Hindawi Publishing Corporation
Journal of Applied Mathematics
Volume 2013, Article ID 423619, 17 pages
http://dx.doi.org/10.1155/2013/423619

Research Article

Wiretap Channel in the Presence of Action-Dependent States and Noiseless Feedback

Bin Dai,[1] A. J. Han Vinck,[2] and Yuan Luo[3]

[1] *School of Information Science and Technology, Southwest JiaoTong University, Chengdu 610031, China*
[2] *Institute for Experimental Mathematics, Duisburg-Essen University, Ellernstraße 29, 45326 Essen, Germany*
[3] *Computer Science and Engineering Department, Shanghai Jiao Tong University, Shanghai 200240, China*

Correspondence should be addressed to Bin Dai; daibinsjtu@gmail.com

Received 29 November 2012; Accepted 3 January 2013

Academic Editor: Jin Liang

We investigate the wiretap channel in the presence of action-dependent states and noiseless feedback. Given the message to be communicated, the transmitter chooses an action sequence that affects the formation of the channel states and then generates the channel input sequence based on the state sequence, the message, and the noiseless feedback, where the noiseless feedback is from the output of the main channel to the channel encoder. The main channel and the wiretap channel are two discrete memoryless channels (DMCs), and they are connected with the legitimate receiver and the wiretapper, respectively. The transition probability distribution of the main channel depends on the channel state. Measuring wiretapper's uncertainty about the message by equivocation, the capacity equivocation regions are provided both for the case where the channel inputs are allowed to depend noncausally on the state sequence and the case where they are restricted to causal dependence. Furthermore, the secrecy capacities for both cases are formulated, which provide the best transmission rate with perfect secrecy. The result is further explained via a binary example.

1. Introduction

Equivocation was first introduced into channel coding by Wyner in his study of wiretap channel [1], see Figure 1. It is a kind of degraded broadcast channels. The object is to transmit messages to the legitimate receiver, while keeping the wiretapper as ignorant of the messages as possible.

After the publication of Wyner's work, Csiszár and Körner [2] investigated a more general situation: the broadcast channels with confidential messages, see Figure 2. The model of [2] is to transmit confidential messages to receiver 1 at rate R_1 and common messages to both receivers at rate R_0, while keeping receiver 2 as ignorant of the confidential messages as possible. Measuring ignorance by equivocation, a single-letter characterization of all the achievable triples (R_1, R_e, R_0) was provided in [2], where R_e is the second receiver's equivocation to the confidential messages. Note that the model of [2] is also a generalization of [3], where no confidentiality condition is imposed. In addition, Merhav [4] studied a specified wiretap channel and obtained the capacity region, where both the legitimate receiver and the wiretapper have access to some leaked symbols from the source, but the channels for the wiretapper are more noisy than the legitimate receiver, which shares a secret key with the encoder.

In communication systems there is often a feedback link from the receiver to the transmitter, for example, the two-way channels for telephone connections. It is well known that feedback does not increase the capacity of discrete memoryless channel (DMC). However, does the feedback increase the capacity region of the wiretap channel? In order to solve this problem, Ahlswede and Cai studied the general wiretap channel (the wiretap channel does not need to be degraded) with noiseless feedback from the legitimate receiver to the channel encoder [5] (see Figure 3), and both upper and lower bounds of the secrecy capacity were provided. Specifically, for the degraded case, they showed that the secrecy capacity is larger than that of Wyner's wiretap channel (without feedback). In the achievability proof, Ahlswede and Cai [5] used the noiseless feedback as a secret key shared by the transmitter and the legitimate receiver, while the wiretapper had no

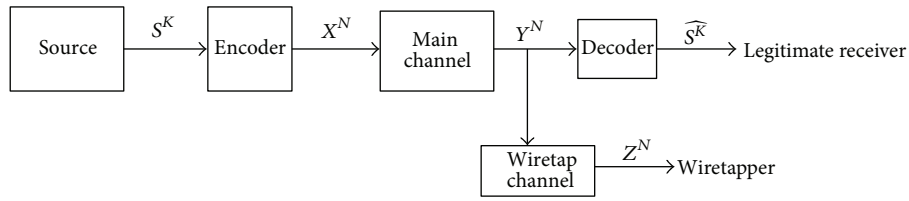

Figure 1: Wiretap channel.

