..., L\}$) is the j-th legitimate receiver's channel output feedback at time i.

For the j-th ($j \in \{1, 2, ..., L\}$) legitimate receiver, after receiving $Y_{j,1}^N$, he/she produces an estimation $\hat{W}^{(j)} = \psi_j(Y_{j,1}^N)$ (ψ_j is the j-th legitimate receiver's decoding function), and the average decoding error probability equals

$$P_{e,j} = \frac{1}{|\mathcal{W}|} \sum_{i \in \mathcal{W}} Pr\{\psi_j(y_{j,1}^N) \neq i | i \text{ sent}\}. \tag{3}$$

The secrecy level of the transmitted message W at the eavesdropper is formulated as

$$\Delta = \frac{1}{N} H(W|Z^N). \tag{4}$$

Given a non-negative number R, if for any $\epsilon > 0$, there exist

encoder and decoders such that

$$\frac{\log|\mathcal{W}|}{N} \geq R-\epsilon,\ \Delta \geq R-\epsilon,$$
$$P_{e,j} \leq \epsilon \text{ for all } j \in \{1,2,...,L\}, \quad (5)$$

R is achievable under weak perfect secrecy constraint. The secrecy capacity $\mathcal{C}_s^f$ is composed of all such achievable R defined in (5), and bounds on $\mathcal{C}_s^f$ are given in the remainder of this section.

Theorem 1: **Lower bound on $\mathcal{C}_s^f$**: $\mathcal{C}_s^f \geq R_s^f$, where

$$R_s^f = \max \min_j \min\{[I(V_j;Y_j,U_j) - I(V_j;Z)]^+, I(V_j;Y_j)\}, \quad (6)$$

$[a]^+ = a$ for $a \geq 0$, $[a]^+ = 0$ for $a < 0$, and the joint probability is defined as

$$\begin{aligned} &P(z,y_1,...,y_L,x,v_1,...,v_L,u_1,...,u_L) \\ &= P(z|x)P(u_1,...,u_L|v_1,...,v_L,y_1,...,y_L)\cdot \\ &P(y_1,...,y_L|x)P(x|v_1,...,v_L)P(v_1,...,v_L) \\ &= P(z|x)P(x|v_1,...,v_L)P(v_1,...,v_L)\cdot \\ &\prod_{j=1}^{L} P(u_j|v_j,y_j)P(y_j|x). \end{aligned} \quad (7)$$

Proof sketch:

The lower bound R_s^f is constructed according to a block Markov coding scheme, and the encoding-decoding procedure of each block is briefly explained by the following Figure 4. From Figure 4, we see that in each block, after receiving the channel feedback of Receiver j ($j \in \{1,2,...,L\}$), the transmitter encodes the transmitted message of current block and the feedback of Receiver j as a codeword $v_{j,1}^N$, and the channel input of current block is generated via an auxiliary discrete memoryless channel with inputs $v_{1,1}^N$,...,$v_{L,1}^N$, and output x^N. For Receiver j, after receiving the channel outputs of all blocks, he/she uses the backward and jointly typical decoding scheme to decode the transmitted $v_{j,1}^N$ for all blocks. If $v_{j,1}^N$ is decoded without error, the messages of all blocks can be obtained by Receiver j. The details about the encoding-decoding scheme of Theorem 1 are in Appendix A.

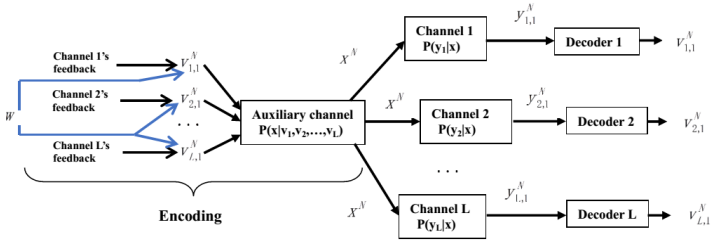

Fig. 4: A feedback scheme for the compound wiretap channel

Remark 1: 1) In [29, Theorem 1], it has been shown that a lower bound R_s on the secrecy capacity $\mathcal{C}_s$ of the compound wiretap channel is given by

$$\mathcal{C}_s \geq R_s = \max \min_j [I(U;Y_j) - I(U;Z)]^+, \quad (8)$$

where the joint distribution is denoted by

$$\begin{aligned} &P(z,y_1,...,y_L,x,u) \\ &= P(z|x)P(y_1,...,y_L|x)P(x|u)P(u) \\ &= P(z|x)P(x|u)P(u)\prod_{j=1}^{L} P(y_j|x). \end{aligned} \quad (9)$$

In general, we do not know whether R_s^f is larger than R_s or not. In Section IV, from a Gaussian fading example, we show that R_s^f is larger than R_s, which indicates that CF may increase the achievable secrecy rate of the compound wiretap channel.

2) For the compound wiretap channel with noiseless feedback, it is natural to ask: why not directly using the noiseless feedback channels for secure communication, is the total secrecy rate of the original compound wiretap channel and the noiseless feedback channels larger than R_s^f given in Theorem 1? To answer this question, a binary symmetric case of the model of Figure 2 is investigated (see Figure 5), and we show that for this special case, directly using noiseless feedback channels for secure communication may not always be the best choice.

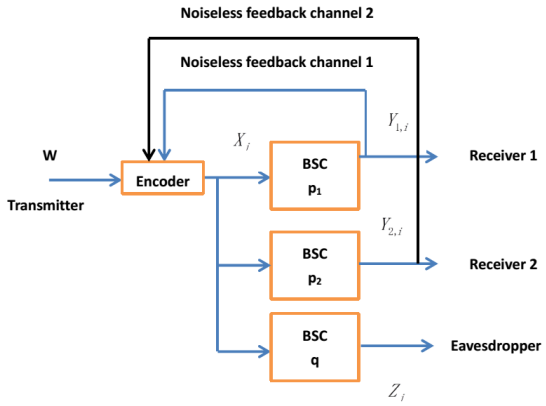

Fig. 5: A binary symmetric case of the compound wiretap channel with noiseless feedback

In Figure 5, one transmitter wishes to send a message W to two legitimate receivers via two BSCs with crossover probabilities p_1 and p_2, respectively, and an eavesdropper tries to eavesdrop W via another BSC with crossover probability q. In addition, the legitimate receivers send their received signals back to the transmitter via two noiseless feedback channels. From Theorem 1, we see that an achievable secrecy rate R_s^f for the model of Figure 5 is given by

$$R_s^f = \max \min_{j=1,2} \min\{[I(V_j;Y_j,U_j) - I(V_j;Z)]^+, I(V_j;Y_j)\}. \quad (10)$$

Then defining $P(V_1=0)=\alpha$, $P(V_1=1)=1-\alpha$, $P(V_2=0)=\beta$, $P(V_2=1)=1-\beta$, V_1 is independent of V_2, and letting $X = V_1 + V_2$, $U_1 = V_1 + Y_1$, $U_2 =$

$V_2 + Y_2$, we have

$$R_s^f = \max_{\alpha,\beta} \min \begin{pmatrix} \min\{[H(\alpha) - H(\alpha \star \beta \star q) \\ +H(\beta \star q)]^+, \\ H(\alpha \star \beta \star p_1) - H(\beta \star p_1)\}, \\ \min\{[H(\beta) - H(\alpha \star \beta \star q) \\ +H(\alpha \star q)]^+, \\ H(\alpha \star \beta \star p_2) - H(\alpha \star p_2)\} \end{pmatrix} \quad (11)$$

where $H(a) = -a\log(a) - (1-a)\log(1-a)$ and $a \star b = a(1-b) + (1-a)b$.

Next, if the noiseless feedback channels of Figure 5 are used for direct transmission, the model of Figure 5 should be revised as the model in the following Figure 6. In Figure 6, two messages W_1 and W_2 are transmitted, where W_1 is transmitted through the binary symmetric compound wiretap channel in the model of Figure 5 without feedback, and W_2 is transmitted through the noiseless feedback channels and due to the broadcast nature of wireless communication, W_2 can also be eavesdropped by the eavesdropper via a binary symmetric wiretap channel with crossover probability q. From [29, Theorem 1], an achievable secrecy rate R_1^* of W_1 is given by

$$R_1^* = \max_{P(x)} \min_{j=1,2} [I(X;Y_j) - I(X;Z)]^+$$
$$\stackrel{(a)}{=} \min\{[H(q) - H(p_1)]^+, [H(q) - H(p_2)]^+\}, \quad (12)$$

where (a) is from defining $P(X=0) = \alpha$, $P(X=1) = 1-\alpha$, and using the fact that the maximum is achieved when $\alpha = \frac{1}{2}$. Analogously, an achievable secrecy rate R_2^* of W_2 is given by

$$R_2^* = \max_{P(x')} \min_{j=1,2} [I(X';Y_j') - I(X';Z')]^+$$
$$\stackrel{(b)}{=} H(q), \quad (13)$$

where (b) follows from substituting $p_1 = p_2 = 0$ into (12). Hence the total secrecy rate $R^* = R_1^* + R_2^*$ is given by

$$R^* = \min\{[H(q) - H(p_1)]^+, [H(q) - H(p_2)]^+\} + H(q). \quad (14)$$

The following Figure 7 plots the achievable secrecy rate R_s^f of the model of Figure 5 and the total secrecy rate R^* of the model of Figure 6 for $p_1 = 0.0001$, $p_2 = 0.5$ and several values of q. From this figure, we see that the feedback scheme proposed in Theorem 1 performs no better than directly using the feedback channels for secure transmission when one legitimate receiver's channel is completely noisy (i.e., $p_2 = 0.5$).

The following Figure 8 plots R_s^f and R^* for $p_1 = 0.0001$, $p_2 = 0.3$ and several values of q. From this figure, we see that the feedback scheme proposed in Theorem 1 performs better than directly using the feedback channel for secure transmission when q is very small.

From the above figures, we see that directly using the feedback channels for secure transmission may not always be the best choice, and sometimes the feedback scheme of Theorem 1 may perform better.

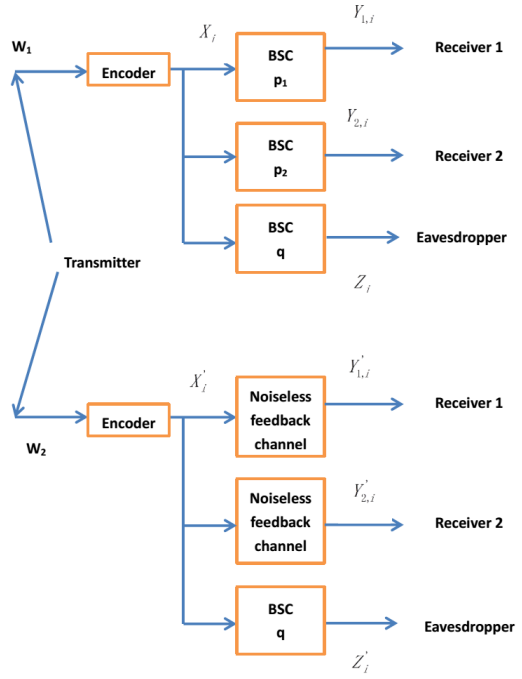

Fig. 6: The binary symmetric compound wiretap channel model with noiseless feedback channels for direct transmission

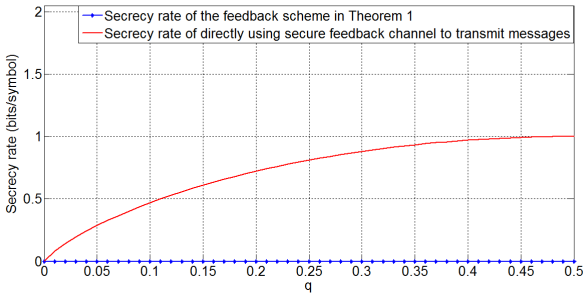

Fig. 7: Comparison of the secrecy rates R_s^f and R^* for $p_1 = 0.0001$, $p_2 = 0.5$ and several values of q

Theorem 2: **Upper bound on $\mathcal{C}_s^f$:** $\mathcal{C}_s^f \le \mathcal{C}_s^{f-out}$, where

$$\mathcal{C}_s^{f-out} = \min_j \max_{P(x)} I(X;Y_j), \quad (15)$$

and the joint probability is defined as

$$P(z, y_1, ..., y_L, x) = P(z|x) \prod_{i=1}^{L} P(y_i|x). \quad (16)$$

Proof: This outer bound can be directly obtained by using the fact that the secrecy capacity can not exceed the capacity of each channel and feedback does not increase the capacity of a discrete memoryless channel. Hence the secrecy capacity $\mathcal{C}_s^f$

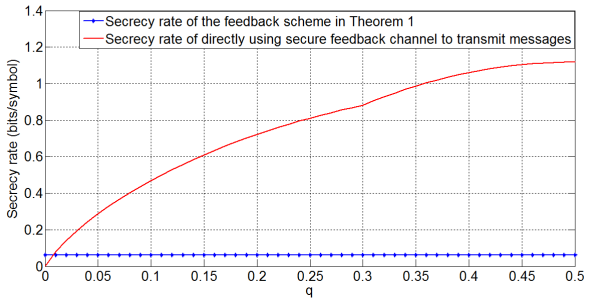

Fig. 8: Comparison of the secrecy rates R_s^f and R^* for $p_1 = 0.0001$, $p_2 = 0.3$ and several values of q

is upper bounded by the minimum of each channel's capacity (here note that the capacity of channel j is $\max_{P(x)} I(X;Y_j)$), and the proof is completed. ■

Here note that the compound wiretap channel with feedback investigated in this section is only an ideal model for the PLS in the down-link transmission of the IoT system. In the next section, we will study a more practical model, which we call the finite state compound wiretap channel with delayed feedback. The lower bound on the secrecy capacity of this more practical model is constructed according to a variation of the feedback strategy in Theorem 1, see the remainder of this paper.

III. The Finite State Compound Wiretap Channel with Delayed Feedback

The practical IoT systems often consist of time-varying and fading channels, and the states of these channels are often obtained by the transmitter via receivers's delayed feedback. In [33], the time-varying fading channel in the presence of one transmitter, one legitimate receiver, one eavesdropper and delayed channel feedback is modeled as the finite state Markov wiretap channel (FSM-WTC) with delayed feedback. In this section, we extend the FSM-WTC with delayed feedback to a more general case, i.e., the finite state compound wiretap channel with delayed feedback, see Figure 3. In Figure 3, the channel consists of multiple legitimate receivers and one eavesdropper, and each legitimate receiver sends his/her received signal back to the transmitter via a corresponding feedback channel with different delayed feedback time. In the remainder of this section, we first give formal definition of the model of Figure 3, and then we show bounds on the secrecy capacity of this new model.

Model formulation:

- The overall channel transition probability of the model of Figure 3 is given by

$$\begin{aligned} &P(y_{1,1}^N, y_{2,1}^N, ..., y_{L,1}^N, z^N | x^N, s_{1,1}^N, ..., s_{L,1}^N, s_{e,1}^N) \\ &= P(z^N | x^N, s_{e,1}^N) \prod_{j=1}^{L} P(y_{j,1}^N | x^N, s_{j,1}^N) \\ &= \prod_{i=1}^{N} \left(P(z_i | x_i, s_{e,i}) \prod_{j=1}^{L} P(y_{j,i} | x_i, s_{j,i}) \right), \end{aligned} \quad (17)$$

where $x_i \in \mathcal{X}$, $y_{j,i} \in \mathcal{Y}_j$, $z_i \in \mathcal{Z}$, $s_{e,i} \in \mathcal{S}_e$ and $s_{j,i} \in \mathcal{S}_j$.

- The finite state processes $\{S_{e,i}\}$ and $\{S_{j,i}\}$ ($j \in \{1, 2, ..., L\}$) are supposed to be stationary irreducible aperiodic ergodic Markov chains. The state processes are independent of one another, and they are independent of the transmitted message. Moreover, the state process $\{S_{j,i}\}$ is independent of the channel input and outputs given the previous states, i.e.,

$$P(s_{j,i} | x^i, y_{1,1}^i, ..., y_{L,1}^i, s_{j,1}^{i-d_j}) = P(s_{j,i} | s_{j,i-d_j}), \quad (18)$$

where $1 \leq d_j \leq i-1$. The state process $\{S_{e,i}\}$ is independent of the channel input and the eavesdropper's channel output given the previous states, i.e.,

$$P(s_{e,i} | x^i, z^i, s_{e,1}^{i-1}) = P(s_{e,i} | s_{e,i-1}). \quad (19)$$

Define the one-step transition probability matrix of the state process $\{S_{j,i}\}$ as K_j. Denote the steady state probabilities of $\{S_{j,i}\}$ and $\{S_{e,i}\}$ by π_j and π_e, respectively. Note that

$$Pr\{S_{j,i} = s_m, S_{j,i-d_j} = s_l\} = \pi_j(s_l) K_j^{d_j}(s_l, s_m), \quad (20)$$

where $s_m, s_l \in \mathcal{S}_j$, s_m is the m-th element of $\mathcal{S}_j$, s_l is the l-th element of $\mathcal{S}_j$, $K_j^{d_j}(s_l, s_m)$ is the (l,m)-th element of the d_j-step transition probability matrix $K_j^{d_j}$ of the Markov process.

- The transmitted message W is uniformly distributed over $\mathcal{W} = \{1, 2, ..., |\mathcal{W}|\}$, and it is independent of the state processes $\{S_{e,i}\}$ and $\{S_{j,i}\}$ ($j \in \{1, 2, ..., L\}$). Receiver $j \in \{1, 2, ..., L\}$ sends his/her received signal back to the transmitter via a feedback channel after a delay time d_j. *Without loss of generality (W.L.O.G.), assume that* $1 \leq d_1 \leq d_2 \leq ... \leq d_L \leq N$. For the case that all legitimate receivers only send their received channel states back to the transmitter via feedback channels with delay times d_1,...,d_L, the i-th ($i \in \{1, 2, ..., N\}$) channel input X_i is given by

$$X_i = \begin{cases} f_i(W), & 1 \leq i \leq d_1, \\ f_i(W, S_{1,1}^{i-d_1}), & d_1 \leq i \leq d_2, \\ ....., & ....., \\ f_i(W, S_{1,1}^{i-d_1}, ..., S_{L-1,1}^{i-d_{L-1}}), & d_{L-1} \leq i \leq d_L, \\ f_i(W, S_{1,1}^{i-d_1}, ..., S_{L,1}^{i-d_L}), & d_L \leq i \leq N. \end{cases} \quad (21)$$

For the case that all legitimate receivers send their received channel outputs and channel states back to the transmitter via feedback channels with delay times d_1,...,d_L, the i-th ($i \in \{1, 2, ..., N\}$) channel input X_i is given by

$$X_i = \begin{cases} f_i(W), & 1 \leq i \leq d_1, \\ f_i(W, S_{1,1}^{i-d_1}, Y_{1,1}^{i-d_1}), & d_1 \leq i \leq d_2, \\ ....., & ....., \\ f_i\left(\begin{matrix} W, S_{1,1}^{i-d_1}, Y_{1,1}^{i-d_1}, \\ ..., S_{L,1}^{i-d_L}, Y_{L,1}^{i-d_L} \end{matrix} \right), & d_L \leq i \leq N. \end{cases} \quad (22)$$

Here note that f_i in (21) and (22) is a stochastic encoding function.

- For Receiver j ($j \in \{1,2,...,L\}$), after receiving $Y_{j,1}^N$ and $S_{j,1}^N$, he/she produces an estimation $\hat{W}^{(j)} = \psi_j(Y_{j,1}^N, S_{j,1}^N)$, and his/her average decoding error probability is defined as

$$P_{e,j} = \frac{1}{|\mathcal{W}|}\sum_{i\in\mathcal{W}} Pr\{\psi_j(y_{j,1}^N, s_{j,1}^N) \neq i | i \text{ sent}\}. \quad (23)$$

The secrecy level of the transmitted message W at the eavesdropper is formulated as

$$\Delta = \frac{1}{N}H(W|Z^N, S_{e,1}^N). \quad (24)$$

The definition of a non-negative number R achieving weak perfect secrecy is the same as that in (5).

The secrecy capacity of the model of Figure 3 with delayed channel output feedback is denoted by $\mathcal{C}_s^{f-dy}$, and without delayed channel output feedback is denoted by $\mathcal{C}_s^{f-d}$. Bounds on $\mathcal{C}_s^{f-dy}$ and $\mathcal{C}_s^{f-d}$ are given in the following theorems.

Theorem 3: **Lower bound on** $\mathcal{C}_s^{f-dy}$: $\mathcal{C}_s^{f-dy} \geq R_s^{f-dy}$, where

$$R_s^{f-dy} = \max \min_j \min\{[I(V_j;Y_j,U_j|S_j,\tilde{S}_j) - I(V_j;Z|S_e)]^+,\ I(V_j;Y_j|S_j,\tilde{S}_j)\}, \quad (25)$$

the auxiliary random variable $\tilde{S}_j$ represents $S_{j,i-d_L}$, S_j represents $S_{j,i}$, and the joint probability is defined as

$$\begin{aligned} &P(z,y_1,...,y_L,x,v_1,...,v_L,u_1,...,u_L,s_1,...,s_L,\\ &\tilde{s}_1,...,\tilde{s}_L,s_e)\\ &= P(z|x,s_e)P(x|v_1,...,v_L)P(s_e)\cdot\\ &\prod_{j=1}^{L}(P(y_j|x,s_j)P(u_j|v_j,y_j,\tilde{s}_j)P(v_j|\tilde{s}_j)\cdot\\ &P(\tilde{s}_j)K_j^{d_L}(\tilde{s}_j,s_j)). \end{aligned} \quad (26)$$

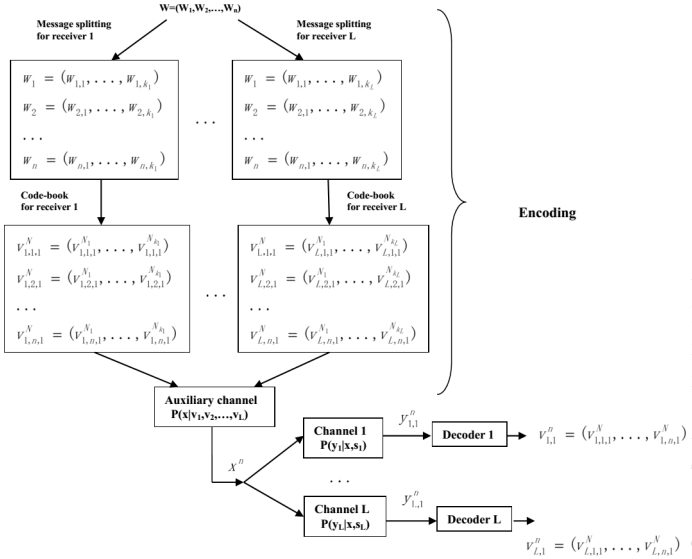

Fig. 9: A feedback scheme for the finite state compound wiretap channel with delayed states and channel output feedback

Proof sketch:

The lower bound R_s^{f-dy} is constructed by combining the coding scheme of Theorem 1 with the multiplexing encoding-decoding scheme for the finite state Markov wiretap channel with delayed feedback [33], and the encoding-decoding procedure is briefly explained by the following Figure 9. From Figure 9, we see that in block i ($i \in \{1,2,...,N\}$), the transmitted message W_i for Receiver j ($j \in \{1,2,...,L\}$) is further divided into k_j sub-messages ($W_i = (W_{i,1}, W_{i,2}, ..., W_{i,k_j})$), where k_j is the size of the alphabet $\mathcal{S}_j$, i.e., $k_j = |\mathcal{S}_j|$. Moreover, in block i, after receiving the delayed feedback channel output and state of Receiver j, the transmitter encodes each sub-message $W_{i,l}$ ($l \in \{1,2,...,k_j\}$) and the delayed feedback as a sub-codeword $v_{j,i,1}^{N_l}$ (similar to the coding scheme in the proof of Theorem 1, see Appendix A), where N_l is the sub-codeword length for $W_{i,l}$, and $\sum_{l=1}^{k_j} N_l = N$. Hence the total codeword $v_{j,i,1}^N$ for W_i is the multiplexing of all sub-codewords $v_{j,i,1}^{N_l}$ for $l \in \{1,2,...,k_j\}$, i.e., $v_{j,i,1}^N = (v_{j,i,1}^{N_1}, v_{j,i,1}^{N_2}, ..., v_{j,i,1}^{N_{k_j}})$. Similar to the coding scheme of Theorem 1, the channel input of block i is generated via an auxiliary discrete memoryless channel with inputs $v_{1,i,1}^N$,...,$v_{L,i,1}^N$, and output x^N.

In the decoding procedure, for Receiver j, after receiving the channel outputs and states of all blocks, he/she uses the backward, de-multiplexing and jointly typical decoding scheme to decode the transmitted $v_{j,1}^n = (v_{j,1,1}^N, v_{j,2,1}^N, ..., v_{j,n,1}^N)$ of all blocks. If $v_{j,1}^n$ is decoded without error, the messages for all blocks can be obtained by Receiver j. The details about the encoding-decoding scheme of Theorem 3 are in Appendix B.

Theorem 4: **Lower bound on** $\mathcal{C}_s^{f-d}$: $\mathcal{C}_s^{f-d} \geq R_s^{f-d}$, where

$$R_s^{f-d} = \max \min_j [I(V_j;Y_j|S_j,\tilde{S}_j) - I(V_j;Z|S_e)]^+, \quad (27)$$

the auxiliary random variable $\tilde{S}_j$ represents $S_{j,i-d_L}$, S_j represents $S_{j,i}$, and the joint probability is defined as

$$\begin{aligned} &P(z,y_1,...,y_L,x,v_1,...,v_L,s_1,...,s_L,\tilde{s}_1,...,\tilde{s}_L,s_e)\\ &= P(z|x,s_e)P(x|v_1,...,v_L)P(s_e)\cdot\\ &\prod_{j=1}^{L}(P(y_j|x,s_j)P(v_j|\tilde{s}_j)P(\tilde{s}_j)K_j^{d_L}(\tilde{s}_j,s_j)). \end{aligned} \quad (28)$$

Proof: First, recall that in the proof of Theorems 1 and 3, the auxiliary random variable U_1,...,U_L are generated by the channel output feedback and they are used to improve the legitimate receivers' decoding performance. Then, note that in Theorem 4, there is no channel output feedback, which indicates that U_1,...,U_L are useless. Finally, substituting $U_1 = U_2 = ... = U_L = const$ into R_s^{f-dy}, and along the lines of the proof of Theorem 3, the lower bound R_s^{f-d} is obtained. The proof of Theorem 4 is completed. ■

The following Theorem 5 provides an upper bound for both $\mathcal{C}_s^{f-dy}$ and $\mathcal{C}_s^{f-d}$, see the followings.

Theorem 5: **Upper bound on both** $\mathcal{C}_s^{f-dy}$ **and** $\mathcal{C}_s^{f-d}$: $\mathcal{C}_s^{f-dy} \leq \mathcal{C}_s^{f-out}$ and $\mathcal{C}_s^{f-d} \leq \mathcal{C}_s^{f-out}$, where

$$\mathcal{C}_s^{f-out} = \min_j \max_{P(x|\tilde{s}_j^*)} I(X;Y_j|S_j,\tilde{S}_j^*), \quad (29)$$

the auxiliary random variable $\tilde{S}_j^*$ represents $S_{j,i-d_j}$, S_j represents $S_{j,i}$,

$$P(y_j, x, s_j, \tilde{s}_j^*) = P(y_j|x, s_j)P(x|\tilde{s}_j^*)P(\tilde{s}_j^*)K_j^{d_j}(\tilde{s}_j^*, s_j), \quad (30)$$

for all $j \in \{1, 2, ..., L\}$.