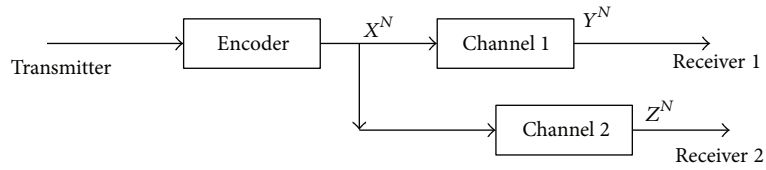

Figure 2: Broadcast channels with confidential messages.

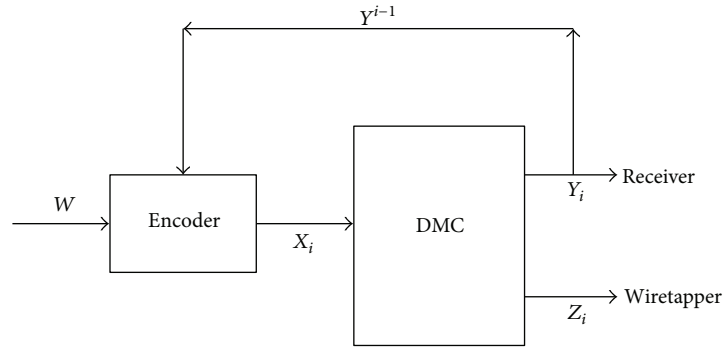

Figure 3: The general wiretap channel with noiseless feedback.

additional knowledge about the key except his own received symbols. Based on the work of [5], Dai et al. [6] studied a special case of the general wiretap channel with noiseless feedback and found that the noiseless feedback enhances the secrecy capacity of the nondegraded wiretap channel. Besides Ahlswede and Cai's work, the wiretap channel with noisy feedback was studied in [7], and the wiretap channel with secure rate-limited feedback was studied in [8], and both of them focused on bounds of the secrecy capacity. Since the feedback in the model of wiretap channel is often used as a shared secret key, the techniques used in the secret sharing scheme play an important role in the construction of the practical secure communication systems, see [9–11].

Communication through state-dependent channels, with states known at the transmitter, was first investigated by Shannon [12] in 1958. In [12], the capacity of the discrete memoryless channel with causal (past and current) channel state information at the encoder was totally determined. After that, in order to solve the problem of coding for a computer memory with defective cells, Kuznecov and Cybakov [13] considered a channel in the presence of noncausal channel state information at the transmitter. They provided some coding techniques without the determination of the capacity. The capacity was found in 1980 by Gel'efand and Pinsker [14]. Furthermore, Costa [15] investigated a power-constrained additive noise channel, where part of the noise is known at the transmitter as side information. This channel is also called dirty paper channel. The assumption in these seminar papers, as well as in the work on communication with state-dependent channels that followed, is that the channel states are generated by nature and cannot be affected or controlled by the communication system.

In 2010, Weissman [16] revisited the above problem setting for the case where the transmitter can take actions that affect the formation of the states, see Figure 4. Specifically, Weissman considered a communication system where encoding is in two parts: given the message, an action sequence is created. The actions affect the formation of the channel states, which are accessible to the transmitter when producing the channel input sequence. The capacity of this model is totally determined both for the case where the channel inputs are allowed to depend noncausally on the state sequence and the case where they are restricted to causal dependence. Meanwhile, Weissman [16] found that the feedback from the channel output to the channel encoder cannot increase the channel capacity. This framework captures

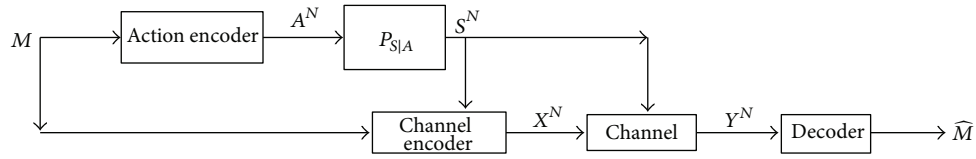

Figure 4: Channel with action-dependent states.