Proof: This outer bound can be directly obtained by using the fact that the secrecy capacities $\mathcal{C}_s^{f-dy}$ and $\mathcal{C}_s^{f-d}$ can not exceed the capacity of each channel without secrecy constraint. To be specific, first, note that in the model of Figure 3, the channel j ($j \in \{1, 2, ..., L\}$) with delayed feedback and without eavesdropper has already been investigated by [34]. It has been shown in [34] that the capacity $\mathcal{C}^{f-d}$ of each channel with only delayed state feedback equals the capacity $\mathcal{C}^{f-dy}$ of the same channel with delayed both state and channel output feedback, and they are given by

$$\mathcal{C}^{f-d} = \mathcal{C}^{f-dy} = \max_{P(x|\tilde{s}_j^*)} I(X; Y_j|S_j, \tilde{S}_j^*). \quad (31)$$

Then, using the fact that $\mathcal{C}_s^{f-dy}$ and $\mathcal{C}_s^{f-d}$ can not exceed (31) for all $j \in \{1, 2, ..., L\}$, the upper bound $\mathcal{C}_s^{f-out}$ is obtained. The proof is completed. ∎

IV. A Gaussian Fading Example of the Finite State Compound Wiretap Channel with Delayed Feedback

In this section, we compute the capacity bounds in Section III via a Gaussian fading example, and we would like to know how the delayed feedback time affects the capacity bounds. The remainder of this section is organized as follows. In Subsection IV-A, we show bounds on the secrecy capacities of the Gaussian fading case of the model of Figure 3 with or without delayed channel output feedback. In Subsection IV-B, the bounds in Subsection IV-A are further explained via numerical results.

A. Gaussian fading case of the model of Figure 3

For the Gaussian fading case of the model of Figure 3, at time i ($1 \le i \le N$), the channel inputs and outputs are given by

$$Y_{j,i} = h_j(s_{j,i})X_i + N_{s_{j,i}},\ Z_i = g(s_{e,i})X_i + N_{s_{e,i}}, \quad (32)$$

where $j \in \{1, 2, ..., L\}$, and $s_{j,i}$, $h_j(s_{j,i})$, $s_{e,i}$, $g(s_{e,i})$, $N_{s_{j,i}}$ and $N_{s_{e,i}}$ are defined as follows.

- $s_{j,i}$ is the i-th time state of the channel for Receiver j, and $h_j(s_{j,i})$ is the fading coefficient of the channel from the transmitter to Receiver j and it depends on the i-th time state $s_{j,i}$.
- $g(s_{e,i})$ is the fading coefficient of the channel from the transmitter to the eavesdropper and it depends on the i-th time state $s_{e,i}$.
- $N_{s_{j,i}} \sim \mathcal{N}(0, \sigma^2_{s_{j,i}})$ is the noise of the channel from the transmitter to Receiver j and it is Gaussian distributed with zero mean and variance $\sigma^2_{s_{j,i}}$ which depends on the i-th time state $s_{j,i}$.
- $N_{s_{e,i}} \sim \mathcal{N}(0, \sigma^2_{s_{e,i}})$ is the noise of the channel from the transmitter to the eavesdropper and it is Gaussian distributed with zero mean and variance $\sigma^2_{s_{e,i}}$ which depends on the i-th time state $s_{e,i}$.

Let $\mathcal{P}$ be the transmitter's power constraint satisfying

$$E[X^2] \le \mathcal{P}. \quad (33)$$

At the i-th time, Receiver $j \in \{1, 2, ..., L\}$ obtains $S_{j,i}$ and the channel output $Y_{j,i}$, and then he/she transmits $S_{j,i}$ (or $S_{j,i}$ and $Y_{j,i}$) back to the transmitter via a feedback channel after a delay time d_j (W.L.O.G., assume that $1 \le d_1 \le d_2 \le ... \le d_L \le N$). The following Corollary 1 shows the lower bound R_s^{f-dy*} on the secrecy capacity $\mathcal{C}_s^{f-dy*}$ of the Gaussian fading case of the model of Figure 3 with delayed states and channel output feedback. Corollary 2 shows the lower bound R_s^{f-d*} on the secrecy capacity $\mathcal{C}_s^{f-d*}$ of the Gaussian fading case of the model of Figure 3 with only delayed state feedback. Corollary 3 shows an upper bound for both $\mathcal{C}_s^{f-dy*}$ and $\mathcal{C}_s^{f-d*}$.

Corollary 1: A lower bound R_s^{f-dy*} on $\mathcal{C}_s^{f-dy*}$ is given by

$$\begin{aligned} R_s^{f-dy*} = & \max_{\substack{\mathcal{P}(\tilde{s}_1), ..., \mathcal{P}(\tilde{s}_L), \alpha_1, ..., \alpha_L: \\ \sum_{\tilde{s}_1} \pi_1(\tilde{s}_1)\mathcal{P}(\tilde{s}_1) \le \mathcal{P} \\ ..... \\ \sum_{\tilde{s}_L} \pi_L(\tilde{s}_L)\mathcal{P}(\tilde{s}_L) \le \mathcal{P} \\ \alpha_1 + ... + \alpha_L = 1,\ \alpha_1, ..., \alpha_L \ge 0}} \min_j \min \\ & \left\{ \left[\sum_{\tilde{s}_j} \sum_{s_j} \pi_j(\tilde{s}_j) K_j^{d_L}(\tilde{s}_j, s_j) \frac{1}{2} \log(2\pi e \alpha_j \mathcal{P}(\tilde{s}_j)) \right. \right. \\ & \left. \left. - \sum_{s_e} \pi_e(s_e) \frac{1}{2} \log \frac{g^2(s_e)\mathcal{P} + \sigma^2_{s_e}}{g^2(s_e)\mathcal{P}(1-\alpha_j) + \sigma^2_{s_e}} \right]^+, \right. \\ & \left. \sum_{\tilde{s}_j} \sum_{s_j} \pi_j(\tilde{s}_j) K_j^{d_L}(\tilde{s}_j, s_j) \cdot \right. \\ & \left. \frac{1}{2} \log \frac{h_j^2(s_j)\mathcal{P}(\tilde{s}_j) + \sigma^2_{s_j}}{h_j^2(s_j)\mathcal{P}(\tilde{s}_j)(1-\alpha_j) + \sigma^2_{s_j}} \right\}, \end{aligned} \quad (34)$$

where $[x]^+ = x$ for $x \ge 0$, $[x]^+ = 0$ for $x < 0$, and $\mathcal{P}(\tilde{s}_j)$ ($j \in \{1, 2, ..., L\}$) is the transmitter's power allocated to the state $\tilde{s}_j$.

Proof: First, for $j \in \{1, 2, ..., L\}$, define

$$X = \sum_{j=1}^{L} V_j, \quad (35)$$

where $V_j \sim \mathcal{N}(0, \mathcal{P}_j)$, $\mathcal{P}_j \le \alpha_j \mathcal{P}$, $\alpha_j \ge 0$ and $\sum_{j=1}^{L} \alpha_j = 1$. Here note that V_1,...,V_L are independent of one another. From the definition (35), it is easy to check that the power constraint (33) holds. Next, define

$$E[X^2|\tilde{s}_j] = \mathcal{P}(\tilde{s}_j), \quad (36)$$

where $\mathcal{P}(\tilde{s}_j)$ is the transmitter's power allocated to the state $\tilde{s}_j$, and it satisfies

$$\sum_{\tilde{s}_j} \pi_j(\tilde{s}_j)\mathcal{P}(\tilde{s}_j) = \sum_{\tilde{s}_j} \pi_j(\tilde{s}_j) E[X^2|\tilde{s}_j] = E[X^2] \le \mathcal{P}. \quad (37)$$

Further define

$$E[V_j^2|\tilde{s}_j] = \alpha_j \mathcal{P}(\tilde{s}_j). \tag{38}$$

From (37) and (38), it is easy to check that

$$\begin{aligned}\mathcal{P}_j &= E[V_j^2] = \sum_{\tilde{s}_j} \pi_j(\tilde{s}_j) E[V_j^2|\tilde{s}_j] \\ &= \sum_{\tilde{s}_j} \pi_j(\tilde{s}_j)\alpha_j \mathcal{P}(\tilde{s}_j) = \alpha_j \sum_{\tilde{s}_j} \pi_j(\tilde{s}_j)\mathcal{P}(\tilde{s}_j) \\ &\le \alpha_j \mathcal{P}. \end{aligned} \tag{39}$$

Finally, note that for $j \in \{1, 2, ..., L\}$, U_j is generated from the feedback Y_j and the transmitted codeword V_j, define

$$U_j = V_j + Y_j. \tag{40}$$

Now substituting the above definitions (35), (36), (38), (40) and (32) into Theorem 3, R_s^{f-dy*} is obtained. The proof of Corollary 1 is completed. ∎

Corollary 2: A lower bound R_s^{f-d*} on $\mathcal{C}_s^{f-d*}$ is given by

$$\begin{aligned} R_s^{f-d*} = \max_{\substack{\mathcal{P}(\tilde{s}_1),\ldots,\mathcal{P}(\tilde{s}_L),\alpha_1,\ldots,\alpha_L: \\ \sum_{\tilde{s}_1} \pi_1(\tilde{s}_1)\mathcal{P}(\tilde{s}_1) \le \mathcal{P} \\ \ldots\ldots \\ \sum_{\tilde{s}_L} \pi_L(\tilde{s}_L)\mathcal{P}(\tilde{s}_L) \le \mathcal{P} \\ \alpha_1 + \ldots + \alpha_L = 1,\ \alpha_1,\ldots,\alpha_L \ge 0}} \min_j \Bigg[\sum_{\tilde{s}_j}\sum_{s_j} \\ \pi_j(\tilde{s}_j) K_j^{d_L}(\tilde{s}_j, s_j) \frac{1}{2}\log \frac{h_j^2(s_j)\mathcal{P}(\tilde{s}_j) + \sigma_{s_j}^2}{h_j^2(s_j)\mathcal{P}(\tilde{s}_j)(1-\alpha_j) + \sigma_{s_j}^2} \\ - \sum_{s_e} \pi_e(s_e) \frac{1}{2}\log \frac{g^2(s_e)\mathcal{P} + \sigma_{s_e}^2}{g^2(s_e)\mathcal{P}(1-\alpha_j) + \sigma_{s_e}^2} \Bigg]^+, \end{aligned} \tag{41}$$

where $[x]^+ = x$ for $x \ge 0$, $[x]^+ = 0$ for $x < 0$, and $\mathcal{P}(\tilde{s}_j)$ ($j \in \{1, 2, ..., L\}$) is the transmitter's power allocated to the state $\tilde{s}_j$.

Proof: Substituting the definitions (35), (36), (38) and (32) into Theorem 4, R_s^{f-dy*} is obtained. The proof of Corollary 2 is completed. ∎

The following Corollary 3 provides an upper bound for both $\mathcal{C}_s^{f-dy*}$ and $\mathcal{C}_s^{f-d*}$, see the followings.

Corollary 3: An upper bound $\mathcal{C}_s^{f-out*}$ on both $\mathcal{C}_s^{f-dy*}$ and $\mathcal{C}_s^{f-d*}$ is given by

$$\begin{aligned} \mathcal{C}_s^{f-out*} = \min_j \max_{\substack{\mathcal{P}(\tilde{s}_j): \\ \sum_{\tilde{s}_j} \pi_j(\tilde{s}_j)\mathcal{P}(\tilde{s}_j) \le \mathcal{P}}} \sum_{\tilde{s}_j}\sum_{s_j} \pi_j(\tilde{s}_j) K_j^{d_j}(\tilde{s}_j, s_j \\ \cdot \frac{1}{2}\log \frac{h_j^2(s_j)\mathcal{P}(\tilde{s}_j) + \sigma_{s_j}^2}{\sigma_{s_j}^2}. \end{aligned} \tag{42}$$

Proof: Substituting the definitions (36) and (32) into Theorem 5, $\mathcal{C}_s^{f-out*}$ is obtained. The proof of Corollary 3 is completed. ∎

B. Numerical results

In this subsection, we investigate a two-state example, i.e., for $j \in \{1, 2, ..., L\}$, $\mathcal{S}_j$ consists of two elements G_j (good state) and B_j (bad state), and $\mathcal{S}_e$ also consists of two elements G_e and B_e. The state process of $\{S_j\}$ is given by

$$\begin{aligned} P(G_j|G_j) = 1 - b_j, P(B_j|G_j) = b_j, \\ P(B_j|B_j) = 1 - g_j, P(G_j|B_j) = g_j, \end{aligned} \tag{43}$$

and the steady probabilities of G_j and B_j are given by

$$\pi_j(G_j) = \frac{g_j}{g_j + b_j},\ \pi_j(B_j) = \frac{b_j}{g_j + b_j}. \tag{44}$$

In addition, the state process of $\{S_e\}$ is given by

$$\begin{aligned} P(G_e|G_e) = 1 - b_e, P(B_e|G_e) = b_e, \\ P(B_e|B_e) = 1 - g_e, P(G_e|B_e) = g_e, \end{aligned} \tag{45}$$

and the steady probabilities of G_e and B_e are given by

$$\pi_e(G_e) = \frac{g_e}{g_e + b_e},\ \pi_e(B_e) = \frac{b_e}{g_e + b_e}. \tag{46}$$

For the noise N_{s_j} of the channel from the transmitter to Receiver j, its variance $\sigma_{s_j}^2$ in state G_j is $\sigma_{G_j}^2$, and in state B_j is $\sigma_{B_j}^2$. Similarly, the noise N_{s_e} of the channel from the transmitter to the eavesdropper, its variance $\sigma_{s_e}^2$ in state G_e is $\sigma_{G_e}^2$, and in state B_e is $\sigma_{B_e}^2$.

For $L = 3$, which indicates that there are 3 legitimate receivers in the model of Figure 3, the following Figure 10 plots the lower and upper bounds on the secrecy capacities of the Gaussian fading case of the model of Figure 3 with delayed state feedback, and with or without delayed channel output feedback for $\sigma_{G_1}^2 = 0.1$, $\sigma_{B_1}^2 = 1$, $\sigma_{G_2}^2 = 0.3$, $\sigma_{B_2}^2 = 0.66$, $\sigma_{G_3}^2 = 1$, $\sigma_{B_3}^2 = 2$, $\sigma_{G_e}^2 = 2000$, $\sigma_{B_e}^2 = 6000$, $g_1 = 0.05$, $b_1 = 0.05$, $g_2 = 0.1$, $b_2 = 0.08$, $g_3 = 0.2$, $b_3 = 0.08$, $g_e = 0.3$, $b_e = 0.5$, $h_1(G_1) = 1$, $h_1(B_1) = 0.5$, $h_2(G_2) = 0.9$, $h_2(B_2) = 0.6$, $h_3(G_3) = 0.8$, $h_3(B_3) = 0.4$, $g(G_e) = 0.8$, $g(B_e) = 0.7$, $d_1 = 1$, $d_2 = 2$, $d_3 = 3$ and several values of $\mathcal{P}$. As depicted in this figure, if the eavesdropper's channel noise variance ($\sigma_{G_e}^2 = 2000$, $\sigma_{B_e}^2 = 6000$) is large, the lower bound R_s^{f-dy*} meets the upper bound $\mathcal{C}_s^{f-out*}$, which indicates that the secrecy capacity of the Gaussian fading case of the model of Figure 3 with delayed both states and channel output feedback is determined. Moreover, we see that channel output feedback helps to enhance the achievable secrecy rate of the Gaussian fading case of the model of Figure 3 with only delayed state feedback.

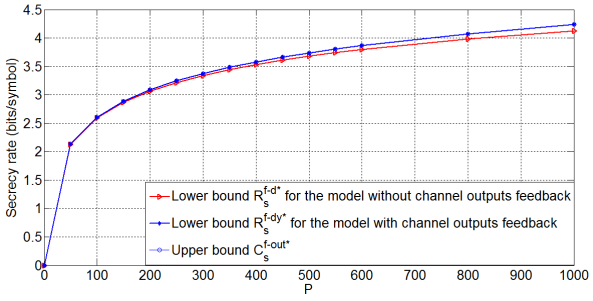

Fig. 10: Comparison of the bounds in Theorems 1-3 for $\sigma_{G_1}^2 = 0.1$, $\sigma_{B_1}^2 = 1$, $\sigma_{G_2}^2 = 0.3$, $\sigma_{B_2}^2 = 0.66$, $\sigma_{G_3}^2 = 1$, $\sigma_{B_3}^2 = 2$, $\sigma_{G_e}^2 = 2000$, $\sigma_{B_e}^2 = 6000$, $g_1 = 0.05$, $b_1 = 0.05$, $g_2 = 0.1$, $b_2 = 0.08$, $g_3 = 0.2$, $b_3 = 0.08$, $g_e = 0.3$, $b_e = 0.5$, $h_1(G_1) = 1$, $h_1(B_1) = 0.5$, $h_2(G_2) = 0.9$, $h_2(B_2) = 0.6$, $h_3(G_3) = 0.8$, $h_3(B_3) = 0.4$, $g(G_e) = 0.8$, $g(B_e) = 0.7$, $d_1 = 1$, $d_2 = 2$, $d_3 = 3$ and several values of $\mathcal{P}$

Figure 11 plots the bounds for the same values of the parameters given in Figure 10 except that $\sigma_{G_e}^2 = 1$ and

$\sigma^2_{B_e} = 2.5$. As depicted in this figure, if the eavesdropper's channel noise variance ($\sigma^2_{G_e} = 1$, $\sigma^2_{B_e} = 2.5$) is decreasing, the gap between the lower and upper bounds on $\mathcal{C}_s^{f-dy*}$ is increasing. In addition, we see that channel output feedback still helps to enhance the achievable secrecy rate R_s^{f-d*} of the Gaussian fading case of the model of Figure 3 with only delayed state feedback.

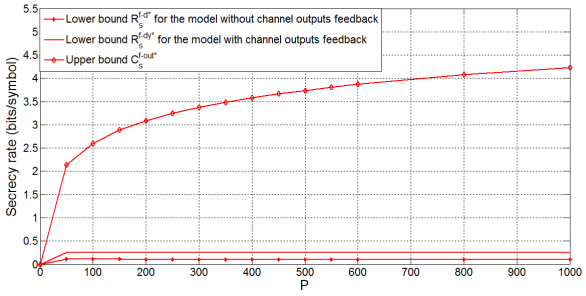

Fig. 11: Comparison of the bounds in Theorems 1-3 for the same values of the parameters in Figure 10 except that $\sigma^2_{G_e} = 1$ and $\sigma^2_{B_e} = 2.5$

Figure 12 plots the bounds for the same values of the parameters given in Figure 10 except that $\sigma^2_{G_e} = 0.03$ and $\sigma^2_{B_e} = 0.05$. As depicted in this figure, if the eavesdropper's channel noise variance ($\sigma^2_{G_e} = 0.03$, $\sigma^2_{B_e} = 0.05$) is sufficiently small, the achievable secrecy rate R_s^{f-d*} of the Gaussian fading case of the model of Figure 3 with only delayed state feedback equals 0, which implies that the perfect secrecy can not be guaranteed for this case. Using the channel output feedback, the positive achievable secrecy rate R_s^{f-dy*} is derived, and hence the PLS of the Gaussian fading case of the model of Figure 3 with only delayed state feedback is enhanced. In addition, we should notice that there still exists a huge gap between the lower and upper bounds on $\mathcal{C}_s^{f-dy*}$.

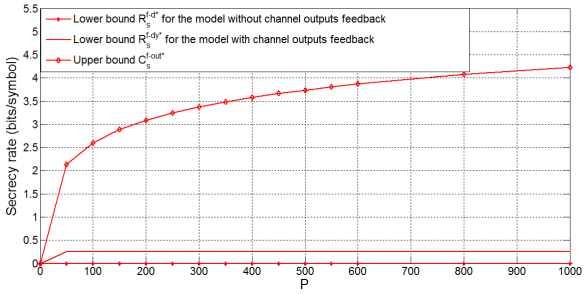

Fig. 12: Comparison of the bounds in Theorems 1-3 for the same values of the parameters in Figure 10 except that $\sigma^2_{G_e} = 0.03$ and $\sigma^2_{B_e} = 0.05$

To investigate how the delayed feedback time d_j ($j \in \{1, 2, ..., L\}$) affects the secrecy rates of the model of Figure 3, the following Figure 13 plots the lower bounds in Theorems 1 and 2 for the case that $L = 3$ (three legitimate receivers) and $d_1 = d_2 = d_3 = d$, which implies that the delayed feedback times of the three legitimate receivers are the same and equal d. As depicted in this figure, the achievable secrecy rates of the Gaussian fading case of the model of Figure 3 with or without delayed channel output feedback are decreasing while the delay time d is increasing. However, we should notice that both R_s^{f-dy*} and R_s^{f-d*} are approaching their infinite asymptotes while d is sufficiently large.

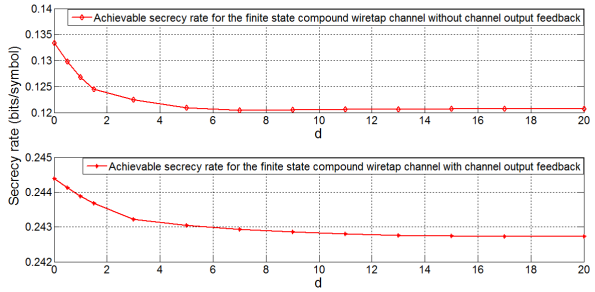

Fig. 13: Comparison of the lower bounds in Theorems 1 and 2 for $\mathcal{P} = 20$, $\sigma^2_{G_1} = 0.1$, $\sigma^2_{B_1} = 1$, $\sigma^2_{G_2} = 0.3$, $\sigma^2_{B_2} = 0.66$, $\sigma^2_{G_3} = 1$, $\sigma^2_{B_3} = 2$, $\sigma^2_{G_e} = 1$, $\sigma^2_{B_e} = 2.5$, $g_1 = 0.05$, $b_1 = 0.05$, $g_2 = 0.1$, $b_2 = 0.08$, $g_3 = 0.2$, $b_3 = 0.08$, $g_e = 0.3$, $b_e = 0.5$, $h_1(G_1) = 1$, $h_1(B_1) = 0.5$, $h_2(G_2) = 0.9$, $h_2(B_2) = 0.6$, $h_3(G_3) = 0.8$, $h_3(B_3) = 0.4$, $g(G_e) = 0.8$, $g(B_e) = 0.7$ and several values of d ($d_1 = d_2 = d_3 = d$)

The following Figure 14 plots the lower bounds in Theorems 1 and 2 for the case that $L = 3$ (three legitimate receivers) and $d_1 = 0$, $d_2 = d$, $d_3 = 2d$, which implies that the delayed feedback times of the three legitimate receivers are different. Similar to Figure 13, R_s^{f-dy*} and R_s^{f-d*} are monotonic decreasing functions of d, and both of them approach their infinite asymptotes when d is sufficiently large.

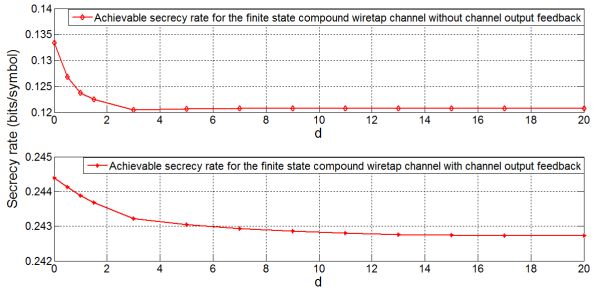

Fig. 14: Comparison of the lower bounds in Theorems 1 and 2 for $\mathcal{P} = 20$, $\sigma^2_{G_1} = 0.1$, $\sigma^2_{B_1} = 1$, $\sigma^2_{G_2} = 0.3$, $\sigma^2_{B_2} = 0.66$, $\sigma^2_{G_3} = 1$, $\sigma^2_{B_3} = 2$, $\sigma^2_{G_e} = 1$, $\sigma^2_{B_e} = 2.5$, $g_1 = 0.05$, $b_1 = 0.05$, $g_2 = 0.1$, $b_2 = 0.08$, $g_3 = 0.2$, $b_3 = 0.08$, $g_e = 0.3$, $b_e = 0.5$, $h_1(G_1) = 1$, $h_1(B_1) = 0.5$, $h_2(G_2) = 0.9$, $h_2(B_2) = 0.6$, $h_3(G_3) = 0.8$, $h_3(B_3) = 0.4$, $g(G_e) = 0.8$, $g(B_e) = 0.7$ and several values of d ($d_1 = 0$, $d_2 = d$, $d_3 = 2d$)

V. Conclusion

This paper establishes a general framework for enhancing the PLS in the down-link transmission of IoT systems via feedback. Two models including the compound wiretap channel with feedback and the finite state compound wiretap channel with delayed feedback are studied, and bounds on

the secrecy capacities of the two models are given. From a Gaussian fading example, we see that the delayed channel output feedback enhances the lower bound on the secrecy capacity of the finite state compound wiretap channel with only delayed state feedback, and the corresponding feedback strategy may achieve the secrecy capacity if the eavesdropper's channel noise variance is sufficiently large. Moreover, numerical results indicate that the secrecy rates are decreasing while the feedback delay time is increasing, and the secrecy rates are approaching their infinite asymptotes while the feedback delay time is sufficiently large. However, we should notice that all the capacity results given in this paper only work well under the perfect weak secrecy condition, and how to design the corresponding encoding-decoding schemes under the strong perfect secrecy condition is of further interest to us.

APPENDIX A
PROOF OF THEOREM 1

The messages are conveyed to the receivers via n blocks. The blocklength of block $i \in \{1, 2, ..., n-1\}$ is N, and for block n (the last block), its blocklength is γN, where γ is a positive real number and will be determined later. In block i ($i \in \{1, 2, ..., n-1\}$), the random sequences X^N, Z^N, $Y_{1,1}^N$, $Y_{2,1}^N$,...,$Y_{L,1}^N$, $U_{1,1}^N$, $U_{2,1}^N$,...,$U_{L,1}^N$, $V_{1,1}^N$, $V_{2,1}^N$,...,$V_{L,1}^N$, are denoted by $\bar{X}_i$, $\bar{Z}_i$, $\bar{Y}_{1,i}$, $\bar{Y}_{2,i}$,...,$\bar{Y}_{L,i}$, $\bar{U}_{1,i}$, $\bar{U}_{2,i}$,...,$\bar{U}_{L,i}$, $\bar{V}_{1,i}$, $\bar{V}_{2,i}$,...,$\bar{V}_{L,i}$, respectively. Similarly, in block n, the random sequences $X^{\gamma N}$, $Z^{\gamma N}$, $Y_{1,1}^{\gamma N}$, $Y_{2,1}^{\gamma N}$,...,$Y_{L,1}^{\gamma N}$, $V_{1,1}^{\gamma N}$, $V_{2,1}^{\gamma N}$,...,$V_{L,1}^{\gamma N}$, are denoted by $\bar{X}_n$, $\bar{Z}_n$, $\bar{Y}_{1,n}$, $\bar{Y}_{2,n}$,...,$\bar{Y}_{L,n}$, $\bar{V}_{1,n}$, $\bar{V}_{2,n}$,...,$\bar{V}_{L,n}$, respectively. In addition, the value of the random vector is written in lower case letter.