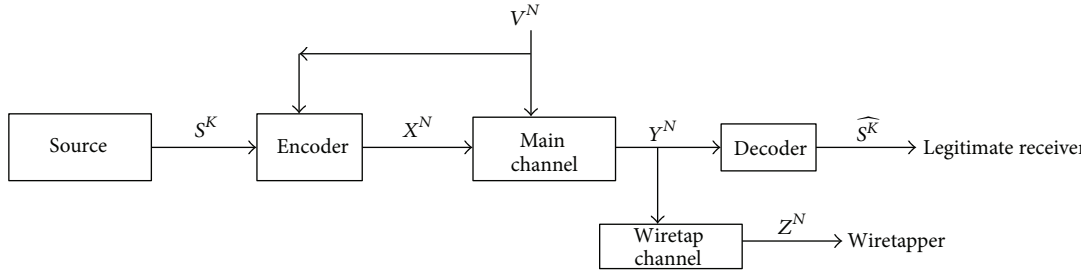

Figure 5: Wiretap channel with noncausal channel state information.

various new channel coding scenarios that may arise naturally in recording for magnetic storage devices or coding for computer memories with defects.

Inspired by the above works, Mitrpant et al. [17] studied the transmission of confidential messages through the channels with channel state information (CSI). In [17], an inner bound on the capacity-equivocation region was provided for the Gaussian wiretap channel with CSI. Furthermore, Chen and Vinck [18] investigated the discrete memoryless wiretap channel with noncausal CSI (see Figure 5) and also provided an inner bound on the capacity-equivocation region. Note that the coding scheme of [18] is a combination of those in [1, 14]. Based on the work of [18], Dai and Luo [19] provided an outer bound on the wiretap channel with noncausal CSI and determined the capacity-equivocation region for the memoryless case.

In this paper, we study the wiretap channel in the presence of action-dependent states and noiseless feedback, see Figure 6. This work is inspired by the wiretap channel with CSI [18], the channel with action-dependent CSI [16], and the wiretap channel with noiseless feedback [5]. The motivation of this work is to investigate the transmission of confidential messages through the channel with action-dependent CSI and noiseless feedback.

More concretely, in Figure 6, the transmitted message W is encoded as an action sequence A^N, and A^N is the input of a discrete memoryless channel (DMC). The output of this DMC is the channel state sequence S^N. The main channel is a DMC with inputs X^N and S^N and output Y^N. The wiretap channel is also a DMC with input Y^N and output Z^N. Moreover, there exists a noiseless feedback from the output of the main channel to the channel encoder; that is, the inputs of the channel encoder are the transmitted message W, the state sequence S^N, and the noiseless feedback, while the output is X^N. Since the action-dependent state captures various new coding scenarios for channels with a rewrite option that may arise naturally in storage for computer memories with defects or in magnetic recording, it is natural to ask the following two questions.

(i) How about the security of these channel models in the presence of a wiretapper?

(ii) What can the noiseless feedback do in the model of wiretap channel with action-dependent CSI?

Measuring wiretapper's uncertainty about the transmitted message by equivocation, the capacity-equivocation regions of the model of Figure 6 are provided both for the case where the channel input is allowed to depend noncausally on the state sequence and the case where it is restricted to causal dependence.

The contribution of this paper is as follows.

(i) Compared with the existing model of wiretap channel with side information [18] (see Figure 5), the channel state information in [18] is a special case of that in our new model; that is, the model of [18] is included in the model of Figure 6. Therefore, our result generalizes the result of [18].

(ii) Our new model also extends the model of Figure 4 by considering an additional secrecy constraint, and therefore our result also generalizes the result of [16].