Code-book construction:

- The message W is sent to all legitimate receivers via n blocks, i.e., the message W is composed of n components ($W = (W_1, ..., W_n)$), and each component W_i ($i \in \{1, 2, ..., n\}$) is the message for block i. Here W_i takes values in the set $\{1, ..., 2^{NR}\}$.
- The dummy message $W^{'}$, which is used to confuse the eavesdropper, is composed of n components ($W^{'} = (W_1^{'}, ..., W_n^{'})$), and the component $W_i^{'}$ ($i \in \{1, 2, ..., n\}$) is transmitted in block i. Here note that $W_i^{'}$ is randomly chosen from the set $\{1, ..., 2^{NR^{'}}\}$, i.e., $Pr\{W_i^{'} = l\} = 2^{-NR^{'}}$, where $l \in \{1, ..., 2^{NR^{'}}\}$.
- The auxiliary message W_j^*, where $j \in \{1, 2, ..., L\}$, is used to improve Receiver j's decoding performance. Here note that W_j^* is composed of $n-1$ components ($W_j^* = (W_{j,1}^*, ..., W_{j,n-1}^*)$), where $W_{j,i}^*$ ($i \in \{1, 2, ..., n-1\}$) takes values in the set $\{1, ..., 2^{NR_j^*}\}$.
- In block i ($1 \le i \le n$), randomly generate $2^{N(R+R^{'}+R^*)}$ independent identically distributed (i.i.d.) sequences $\bar{v}_{j,i}$ ($j \in \{1, 2, ..., L\}$) according to the probability $P(v_j)$, and label them as $\bar{v}_{j,i}(w_i, w_i^{'}, w_{j,i-1}^*)$, where $w_i \in \{1, 2, ..., 2^{NR}\}$, $w_i^{'} \in \{1, 2, ..., 2^{NR^{'}}\}$ and $w_{j,i-1}^* \in \{1, 2, ..., 2^{NR_j^*}\}$.
- In block i ($1 \le i \le n-1$), for each possible value of $\bar{v}_{j,i}(w_i, w_i^{'}, w_{j,i-1}^*)$ and $\bar{y}_{j,i}$, randomly generate $2^{N\tilde{R}_j}$ i.i.d. codewords $\bar{u}_{j,i}$ according to the probability $P(u_j|v_j, y_j)$. Then label these $\bar{u}_{j,i}$ as $\bar{u}_{j,i}(w_{j,i}^*, w_{j,i}^{**})$, where $w_{j,i}^* \in \{1, 2, ..., 2^{NR_j^*}\}$ and $w_{j,i}^{**} \in \{1, 2, ..., 2^{N(\tilde{R}_j - R_j^*)}\}$.
- In block i ($1 \le i \le n$), for given $\bar{v}_{1,i}$,...,$\bar{v}_{L,i}$, the channel input $\bar{x}_i$ is i.i.d. generated according to $P(x|v_1, ..., v_L)$.

For convenience, the following Figure 15 provides some important notations in the proof of Theorem 1.

Notation	Meaning
X^N	Random vector with length N
$\bar{X}_i$	The i-th block's random vector
$\bar{X}_{j,i}$	The i-th block's random vector for receiver j
W_i	The transmitted message for block i
$W_i^{'}$	The dummy message for block i
$W_{j,i}^*$	The auxiliary message for receiver j and block i

Fig. 15: Some important notations in the proof of Theorem 1

Encoding procedure:

- At block 1, the transmitter selects $\bar{v}_{j,1}(w_1, w_1^{'}, 1)$ ($j \in \{1, 2, ..., L\}$). Here notice that $w_1^{'}$ is randomly chosen from the set $\{1, 2, ..., 2^{NR^{'}}\}$.
- At block i ($i \in \{2, 3, ..., n-1\}$), once the transmitter receives $\bar{y}_{j,i-1}$ ($j \in \{1, 2, ..., L\}$), he seeks a $\bar{u}_{j,i-1}$ such that the triplet $(\bar{u}_{j,i-1}, \bar{v}_{j,i-1}, \bar{y}_{j,i-1})$ is jointly typical. If there exist multiple $\bar{u}_{j,i-1}$, randomly choose one, and if no such $\bar{u}_{j,i-1}$ exists, an error occurs. Based on the covering Lemma [32], if

$$\tilde{R}_j \ge I(U_j; V_j, Y_j), \tag{A1}$$

this encoding error vanishes as N tends to infinity. Once the transmitter decodes such a $\bar{u}_{j,i-1}(w_{j,i-1}^*, w_{j,i-1}^{**})$, he chooses $\bar{v}_{j,i}(w_i, w_i^{'}, w_{j,i-1}^*)$ for transmission.
- At block n, once the transmitter receives the feedback $\bar{y}_{j,n-1}$ ($j \in \{1, 2, ..., L\}$), he seeks a $\bar{u}_{j,n-1}$ such that the triplet $(\bar{u}_{j,n-1}, \bar{v}_{j,n-1}, \bar{y}_{j,n-1})$ is jointly typical, and the corresponding encoding error vanishes as N tends to infinity if (A1) is guaranteed. Once the transmitter decodes such a $\bar{u}_{j,n-1}(w_{j,n-1}^*, w_{j,n-1}^{**})$, he chooses $\bar{v}_{j,n}(1, 1, w_{j,n-1}^*)$ for transmission.

Decoding procedure:

Receiver j's ($j \in \{1, 2, ..., L\}$) decoding scheme begins from the last block. At block n, first, note that due to the reason that the blocklength is γN, the actual rate of $W_{j,n-1}^*$ is given by

$$\frac{H(W_{j,n-1}^*)}{\gamma N} = \frac{NR_j^*}{\gamma N} = \frac{R_j^*}{\gamma}. \tag{A2}$$

Next, Receiver j selects a unique $\bar{v}_{j,n}$ such that $(\bar{v}_{j,n}, \bar{y}_{j,n})$ are joint typical. If multiple or no such $\bar{v}_{j,n}$ exists, an error occurs. From the packing lemma [32], this error vanishes if

$$\frac{R_j^*}{\gamma} \le I(V_j; Y_j). \tag{A3}$$

Since $I(V_j; Y_j)$ of (A3) is finite, for any given R_j^*, we can choose a sufficiently large γ such that (A3) is guaranteed.

After decoding $\bar{v}_{j,n}$, Receiver j picks out $w^*_{j,n-1}$ from it. Then he/she tries to choose a unique $\bar{u}_{j,n-1}$ such that given $w^*_{j,n-1}$, $\bar{u}_{j,n-1}$ and $\bar{y}_{j,n-1}$ are jointly typical. If multiple or no such $\bar{u}_{j,n-1}$ exists, an error occurs. From the packing lemma [32], this error vanishes if

$$\tilde{R}_j - R^*_j \leq I(U_j; Y_j). \tag{A4}$$

Once such unique $\bar{u}_{j,n-1}$ is obtained, Receiver j seeks a unique $\bar{v}_{j,n-1}$ such that $(\bar{v}_{j,n-1}, \bar{y}_{j,n-1}, \bar{u}_{j,n-1})$ are jointly typical. From the packing lemma [32], this error vanishes if

$$R + R^{'} + R^*_j \leq I(V_j; U_j, Y_j). \tag{A5}$$

After decoding $\bar{v}_{j,n-1}$, Receiver j picks out w_{n-1}, and $w^*_{j,n-2}$ from it. Repeating the above decoding procedure, the messages of all blocks are decoded by Receiver j. The decoding procedure is completed. The following Figures 16 and 17 illustrate the encoding-decoding procedure for Receiver j ($j \in \{1, 2, ..., L\}$).

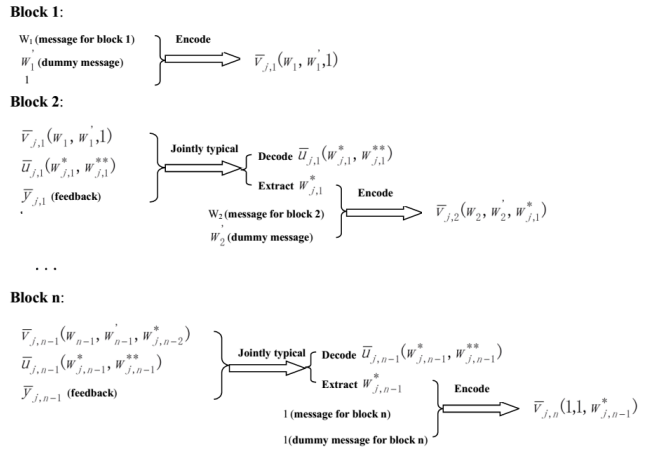

Fig. 16: The encoding procedure for $j \in \{1, 2, ..., L\}$

Equivocation Analysis:

The eavesdropper's equivocation Δ, defined as $\Delta = \frac{1}{(n-1)N+\gamma N} H(W|\bar{Z}_1, ..., \bar{Z}_n)$ (the overall length of all blocks is $(n-1)N + \gamma N$), follows that

$$\begin{aligned}
\Delta &= \frac{1}{(n-1)N+\gamma N} H(W|\bar{Z}_1, ..., \bar{Z}_n) \\
&\overset{(a)}{=} \frac{1}{(n-1)N+\gamma N} \sum_{i=1}^{n-1} H(W_i|W_1, ..., W_{i-1}, \bar{Z}_1, ..., \bar{Z}_n) \\
&\overset{(b)}{=} \frac{1}{(n-1)N+\gamma N} \sum_{i=1}^{n-1} H(W_i|\bar{Z}_i) \\
&= \frac{1}{(n-1)N+\gamma N} \sum_{i=1}^{n-1} (H(W_i, \bar{Z}_i) - H(\bar{Z}_i)) \\
&= \frac{1}{(n-1)N+\gamma N} \sum_{i=1}^{n-1} (H(W_i, \bar{Z}_i, \bar{V}_{j,i}) \\
&- H(\bar{V}_{j,i}|W_i, \bar{Z}_i) - H(\bar{Z}_i)) \\
&\overset{(c)}{=} \frac{1}{(n-1)N+\gamma N} \sum_{i=1}^{n-1} (H(\bar{Z}_i|\bar{V}_{j,i}) + H(\bar{V}_{j,i}) \\
&- H(\bar{V}_{j,i}|W_i, \bar{Z}_i) - H(\bar{Z}_i)) \\
&= \frac{1}{(n-1)N+\gamma N} \sum_{i=1}^{n-1} (H(\bar{V}_{j,i}) - I(\bar{V}_{j,i}; \bar{Z}_i) \\
&- H(\bar{V}_{j,i}|W_i, \bar{Z}_i)) \\
&= \frac{1}{(n-1)N+\gamma N} (\sum_{i=1}^{n-1} H(\bar{V}_{j,i}) - \sum_{i=1}^{n-1} I(\bar{V}_{j,i}; \bar{Z}_i) \\
&- \sum_{i=1}^{n-1} H(\bar{V}_{j,i}|W_i, \bar{Z}_i)) \\
&\overset{(d)}{\geq} \frac{1}{(n-1)N+\gamma N} (\sum_{i=1}^{n-1} H(\bar{V}_{j,i}) \\
&- (n-1)N(I(V_j; Z) + \epsilon_1) - \sum_{i=1}^{n-1} H(\bar{V}_{j,i}|W_i, \bar{Z}_i)) \\
&= \frac{1}{(n-1)N+\gamma N} (H(\bar{V}_{j,1}) + \sum_{i=2}^{n-1} H(\bar{V}_{j,i}) \\
&- (n-1)N(I(V_j; Z) + \epsilon_1) \\
&- H(\bar{V}_{j,1}|W_1, \bar{Z}_1) - \sum_{i=2}^{n-1} H(\bar{V}_{j,i}|W_i, \bar{Z}_i)) \\
&\overset{(e)}{\geq} \frac{1}{(n-1)N+\gamma N} (N(R + R^{'} - \epsilon_2) \\
&+ (n-2)N(R + R^{'} + R^*_j - \epsilon_3) \\
&- (n-1)N(I(V_j; Z) + \epsilon_1)
\end{aligned}$$

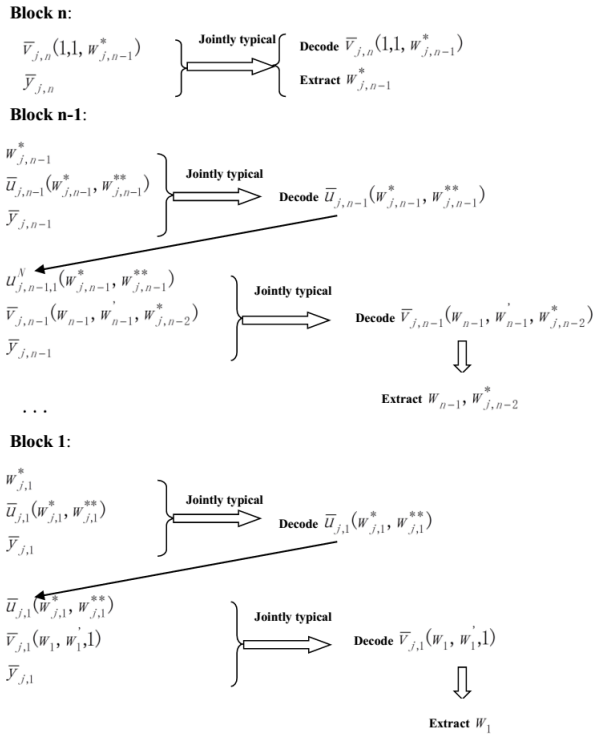

Fig. 17: The decoding procedure for Receiver j ($j \in \{1, 2, ..., L\}$)

$$
\begin{aligned}
&-H(\bar{V}_{j,1}|W_1,\bar{Z}_1)-\sum_{i=2}^{n-1}H(\bar{V}_{j,i}|W_i,\bar{Z}_i))\\
&\overset{(f)}{\geq}\frac{1}{(n-1)N+\gamma N}(N(R+R^{'}-\epsilon_2)\\
&+(n-2)N(R+R^{'}+R_j^*-\epsilon_3)\\
&-(n-1)N(I(V_j;Z)+\epsilon_1)-N\epsilon_4-(n-2)N\epsilon_5)\\
&=\frac{R+R^{'}-\epsilon_2}{(n-1)+\gamma}+\frac{(n-2)(R+R^{'}+R_j^*-\epsilon_3)}{(n-1)+\gamma}\\
&-\frac{(n-1)(I(V_j;Z)+\epsilon_1)}{(n-1)+\gamma}-\frac{\epsilon_4+(n-2)\epsilon_5}{(n-1)+\gamma},
\end{aligned}
\tag{A6}
$$

where (a) follows from the fact that $W=(W_1,...,W_n)$ and W_n is constant, (b) follows from the fact that given $\bar{Z}_i$, the message W_i of block i is independent of other blocks' $(\bar{Z}_1,...,\bar{Z}_{i-1},\bar{Z}_{i+1},...,\bar{Z}_n)$ and previous blocks' messages $(W_1,...,W_{i-1})$, i.e., the Markov chain $W_i\rightarrow\bar{Z}_i\rightarrow(W_1,...,W_{i-1},\bar{Z}_1,...,\bar{Z}_{i-1},\bar{Z}_{i+1},...,\bar{Z}_n)$ holds, (c) follows from the fact that $H(W_i|\bar{V}_{j,i})=0$, (d) follows from a similar argument in [5, Lemma 3], i.e., $I(\bar{V}_{j,i};\bar{Z}_i)\leq N(I(V_j;Z)+\epsilon_1)$, where $\epsilon_1\rightarrow 0$ as $N\rightarrow\infty$, (e) follows from the construction of $\bar{V}_{j,i}$ and a similar argument in [3, equations (16) and (23)], i.e., $H(\bar{V}_{j,1})\geq N(R+R^{'}-\epsilon_2)$, $H(\bar{V}_{j,i})\geq N(R+R^{'}+R_j^*-\epsilon_3)$, where $\epsilon_2,\epsilon_3\rightarrow 0$ as $N\rightarrow\infty$, (f) follows from that given w_1, $\bar{z}_1$, the eavesdropper attempts to find a unique $\bar{v}_{j,1}$ jointly typical with his/her received $\bar{z}_1$, and from the packing lemma [32], this decoding error vanishes if

$$
R^{'}\leq I(V_j;Z),
\tag{A7}
$$

then applying Fano's lemma, $H(\bar{V}_{j,1}|W_1,\bar{Z}_1)\leq N\epsilon_4$ is obtained, where $\epsilon_4\rightarrow 0$ while $N\rightarrow\infty$, and analogously, for $i\in\{2,...,n-1\}$ given w_i, $\bar{z}_i$, the eavesdropper attempts to find a unique $\bar{v}_{j,i}$ jointly typical with his/her received $\bar{z}_i$, and from the packing lemma [32], this decoding error vanishes if

$$
R^{'}+R_j^*\leq I(V_j;Z),
\tag{A8}
$$

then applying Fano's lemma, $H(\bar{V}_{j,i}|W_i,\bar{Z}_i)\leq N\epsilon_5$ is obtained, where $\epsilon_5\rightarrow 0$ while $N\rightarrow\infty$. Here note that (A7) is included in (A8), and thus we only need to use (A8) to derive the final region.

The bound (A6) implies that if

$$
R^{'}+R_j^*\geq I(V_j;Z),
\tag{A9}
$$

$\Delta\geq R-\epsilon$ is satisfied by choosing sufficiently large n and N.

Now it remains to use the above conditions (A1), (A4), (A5), (A8) and (A9) to derive the lower bound in Theorem 1, see the followings.

First, note that from (A1) and (A4), we have

$$
\begin{aligned}
R_j^*&\geq I(U_j;V_j,Y_j)-I(U_j;Y_j)\\
&=I(U_j;V_j|Y_j).
\end{aligned}
\tag{A10}
$$

Next, substituting (A10) into (A5), we get

$$
\begin{aligned}
R\leq R+R^{'}&\leq I(V_j;U_j,Y_j)-I(U_j;V_j|Y_j)\\
&=I(V_j;Y_j).
\end{aligned}
\tag{A11}
$$

Then, note that from (A8) and (A9), we can conclude that

$$
R^{'}+R_j^*=I(V_j;Z).
\tag{A12}
$$

Now substituting (A12) into (A5), we have

$$
R\leq I(V_j;U_j,Y_j)-I(V_j;Z).
\tag{A13}
$$

From the above (A11) and (A13), we have

$$
R\leq\min\{I(V_j;U_j,Y_j)-I(V_j;Z),I(V_j;Y_j)\}.
\tag{A14}
$$

Next, note that if $I(V_j;U_j,Y_j)\leq I(V_j;Z)$, from (A5), we have

$$
R+R^{'}+R_j^*\leq I(V_j;U_j,Y_j)\leq I(V_j;Z).
\tag{A15}
$$

Combining (A15) with (A12), and observing that $R\geq 0$, we can conclude that $R=0$ if $I(V_j;U_j,Y_j)\leq I(V_j;Z)$. Hence (A14) should be re-written as

$$
R\leq\min\{[I(V_j;U_j,Y_j)-I(V_j;Z)]^+,I(V_j;Y_j)\}.
\tag{A16}
$$

Note that (A16) should be satisfied for all $j\in\{1,2,...,L\}$, hence we have

$$
R\leq\min_j\min\{[I(V_j;U_j,Y_j)-I(V_j;Z)]^+,I(V_j;Y_j)\}.
\tag{A17}
$$

Finally, note that the effective transmission rate is

$$
\begin{aligned}
\frac{H(W)}{(n-1)N+\gamma N}&=\frac{\sum_{i=1}^{n-1}H(W_i)}{(n-1)N+\gamma N}\\
&=\frac{(n-1)NR}{(n-1)N+\gamma N}=\frac{n-1}{n-1+\gamma}R,
\end{aligned}
\tag{A18}
$$

which indicates that the effective transmission rate approaches R as the number of blocks $n\rightarrow\infty$, then maximizing the bound in (A17), Theorem 1 is proved, and the proof is completed.

APPENDIX B
PROOF OF THEOREM 3

The encoding-decoding scheme of Theorem 3 combines that of Theorem 1 with the multiplexing encoding-decoding scheme for the finite state Markov channel with delayed feedback [34]. The detail about the coding scheme is given below.

Definition:

Similar to the definitions in the proof of Theorem 1, the messages are transmitted via n blocks. Since $0\leq d_1\leq d_2\leq ...\leq d_L\leq N$, define the blocklength of block $i\in\{1,2,...,n-d_L\}$ is N, and for block $i\in\{n-d_L+1,...,n\}$, its blocklength is $N^{'}=\gamma N$. For $j\in\{1,2,...,L\}$, define the alphabet $\mathcal{S}_j$ as $\mathcal{S}_j=\{1,2,...,k_j\}$ and the steady probability $\pi_j(l)>0$ for any $l\in\mathcal{S}_j$. Moreover, in block $i\in\{1,2,...,n-d_L\}$, define $N_{\tilde{s}_j}$ ($\tilde{s}_j\in\{1,2,...,k_j\}$) as

$$
N_{\tilde{s}_j}=N\pi_j(\tilde{s}_j).
\tag{A19}
$$

Similarly, in block $i \in \{n-d_L+1,...,n\}$, define $N_{\tilde{s}_j}^{'}$ ($\tilde{s}_j \in \{1,2,...,k_j\}$) as

$$N_{\tilde{s}_j}^{'} = \gamma N_{\tilde{s}_j} = \gamma N \pi_j(\tilde{s}_j). \tag{A20}$$

The random sequence X^N of block $i \in \{1,2,...,n-d_L\}$ and $X^{\gamma N}$ of block $i \in \{n-d_L+1,...,n\}$ are denoted by $\bar{X}_i$, and similar convention is applied to other random sequences. In addition, for block $i \in \{1,2,...,n-d_L\}$, the random sequence $X^{N_{\tilde{s}_j}}$ is denoted by $\bar{X}_i^{N_{\tilde{s}_j}}$, and for block $i \in \{n-d_L+1,...,n\}$, the random sequence $X^{N_{\tilde{s}_j}^{'}}$ is denoted by $\bar{X}_i^{N_{\tilde{s}_j}^{'}}$. Similar convention is applied to other random sequences, and the values of the random sequences are written in lower case letter.

The message W is composed of n components ($W = (W_1,...,W_n)$), and the component W_i ($i \in \{1,2,...,n\}$) is the message for block i. Here W_i takes values in the set $\{1,...,2^{NR}\}$. For $j \in \{1,2,...,L\}$ and given $\tilde{s}_j$ ($\tilde{s}_j \in \{1,2,...,k_j\}$), further divide W_i into k_j sub-messages, i.e., $W_i = (W_{i,1},...,W_{i,k_j})$, where $W_{i,\tilde{s}_j}$ takes values in $\{1,2,...,2^{N_{\tilde{s}_j}R(\tilde{s}_j)}\}$. Here note that

$$\sum_{\tilde{s}_j=1}^{k_j} \pi_j(\tilde{s}_j)R(\tilde{s}_j) = R. \tag{A21}$$

The dummy message $W^{'}$ also consists of n components ($W^{'} = (W_1^{'},...,W_n^{'})$), and $W_i^{'}$ ($i \in \{1,2,...,n\}$) is for block i. Here note that $W_i^{'}$ is uniformly drawn from the set $\{1,...,2^{NR^{'}}\}$, i.e., $Pr\{W_i^{'} = l\} = 2^{-NR^{'}}$, where $l \in \{1,...,2^{NR^{'}}\}$. Similarly, for $j \in \{1,2,...,L\}$ and given $\tilde{s}_j$ ($\tilde{s}_j \in \{1,2,...,k_j\}$), further divide $W_i^{'}$ into k_j sub-messages, i.e., $W_i^{'} = (W_{i,1}^{'},...,W_{i,k_j}^{'})$ and $W_{i,\tilde{s}_j}^{'}$ takes values in $\{1,2,...,2^{N_{\tilde{s}_j}R^{'}(\tilde{s}_j)}\}$. Here note that

$$\sum_{\tilde{s}_j=1}^{k_j} \pi_j(\tilde{s}_j)R^{'}(\tilde{s}_j) = R^{'}. \tag{A22}$$

The auxiliary message W_j^* ($j \in \{1,2,...,L\}$) is composed of n components ($W_j^* = (W_{j,1}^*,...,W_{j,n}^*)$), where $W_{j,i}^*$ ($i \in \{1,2,...,n\}$) takes values in $\{1,...,2^{NR_j^*}\}$. Similarly, given $\tilde{s}_j$ ($\tilde{s}_j \in \{1,2,...,k_j\}$), further divide $W_{j,i}^*$ into k_j sub-messages, i.e., $W_{j,i}^* = (W_{j,i,1}^*,...,W_{j,i,k_j}^*)$ and $W_{j,i,\tilde{s}_j}^*$ takes values in $\{1,2,...,2^{N_{\tilde{s}_j}R_j^*(\tilde{s}_j)}\}$. Here note that

$$\sum_{\tilde{s}_j=1}^{k_j} \pi_j(\tilde{s}_j)R_j^*(\tilde{s}_j) = R_j^*. \tag{A23}$$

For convenience, the following Figure 18 provides some important notations in the proof of Theorem 3.

Code-book construction:

- In block $i \in \{1,2,...,n-d_L\}$, for $j \in \{1,2,...,L\}$ and given $\tilde{s}_j \in \{1,2,...,k_j\}$, randomly produce $2^{N_{\tilde{s}_j}(R(\tilde{s}_j)+R^{'}(\tilde{s}_j)+R^*(\tilde{s}_j))}$ i.i.d. sequences $\bar{v}_{j,i}^{N_{\tilde{s}_j}}$ according to the probability $P(v_j|\tilde{s}_j)$, and label them as $\bar{v}_{j,i}^{N_{\tilde{s}_j}}(w_{i,\tilde{s}_j}, w_{i,\tilde{s}_j}^{'}, w_{j,i-d_L,\tilde{s}_j}^*)$, where $w_{i,\tilde{s}_j} \in \{1,2,...,2^{N_{\tilde{s}_j}R(\tilde{s}_j)}\}$, $w_{i,\tilde{s}_j}^{'} \in \{1,2,...,2^{N_{\tilde{s}_j}R^{'}(\tilde{s}_j)}\}$ and $w_{j,i-d_L,\tilde{s}_j}^* \in \{1,2,...,2^{N_{\tilde{s}_j}R_j^*(\tilde{s}_j)}\}$.
- In block $i \in \{1,2,...,n-d_L\}$, for $j \in \{1,2,...,L\}$, $\tilde{s}_j \in \{1,2,...,k_j\}$, and each possible value of $\bar{v}_{j,i}^{N_{\tilde{s}_j}}(w_{i,\tilde{s}_j}, w_{i,\tilde{s}_j}^{'}, w_{j,i-d_L,\tilde{s}_j}^*)$ and $\bar{y}_{j,i}^{N_{\tilde{s}_j}}$, randomly generate $2^{N_{\tilde{s}_j}\tilde{R}_j(\tilde{s}_j)}$ i.i.d. codewords $\bar{u}_{j,i}^{N_{\tilde{s}_j}}$ according to the probability $P(u_j|v_j,y_j,\tilde{s}_j)$. Then label these $\bar{u}_{j,i}^{N_{\tilde{s}_j}}$ as $\bar{u}_{j,i}^{N_{\tilde{s}_j}}(w_{j,i,\tilde{s}_j}^*, w_{j,i,\tilde{s}_j}^{**})$, where $w_{j,i,\tilde{s}_j}^* \in \{1,2,...,2^{N_{\tilde{s}_j}R_j^*(\tilde{s}_j)}\}$ and $w_{j,i,\tilde{s}_j}^{**} \in \{1,2,...,2^{N_{\tilde{s}_j}(\tilde{R}_j(\tilde{s}_j)-R_j^*(\tilde{s}_j))}\}$. Here note that

$$\sum_{\tilde{s}_j=1}^{k_j} \pi_j(\tilde{s}_j)\tilde{R}_j(\tilde{s}_j) = \tilde{R}_j. \tag{A24}$$

- In block $i \in \{1,2,...,n-d_L\}$, $\bar{v}_{j,i}$ ($j \in \{1,2,...,L\}$) is generated by multiplexing different sub-sequences $\bar{v}_{j,i}^{N_{\tilde{s}_j}}$ for all $\tilde{s}_j \in \{1,2,...,k_j\}$, i.e., $\bar{v}_{j,i} = (\bar{v}_{j,i}^{N_1},...,\bar{v}_{j,i}^{N_{k_j}})$. In block $i \in \{n-d_L+1,...,n\}$, $\bar{v}_{j,i}$ is generated by multiplexing different sub-sequences $\bar{v}_{j,i}^{N_{\tilde{s}_j}^{'}}$ for all $\tilde{s}_j \in \{1,2,...,k_j\}$, i.e., $\bar{v}_{j,i} = (\bar{v}_{j,i}^{N_1^{'}},...,\bar{v}_{j,i}^{N_{k_j}^{'}})$. Then for given $\bar{v}_{1,i}$,...,$\bar{v}_{L,i}$, the channel input $\bar{x}_i$ is i.i.d. generated according to $P(x|v_1,...,v_L)$.