In this paper, random variables, sample values, and alphabets are denoted by capital letters, lowercase, letters and calligraphic letters, respectively. A similar convention is applied to the random vectors and their sample values. *For example, U^N denotes a random N-vector $(U_1, \ldots, U_N)$, and $u^N = (u_1, \ldots, u_N)$ is a specific vector value in $\mathcal{U}^N$, that is the, Nth Cartesian power of $\mathcal{U}$. U_i^N denotes a random $N - i + 1$-vector $(U_i, \ldots, U_N)$, and $u_i^N = (u_i, \ldots, u_N)$ is a specific vector value in $\mathcal{U}_i^N$.* Let $p_V(v)$ denote the probability mass function $\Pr\{V = v\}$. Throughout the paper, the logarithmic function is taken to the base 2.

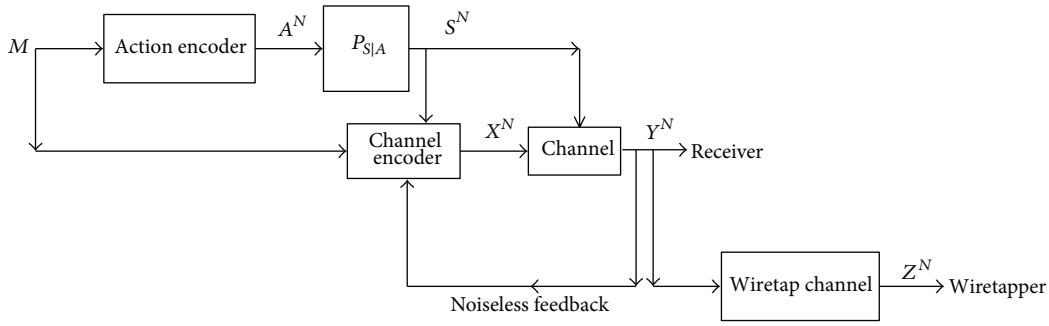

Figure 6: Wiretap channel with action-dependent channel state information and noiseless feedback.

The remainder of this paper is organized as follows. In Section 2, we present the basic definitions and the main result on the capacity-equivocation region of wiretap channel with action-dependent channel state information and noiseless feedback. In Section 3, we provide a binary example of the model of Figure 6. Final conclusions are presented in Section 4.

2. Notations, Definitions, and the Main Results

In this section, the model of Figure 6 is considered into two parts. The model of Figure 6 with noncausal channel state information is described in Section 2.1, and the causal case is described in Section 2.2, see the following.

2.1. The Model of Figure 6 with Noncausal Channel State Information. In this section, a description of the wiretap channel with noncausal action-dependent channel state information is given by Definition 1 to Definition 5. The capacity-equivocation region $\mathscr{C}^n$ composed of all achievable (R, R_e) pairs is given in Theorem 6, where the achievable (R, R_e) pair is defined in Definition 5.

Definition 1 (action encoder). The message W takes values in $\mathscr{W}$, and it is uniformly distributed over its range. The action encoder is a deterministic mapping:

$$f_1^N : \mathscr{W} \longrightarrow \mathscr{A}^N. \tag{1}$$

The input of the action encoder is W, while the output is A^N.

The channel state sequence S^N is generated by a DMC with input A^N and output S^N. The transition probability distribution is given by

$$p_{S^N|A^N}\left(s^N \mid a^N\right) = \prod_{i=1}^{N} p_{S_i|A_i}\left(s_i \mid a_i\right). \tag{2}$$

Note that the components of the state sequence S^N may not be i.i.d. random variables, and this is due to the fact that A^N is not i.i.d. generated. The transmission rate of the message is $\log \|\mathscr{W}\|/N$.