Encoding procedure:

- At block $i \in \{1,...,2d_L\}$, for each $\tilde{s}_j \in \{1,2,...,k_j\}$, the transmitter selects $\bar{v}_{j,i}^{N_{\tilde{s}_j}}(1,1,1)$ for transmission.
- At block $i \in \{2d_L+1,...,n-d_L\}$, for each $\tilde{s}_j \in \{1,2,...,k_j\}$ and $j \in \{1,2,...,L\}$, the delayed feedback state sequences $\bar{s}_{j,i-2d_L}^{N_{\tilde{s}_j}}$, $\bar{s}_{j,i-d_L}^{N_{\tilde{s}_j}}$, the delayed feed-

Notation	Meaning
$S_{j,i}$	The state for receiver j at time i
$\tilde{S}_j$	The state $S_{j,i-d_L}$
$S_{j,1}^N$	The state sequence for receiver j with length N
$S_{j,1}^{N_k}$	The state sequence for receiver j with length N_k
$\bar{S}_{j,i}$	The i-th block's state sequence for receiver j
$\bar{S}_{j,i}^{N_k}$	The i-th block's state sequence for receiver j with length N_k
$W_{i,k}$	The k-th sub-message of W_i for block i
$W_{i,k}^{'}$	The k-th sub-dummy message of $W_i^{'}$ for block i
$W_{j,i,k}^*$	The k-th sub-auxiliary message of $W_{j,i}^*$ for receiver j and block i

Fig. 18: Some important notations in the proof of Theorem 3

back channel output $\bar{y}_{j,i-d_L}$ and the previous transmitted sequence $\bar{v}_{j,i-d_L}^{N_{\tilde{s}_j}}$ have already been known by the transmitter. Here note that $\bar{s}_{j,i-2d_L}^{N_{\tilde{s}_j}}$ is the delayed feedback state de-multiplexing the delayed feedback channel output $\bar{y}_{j,i-d_L}$ into $\bar{y}_{j,i-d_L}^{N_1}$, ..., $\bar{y}_{j,i-d_L}^{N_{k_j}}$. Given $\tilde{s}_j$, the transmitter seeks a $\bar{u}_{j,i-d_L}^{N_{\tilde{s}_j}}$ such that $(\bar{u}_{j,i-d_L}^{N_{\tilde{s}_j}}, \bar{v}_{j,i-d_L}^{N_{\tilde{s}_j}}, \bar{y}_{j,i-d_L}^{N_{\tilde{s}_j}}, \bar{s}_{j,i-d_L}^{N_{\tilde{s}_j}})$ are jointly typical. If there exist multiple $\bar{u}_{j,i-d_L}^{N_{\tilde{s}_j}}$, randomly pick out one; if no such $\bar{u}_{j,i-d_L}^{N_{\tilde{s}_j}}$ exists, an error occurs. Based on the covering Lemma [32], if

$$\tilde{R}_j(\tilde{s}_j) \geq I(U_j; V_j, Y_j | S_j, \tilde{S}_j = \tilde{s}_j), \quad \text{(A25)}$$

this encoding error vanishes as N tends to infinity. Once the transmitter decodes such a $\bar{u}_{j,i-d_L}^{N_{\tilde{s}_j}}(w_{j,i-d_L,\tilde{s}_j}^*, w_{j,i-d_L,\tilde{s}_j}^{**})$, he chooses $\bar{v}_{j,i}^{N_{\tilde{s}_j}}(w_{i,\tilde{s}_j}, w_{i,\tilde{s}_j}^{'}, w_{j,i-d_L,\tilde{s}_j}^*)$ for transmission.

- At block $i \in \{n - d_L + 1, ..., n\}$, using the previous encoding scheme for $i \in \{2d_L + 1, ..., n - d_L\}$, the transmitter decodes $\bar{u}_{j,i-d_L}^{N_{\tilde{s}_j}}(w_{j,i-d_L,\tilde{s}_j}^*, w_{j,i-d_L,\tilde{s}_j}^{**})$ and chooses $\bar{v}_{j,i}^{N_{\tilde{s}_j}^{'}}(w_{i,\tilde{s}_j} = 1, w_{i,\tilde{s}_j}^{'} = 1, w_{j,i-d_L,\tilde{s}_j}^*)$ for transmission.

Decoding procedure:

Once Receiver $j \in \{1, 2, ..., L\}$ obtains all n blocks $\bar{y}_{j,1}$,...,$\bar{y}_{j,n}$ and $\bar{s}_{j,1}$,...,$\bar{s}_{j,n}$, he/she demultiplexes them into sub-sequences according to $\tilde{s}_j \in \{1, 2, ..., k_j\}$. Receiver j does backward decoding, i.e., the decoding procedure starts from block $i \in \{n - d_L + 1, ..., n\}$. Given $\tilde{s}_j \in \{1, 2, ..., k_j\}$, Receiver j selects a unique $\bar{v}_{j,i}^{N_{\tilde{s}_j}^{'}}$ such that $(\bar{v}_{j,i}^{N_{\tilde{s}_j}^{'}}, \bar{y}_{j,i}^{N_{\tilde{s}_j}^{'}}, \bar{s}_{j,i}^{N_{\tilde{s}_j}^{'}})$ are joint typical. If multiple or no such $\bar{v}_{j,i}^{N_{\tilde{s}_j}^{'}}$ exists, an error occurs. From the packing lemma [32], this error vanishes if

$$\begin{aligned} &\frac{H(W_{j,i-d_L,\tilde{s}_j}^*)}{N_{\tilde{s}_j}^{'}} = \frac{H(W_{j,i-d_L,\tilde{s}_j}^*)}{\gamma N_{\tilde{s}_j}} = \frac{N_{\tilde{s}_j} R_j^*(\tilde{s}_j)}{\gamma N_{\tilde{s}_j}} \\ &= \frac{R_j^*(\tilde{s}_j)}{\gamma} \leq I(V_j; Y_j | S_j, \tilde{S}_j = \tilde{s}_j). \end{aligned} \quad \text{(A26)}$$

Since $I(V_j; Y_j | S_j, \tilde{S}_j = \tilde{s}_j)$ of (A26) is finite, for any given $R_j^*(\tilde{s}_j)$, we can choose a sufficiently large γ such that (A26) is guaranteed.

After decoding $\bar{v}_{j,i}^{N_{\tilde{s}_j}^{'}}$ for block $i \in \{n - d_L + 1, ..., n\}$, Receiver j picks out $w_{j,i-d_L,\tilde{s}_j}^*$ from it. Then given $\tilde{s}_j$, he/she tries to choose a unique $\bar{u}_{j,i-d_L}^{N_{\tilde{s}_j}}$ such that given $w_{j,i-d_L,\tilde{s}_j}^*$, $\bar{u}_{j,i-d_L}^{N_{\tilde{s}_j}}$, $\bar{s}_{j,i-d_L}^{N_{\tilde{s}_j}}$ and $\bar{y}_{j,i-d_L}^{N_{\tilde{s}_j}}$ are jointly typical. If multiple or no such $\bar{u}_{j,i-d_L}^{N_{\tilde{s}_j}}$ exists, an error occurs. From the packing lemma [32], this error vanishes if

$$\tilde{R}_j(\tilde{s}_j) - R_j^*(\tilde{s}_j) \leq I(U_j; Y_j | S_j, \tilde{S}_j = \tilde{s}_j). \quad \text{(A27)}$$

Once such unique $\bar{u}_{j,i-d_L}^{N_{\tilde{s}_j}}$ is decoded, given $\tilde{s}_j$, Receiver j seeks a unique $\bar{v}_{j,i-d_L}^{N_{\tilde{s}_j}}$ such that $(\bar{v}_{j,i-d_L}^{N_{\tilde{s}_j}}, \bar{s}_{j,i-d_L}^{N_{\tilde{s}_j}}, \bar{y}_{j,i-d_L}^{N_{\tilde{s}_j}}, \bar{u}_{j,i-d_L}^{N_{\tilde{s}_j}})$ are jointly typical. From the packing lemma [32], this error vanishes if

$$R(\tilde{s}_j) + R^{'}(\tilde{s}_j) + R_j^*(\tilde{s}_j) \leq I(V_j; U_j, Y_j | S_j, \tilde{S}_j = \tilde{s}_j). \quad \text{(A28)}$$

After decoding $\bar{v}_{j,i-d_L}^{N_{\tilde{s}_j}}$, Receiver j picks out $w_{i-d_L,\tilde{s}_j}$ and $w_{j,i-d_L,\tilde{s}_j}^*$ from it. Repeating the above decoding procedure, the messages of all blocks are decoded by Receiver j. The decoding procedure is completed.

Equivocation Analysis:

The eavesdropper's equivocation Δ, defined as $\Delta = \frac{1}{(n-d_L)N + d_L\gamma N} H(W | \bar{Z}_1, ..., \bar{Z}_n, \bar{S}_{e,1}, ..., \bar{S}_{e,n})$ (the overall length of all blocks is $(n - d_L)N + d_L\gamma N$), follows that

$$\begin{aligned} \Delta &= \frac{1}{(n-d_L)N + d_L\gamma N} H(W | \bar{Z}_1, ..., \bar{Z}_n, \bar{S}_{e,1}, ..., \bar{S}_{e,n}) \\ &\overset{(a)}{=} \frac{1}{(n-d_L)N + d_L\gamma N} \sum_{i=2d_L+1}^{n-d_L} H(W_i | W_{2d_L+1}, ..., \\ &W_{i-1}, \bar{Z}_1, ..., \bar{Z}_n, \bar{S}_{e,1}, ..., \bar{S}_{e,n}) \\ &\overset{(b)}{=} \frac{1}{(n-d_L)N + d_L\gamma N} \sum_{i=2d_L+1}^{n-d_L} H(W_i | \bar{Z}_i, \bar{S}_{e,i}), \end{aligned} \quad \text{(A29)}$$

where (a) follows from the fact that W_i is constant for block $i \in \{1, ..., 2d_L\}$ and $i \in \{n - d_L + 1, ..., n\}$, and (b) follows from the fact that given $\bar{Z}_i$ and $\bar{S}_{e,i}$, the message W_i of block i is independent of other blocks' $(\bar{Z}_1, ..., \bar{Z}_{i-1}, \bar{Z}_{i+1}, ..., \bar{Z}_n)$, $(\bar{S}_{e,1}, ..., \bar{S}_{e,i-1}, \bar{S}_{e,i+1}, ..., \bar{S}_{e,n})$ and previous blocks' messages $(W_{2d_L+1}, ..., W_{i-1})$, i.e., the Markov chain $W_i \rightarrow (\bar{Z}_i, \bar{S}_{e,i}) \rightarrow (W_{2d_L+1}, ..., W_{i-1}, \bar{Z}_1, ..., \bar{Z}_{i-1}, \bar{Z}_{i+1}, ..., \bar{Z}_n, \bar{S}_{e,1}, ..., \bar{S}_{e,i-1}, \bar{S}_{e,i+1}, ..., \bar{S}_{e,n})$ holds.