Definition 2 (channel encoder and the main channel). The main channel is a DMC with finite input alphabet $\mathscr{X} \times \mathscr{S}$, finite output alphabet $\mathscr{Y}$, and transition probability $Q_M(y \mid x, s)$, where $x \in \mathscr{X}, s \in \mathscr{S}, y \in \mathscr{Y}$. $Q_M(y^N \mid x^N, s^N) = \prod_{n=1}^{N} Q_M(y_n \mid x_n, s_n)$. The inputs of the main channel are X^N and S^N, while the output is Y^N.

There is a noiseless feedback from the output of the main channel to the channel encoder. At the ith time, the feedback Y^{i-1} (where $2 \le i \le N$ and Y^{i-1} takes values in $\mathscr{Y}^{i-1}$) is the previous $i-1$ time output of the main channel. Since the channel encoder knows the state sequence s^N in a noncausal manner, at the ith time, the inputs of the channel encoder are W, Y^{i-1}, and S^N, while the output is X_i; that is, the ith time channel encoder is a conditional probability $f_{2,i}^N(x_i \mid w, s^N, y^{i-1})$ that the message w, the feedback y^{i-1}, and the channel state sequence s^N are encoded as the ith time channel input x_i.

Definition 3 (wiretap channel). The wiretap channel is also a DMC with finite input alphabet $\mathscr{Y}$, finite output alphabet $\mathscr{Z}$, and transition probability $Q_W(z \mid y)$, where $y \in \mathscr{Y}, z \in \mathscr{Z}$. The input and output of the wiretap channel are Y^N and Z^N, respectively. The equivocation to the wiretapper is defined as

$$\Delta = \frac{H\left(W \mid Z^N\right)}{N}. \tag{3}$$

The cascade of the main channel and the wiretap channel is another DMC with transition probability:

$$Q_{MW}(z \mid x, s) = \sum_{y \in \mathscr{Y}} Q_W(z \mid y)\, Q_M(y \mid x, s). \tag{4}$$

Note that $W \to A^N \to (X^N, S^N) \to Y^N \to Z^N$ is a Markov chain in the model of Figure 6.

Definition 4 (decoder). The decoder for the legitimate receiver is a mapping $f_D : \mathscr{Y}^N \to \mathscr{W}$, with input Y^N and output $\widehat{W}$. Let P_e be the error probability of the receiver, and it is defined as $\Pr\{W \neq \widehat{W}\}$.

Definition 5 (achievable (R, R_e) pair in the model of Figure 6). A pair (R, R_e) (where $R, R_e > 0$) is called achievable if, for any

$\epsilon > 0$ (where ϵ is an arbitrary small positive real number and $\epsilon \to 0$), there exist channel encoders decoders (N, P_e) such that

$$\lim_{N\to\infty} \frac{\log \|\mathscr{W}\|}{N} = R, \qquad \lim_{N\to\infty} \Delta \geq R_e, \quad P_e \leq \epsilon. \tag{5}$$

The capacity-equivocation region $\mathscr{R}^n$ is a set composed of all achievable (R, R_e) pairs, and it is characterized by the following Theorem 6. The proof of Theorem 6 is in Appendices A and B.

Theorem 6. *A single-letter characterization of the region $\mathscr{R}^n$ is as follows:*

$$\begin{aligned} \mathscr{R}^{(n)} = \{&(R, R_e) : 0 \leq R_e \leq R, \\ &R \leq I(U;Y) - I(U;S \mid A), \\ &R_e \leq H(Y \mid Z), \\ &R_e \leq H(A \mid Z)\}, \end{aligned} \tag{6}$$

where $p_{UASXYZ}(u,a,s,x,y,z) = p_{Z|Y}(z \mid y) p_{Y|X,S}(y \mid x,s) p_{UAXS}(u,a,x,s)$, which implies that $(A,U) \to (X,S) \to Y \to Z$.

Remark 7. There are some notes on Theorem 6, see the following.

(i) The region $\mathscr{R}^{(n)}$ is convex, and the proof is directly obtained by introducing a time-sharing random variable into Theorem 6, and, therefore, we omit the proof here.

(ii) The range of the random variable U satisfies

$$\|\mathscr{U}\| \leq \|\mathscr{X}\| \|\mathscr{A}\| \|\mathscr{S}\| + 1. \tag{7}$$

The proof is in Appendix C.

(iii) Without the equivocation parameter, the capacity of the main channel with feedback is given by

$$C_M = \max_{p_{X|U,S}(x|u,s) p_{U|A,S}(u|a,s) p_A(a)} (I(U;Y) - I(U;S \mid A)). \tag{8}$$

The formula (8) is proved by Weissman [16], and it is omitted here.

(iv) *Secrecy capacity*: the points in $\mathscr{R}^{(n)}$ for which $R_e = R$ are of considerable interest, which imply the perfect secrecy $H(W) = H(W \mid Z^N)$.

Definition 8 (the secrecy capacity $C_s^{(n)}$). The secrecy capacity $C_s^{(n)}$ of the model of Figure 6 with noncausal channel state information is denoted by

$$C_s^{(n)} = \max_{(R,R_e=R)\in\mathscr{R}^{(n)}} R. \tag{9}$$

Clearly, we can easily determine the secrecy capacity C_s^n of the model of Figure 6 with noncausal channel state information by

$$\begin{aligned} C_s^n = \max \min \{&I(U;Y) - I(U;S \mid A), \\ &H(Y \mid Z), H(A \mid Z)\}. \end{aligned} \tag{10}$$

Proof of (10). Substituting $R_e = R$ into the region $\mathscr{R}^{(n)}$ in Theorem 6, we have

$$\begin{aligned} R &\leq H(Y \mid Z), \\ R &\leq I(U;Y) - I(U;S \mid A), \\ R &\leq H(A \mid Z). \end{aligned} \tag{11}$$

Therefore, the secrecy capacity $C_s^{(n)} = \max_{(R,R_e=R)\in\mathscr{R}^{(n)}} R = \max \min\{I(U;Y) - I(U;S \mid A), H(Y \mid Z), H(A \mid Z)\}$. Thus the proof is completed. □

2.2. The Model of Figure 6 with Causal Channel State Information. The model of Figure 6 with causal channel state information is similar to the model with noncausal channel state information in Section 2.1, except that the state sequence S^N in Definition 1 is known to the channel encoder in a causal manner, that is, at the ith time $(1 \leq i \leq N)$, the output of the encoder $x_i = f_{2,i}(w, s^i, y^{i-1})$, where $s^i = (s_1, s_2, \ldots, s_i)$ and $f_{2,i}$ is the probability that the message w, the feedback y^{i-1}, and the state s^i are encoded as the channel input x_i at time i.

The capacity-equivocation region $\mathscr{R}^c$ for the model of Figure 6 with causal channel state information is characterized by the following Theorem 9, and it is proved in Appendices D and E.

Theorem 9. *A single-letter characterization of the region $\mathscr{R}^c$ is as follows:*

$$\begin{aligned} \mathscr{R}^{(c)} = \{&(R, R_e) : 0 \leq R_e \leq R, \\ &R \leq I(U;Y), \\ &R_e \leq H(Y \mid Z), \\ &R_e \leq H(A \mid Z)\}, \end{aligned} \tag{12}$$

where $p_{UASXYZ}(u,a,s,x,y,z) = p_{Z|Y}(z \mid y) p_{Y|X,S}(y \mid x,s) p_{UAXS}(u,a,x,s)$, which implies that $(A,U) \to (X,S) \to Y \to Z$.

Remark 10. There are some notes on Theorem 9, see the following.

(i) The region $\mathscr{R}^{(c)}$ is convex.

(ii) The range of the random variable U satisfies

$$\|\mathscr{U}\| \leq \|\mathscr{X}\| \|\mathscr{A}\| \|\mathscr{S}\|. \tag{13}$$

The proof is similar to that in Theorem 6, and it is omitted here.

(iii) Without the equivocation parameter, the capacity of the main channel with feedback is given by

$$C_M^* = \max_{p_{X|U,S}(x|u,s)p_{U|A,S}(u|a,s)p_A(a)} I(U;Y). \tag{14}$$

The formula (14) is proved by Weissman [16], and it is omitted here.