The term $H(W_i | \bar{Z}_i, \bar{S}_{e,i})$ in (A29) is further bounded by

$$\begin{aligned} &H(W_i | \bar{Z}_i, \bar{S}_{e,i}) \\ &= H(W_i, \bar{Z}_i, \bar{S}_{e,i}) - H(\bar{Z}_i, \bar{S}_{e,i}) \\ &= H(W_i, \bar{Z}_i, \bar{S}_{e,i}, \bar{V}_{j,i}) - H(\bar{V}_{j,i} | W_i, \bar{Z}_i, \bar{S}_{e,i}) \\ &- H(\bar{Z}_i, \bar{S}_{e,i}) \\ &\overset{(c)}{=} H(\bar{Z}_i | \bar{S}_{e,i}, \bar{V}_{j,i}) + H(\bar{S}_{e,i}, \bar{V}_{j,i}) - H(\bar{V}_{j,i} | W_i, \bar{Z}_i, \bar{S}_{e,i}) \\ &- H(\bar{Z}_i, \bar{S}_{e,i}) \\ &\overset{(d)}{=} H(\bar{Z}_i | \bar{S}_{e,i}, \bar{V}_{j,i}) + H(\bar{S}_{e,i}) + H(\bar{V}_{j,i}) \\ &- H(\bar{V}_{j,i} | W_i, \bar{Z}_i, \bar{S}_{e,i}) - H(\bar{Z}_i, \bar{S}_{e,i}) \\ &= H(\bar{V}_{j,i}) - I(\bar{V}_{j,i}; \bar{Z}_i | \bar{S}_{e,i}) - H(\bar{V}_{j,i} | W_i, \bar{Z}_i, \bar{S}_{e,i}) \\ &\overset{(e)}{\geq} N(R + R^{'} + R_j^* - \epsilon_1) - I(\bar{V}_{j,i}; \bar{Z}_i | \bar{S}_{e,i}) \\ &- H(\bar{V}_{j,i} | W_i, \bar{Z}_i, \bar{S}_{e,i}) \\ &\overset{(f)}{\geq} N(R + R^{'} + R_j^* - \epsilon_1) - N(I(V_j; Z | S_e) + \epsilon_2) \\ &- H(\bar{V}_{j,i} | W_i, \bar{Z}_i, \bar{S}_{e,i}) \\ &\overset{(g)}{\geq} N(R + R^{'} + R_j^* - \epsilon_1) - N(I(V_j; Z | S_e) + \epsilon_2) \\ &- N\epsilon_3, \end{aligned} \quad \text{(A30)}$$

where (c) follows from the fact that $H(W_i | \bar{V}_{j,i}) = 0$, (d) follows from the fact that $\bar{S}_{e,i}$ is independent of $\bar{V}_{j,i}$, (e) follows from the construction of $\bar{V}_{j,i}$ and a similar argument in [3,

equations (16) and (23)], i.e., $H(\bar{V}_{j,i}) \geq N(R+R^{'}+R_j^*-\epsilon_1)$, where $\epsilon_1 \to 0$ as $N \to \infty$, (f) follows from a similar argument in [5, Lemma 3], i.e., $I(\bar{V}_{j,i};\bar{Z}_i|\bar{S}_{e,i}) \leq N(I(V_j;Z|S_e)+\epsilon_2)$, where $\epsilon_2 \to 0$ as $N \to \infty$, and (g) follows from that given w_i, $\bar{z}_i$, the eavesdropper attempts to find a unique $\bar{v}_{j,i}$ jointly typical with his/her received $\bar{z}_i$, and from the packing lemma [32], this error vanishes if

$$R^{'} + R_j^* \leq I(V_j;Z|S_e), \tag{A31}$$

then applying Fano's lemma, $H(\bar{V}_{j,i}|W_i,\bar{Z}_i,\bar{S}_{e,i}) \leq N\epsilon_3$ is obtained, where $\epsilon_3 \to 0$ while $N \to \infty$.

Substituting (A30) into (A29), we have

$$\begin{aligned}\Delta &= \frac{n-3d_L}{n-d_L+d_L\gamma}(R+R^{'}+R_j^*-I(V_j;Z|S_e)\\ &\quad -\epsilon_1-\epsilon_2-\epsilon_3).\end{aligned} \tag{A32}$$

The bound (A32) implies that if

$$R^{'} + R_j^* \geq I(V_j;Z|S_e), \tag{A33}$$

$\Delta \geq R-\epsilon$ is satisfied by choosing sufficiently large n and N.

Now it remains to use the above conditions (A25), (A27), (A28), (A31) and (A33) to derive the lower bound in Theorem 3, see the followings.

First, note that from (A24) and (A25), we have

$$\tilde{R}_j \geq I(U_j;V_j,Y_j|S_j,\tilde{S}_j). \tag{A34}$$

Analogously, from (A23), (A24) and (A27), we have

$$\tilde{R}_j - R_j^* \leq I(U_j;Y_j|S_j,\tilde{S}_j). \tag{A35}$$

From (A21), (A22), (A23) and (A28), we have

$$R + R^{'} + R_j^* \leq I(V_j;U_j,Y_j|S_j,\tilde{S}_j). \tag{A36}$$

Next, from (A34) and (A35), we get

$$\begin{aligned}R_j^* &\geq I(U_j;V_j,Y_j|S_j,\tilde{S}_j) - I(U_j;Y_j|S_j,\tilde{S}_j)\\ &= I(U_j;V_j|Y_j,S_j,\tilde{S}_j).\end{aligned} \tag{A37}$$

Next, substituting (A37) into (A36), we get

$$\begin{aligned}R &\leq R + R^{'}\\ &\leq I(V_j;U_j,Y_j|S_j,\tilde{S}_j) - I(U_j;V_j|Y_j,S_j,\tilde{S}_j)\\ &= I(V_j;Y_j|S_j,\tilde{S}_j).\end{aligned} \tag{A38}$$

Then, note that from (A31) and (A33), we can conclude that

$$R^{'} + R_j^* = I(V_j;Z|S_e). \tag{A39}$$

Now substituting (A39) into (A36), we have

$$R \leq I(V_j;U_j,Y_j|S_j,\tilde{S}_j) - I(V_j;Z|S_e). \tag{A40}$$

From the above (A38) and (A40), we have

$$\begin{aligned}R \leq \min\{&I(V_j;U_j,Y_j|S_j,\tilde{S}_j) - I(V_j;Z|S_e),\\ &I(V_j;Y_j|S_j,\tilde{S}_j)\}.\end{aligned} \tag{A41}$$

Next, note that if $I(V_j;U_j,Y_j|S_j,\tilde{S}_j) \leq I(V_j;Z|S_e)$, from (A36), we have

$$R + R^{'} + R_j^* \leq I(V_j;U_j,Y_j|S_j,\tilde{S}_j) \leq I(V_j;Z|S_e). \tag{A42}$$

Combining (A42) with (A39), and observing that $R \geq 0$, we can conclude that $R = 0$ if $I(V_j;U_j,Y_j|S_j,\tilde{S}_j) \leq I(V_j;Z|S_e)$. Hence (A41) should be re-written as

$$\begin{aligned}R \leq \min\{&[I(V_j;U_j,Y_j|S_j,\tilde{S}_j) - I(V_j;Z|S_e)]^+,\\ &I(V_j;Y_j|S_j,\tilde{S}_j)\}.\end{aligned} \tag{A43}$$

Note that (A43) should be satisfied for all $j \in \{1,2,...,L\}$, hence we have

$$\begin{aligned}R \leq \min_j \min\{&[I(V_j;U_j,Y_j|S_j,\tilde{S}_j) - I(V_j;Z|S_e)]^+,\\ &I(V_j;Y_j|S_j,\tilde{S}_j)\}.\end{aligned} \tag{A44}$$

Finally, note that the effective transmission rate is

$$\begin{aligned}\frac{H(W)}{(n-d_L)N+d_L\gamma N} &= \frac{\sum_{i=2d_L+1}^{n-d_L} H(W_i)}{(n-d_L)N+d_L\gamma N}\\ &= \frac{(n-3d_L)NR}{(n-d_L)N+d_L\gamma N} = \frac{n-3d_L}{n-d_L+d_L\gamma}R,\end{aligned} \tag{A45}$$

which indicates that the effective transmission rate approaches R as the number of blocks $n \to \infty$, then maximizing the bound in (A44), Theorem 3 is proved, and the proof is completed.

REFERENCES

[1] A. D. Wyner, "The wire-tap channel," *The Bell System Technical Journal*, vol. 54, no. 8, pp. 1355-1387, 1975.

[2] U. Maurer, S. Wolf, "Information-theoretic key agreement: from weak to strong secrecy for free," *International Conference on the Theory and Applications of Cryptographic Techniques*, pp. 351-368, 2000.

[3] I. Csiszár, J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 3, pp. 339-348, May 1978.

[4] S. K. Leung-Yan-Cheong, M. E. Hellman, "The Gaussian wire-tap channel," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 4, pp. 451-456, July 1978.

[5] R. Liu, I. Maric, P. Spasojević, R.D Yates, "Discrete memoryless interference and broadcast channels with confidential messages: secrecy rate regions," *IEEE Trans. Inf. Theory*, vol. IT-54, no. 6, pp. 2493-2507, Jun. 2008.

[6] J. Xu, Y. Cao, B. Chen, "Capacity bounds for broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. IT-55, no. 6, pp. 4529-4542, 2009.

[7] E. Ekrem, S. Ulukus, "Secrecy capacity of a class of broadcast channels with an eavesdropper," *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, no. 1, pp. 1-29, 2009.

[8] G. Bagherikaram, A. S. Motahari, A. K. Khandani, "The secrecy rate region of the broadcast channel," *Proceedings of the Allerton Conference on Communications, Control and Computing*, 2008.

[9] E. Tekin, A. Yener, "The general Gaussian multiple access and two-way wire-tap channels: Achievable rates and cooperative jamming," *IEEE Trans. Inf. Theory*, vol. IT-54, no. 6, pp. 2735-2751, June 2008.

[10] Q. Zhang, X. Huang, Q. Li, J. Qin, "Cooperative jamming aided robust secure transmission for wireless information and power transfer in MISO channels," *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 906-915, 2015.

[11] G. Zheng, L. C. Choo, K. K. Wong, "Optimal cooperative jamming to enhance physical layer security using relays," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1317-1322, 2011.

[12] K. H. Park, T. Wang, M. S. Alouini, "On the jamming power allocation for secure amplify-and-forward relaying via cooperative jamming," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1741-1750, 2013.

[13] X. Zhou, M. R. McKay, "Secure transmission with artificial noise over fading channels: Achievable rate and optimal power allocation.," *IEEE Trans. Veh. Technol*, vol. 59, no. 8, pp. 3831-3842, 2010.

[14] P. H. Lin, S. H. Lai, S. C. Lin, H. J. Su, "On secrecy rate of the generalized artificial-noise assisted secure beamforming for wiretap channels," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1728-1740, 2013.

[15] J. Zhu, J. Mo, M. Tao, "Cooperative secret communication with artificial noise in symmetric interference channel," *IEEE Communications Letters*, vol. 14, no. 10, pp. 885-887, 2010.

[16] L. Hu, H. Wen, B. Wu, F. Pan, R. Liao, H. Song, J. Tang, X. Wang, "Cooperative jamming for physical layer security enhancement in Internet of Things," *IEEE Internet of Things Journal*, vol. 5, no. 1, pp. 219-228, 2018.

[17] J. M. Liang, J. J. Chen, H. H. Cheng, Y. C. Tseng, "An energy-efficient sleep scheduling with QoS consideration in 3GPP LTE-Advanced networks for Internet of Things, *IEEE J. Emerg. Sel. Top. Circuits Syst.*, vol. 3, no. 1, pp. 13-22, 2013.

[18] A. Mukherjee, "Physical-layer security in the Internet of Things: sensing and communication confidentiality under resource constraints, *Proceedings of IEEE*, vol. 103, no. 10, pp. 1747-1761, 2015.

[19] R. Ahlswede, N. Cai, "Transmission, Identification and Common Randomness Capacities for Wire-Tap Channels with Secure Feedback from the Decoder," book chapter in *General Theory of Information Transfer and Combinatorics*, LNCS 4123, pp. 258-275, Berlin: Springer-Verlag, 2006.

[20] E. Ardestanizadeh, M. Franceschetti, T. Javidi, Y. Kim, "Wiretap channel with secure rate-limited feedback," *IEEE Trans. Inf. Theory*, vol. IT-55, no. 12, pp. 5353-5361, December 2009.

[21] R. F. Schaefer, A. Khisti, H. V. Poor, "Secure broadcasting using independent secret keys," *IEEE Trans. Commun.*, vol. 66, no. 2, pp. 644-661, 2018.

[22] A. Cohen, A. Cohen, "Wiretap channel with causal state information and secure rate-limited feedback," *IEEE Trans. Commun.*, vol. 64, no. 3, pp. 1192-1203, 2016.

[23] B. Dai, Z. Ma, X. Fang, "Feedback enhances the security of state-dependent degraded broadcast channels with confidential messages," *IEEE Trans. Inf. Forensics and Security*, vol. 10, No. 7, pp. 1529-1542, 2015.

[24] B. Dai, Z. Ma, M. Xiao, X. Tang, P. Fan, "Secure communication over finite state multiple-access wiretap channel with delayed feedback," *IEEE J. Sel. Areas Commun.*, vol. 36, No. 4, pp. 723-736, 2018.

[25] B. Dai, Y. Luo, "An improved feedback coding scheme for the wiretap channel," *IEEE Trans. Inf. Forensics and Security*, vol. 14, No. 1, pp. 262-271, 2019.

[26] C. Li, Y. Liang, H. V. Poor, S. Shamai, "A coding scheme for colored Gaussian wiretap channels with feedback," *2018 IEEE International Symposium on Information Theory (ISIT)*, pp. 131-135, 2018.

[27] D. Gunduz, D. R. Brown and H. V. Poor, "Secret communication with feedback," *International Symposium on Information Theory and Its Applications, ISITA 2008*, pp. 1-6, 2008.

[28] J. P. M. Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedback. part I: No bandwidth constraint," *IEEE Trans. Inf. Theory*, vol. 12, no. 2, pp. 172-182, 1966.

[29] Y. Liang, G. Kramer, H. V. Poor, S. Shamai, "Compound wiretap channels," *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, Article ID 142374, 2009.

[30] A. Khisti, "Interference alignment for the multiantenna compound wiretap channel," *IEEE Trans. Inf. Theory*, vol. IT-57, no. 5, pp. 2976-2993, 2011.

[31] I. Bjelakovic, H. Boche, J. Sommerfeld, "Secrecy results for compound wiretap channels," *Problems of Information Transmission*, vol. 49, no. 1, pp. 73-98, 2013.

[32] A. El Gamal, Y. Kim, *Network information theory*. Cambridge University Press, 2011.

[33] B. Dai, Z. Ma and Y. Luo, "Finite state Markov wiretap channel with delayed feedback," *IEEE Trans. Inf. Forensics and Security*, Vol. 12, No. 3, pp. 746-760, 2017.

[34] H. Viswanathan, "Capacity of Markov channels with receiver CSI and delayed feedback," *IEEE Trans. Inf. Theory*, vol. IT-45, no. 2, pp. 761-771, 1999.