(iv) *Secrecy capacity*: the points in $\mathscr{R}^{(c)}$ for which $R_e = R$ are of considerable interest, which imply the perfect secrecy $H(W) = H(W \mid Z^N)$.

Definition 11 (the secrecy capacity $C_s^{(c)}$). The secrecy capacity $C_s^{(c)}$ of the model of Figure 6 with causal channel state information is denoted by

$$C_s^{(c)} = \max_{(R,R_e=R)\in\mathscr{R}^{(c)}} R. \tag{15}$$

Clearly, we can easily determine the secrecy capacity $C_s^{(c)}$ of the model of Figure 6 with causal channel state information by

$$C_s^{(c)} = \max \min\{I(U;Y), H(Y \mid Z), H(A \mid Z)\}. \tag{16}$$

Proof of (16). Substituting $R_e = R$ into the region $\mathscr{R}^{(c)}$ in Theorem 9, we have

$$\begin{aligned} R &\le H(Y \mid Z), \\ R &\le I(U;Y), \\ R &\le H(A \mid Z). \end{aligned} \tag{17}$$

Therefore, the secrecy capacity $C_s^{(c)} = \max_{(R,R_e=R)\in\mathscr{R}^{(c)}} R = \max \min\{I(U;Y), H(Y \mid Z), H(A \mid Z)\}$. Thus the proof is completed. □

3. A Binary Example for the Model of Figure 6 with Causal Channel State Information

In this section, we calculate the secrecy capacity of a special case of the model of Figure 6 with causal channel state information.

Suppose that the channel state information S^N is available at the channel encoder in a casual manner. Meanwhile, the random variables U, A, X, S, Y, and Z take values in $\{0,1\}$, and the transition probability of the main channel is defined as follows.

When $s = 0$,

$$p_{Y|X,S}(y \mid x, s=0) = \begin{cases} 1-p, & \text{if } y = x, \\ p, & \text{otherwise.} \end{cases} \tag{18}$$

When $s = 1$,

$$p_{Y|X,S}(y \mid x, s=1) = \begin{cases} p, & \text{if } y = x, \\ 1-p, & \text{otherwise.} \end{cases} \tag{19}$$

Note that here $0 \le p \le 1$.

The wiretap channel is a (binary symmetric channel) BSC with crossover probability q $(0 \le q \le 1/2)$, that is,

$$p_{Z|Y}(z \mid y) = \begin{cases} 1-q, & \text{if } y = x, \\ q, & \text{otherwise.} \end{cases} \tag{20}$$

The channel for generating the state sequence S^N is a BSC with crossover probability r $(0 \le r \le 1)$, that is,

$$p_{S|A}(s \mid a) = \begin{cases} 1-r, & \text{if } y = x, \\ r, & \text{otherwise.} \end{cases} \tag{21}$$

From Remark 10 we know that the secrecy capacity for the causal case is given by

$$C_s^c = \max \min\{I(U;Y), H(Y \mid Z), H(A \mid Z)\}. \tag{22}$$

Note that $\max I(U;Y)$, $\max H(A \mid Z)$, and $\max H(Y \mid Z)$ are achieved if A is a function of U and X is a function of U and S, and this is similar to the argument in [16]. Define $a = g(u)$ and $x = f(u,s)$, then (22) can be written as

$$C_s^c = \max_{f,g,p_U(u)} \min\{I(U;Y), H(A \mid Z), H(Y \mid Z)\}, \tag{23}$$

and this is because the joint probability distribution $p_{AUSXYZ}(a,u,s,x,y,z)$ can be calculated by

$$\begin{aligned} &p_{AUSXYZ}(a,u,s,x,y,z) \\ &\quad = p_{Z|Y}(z \mid y)\, p_{Y|X,S}(y \mid x,s)\, 1_{x=f(u,s)} p_{S|A} \\ &\qquad \times (s \mid a)\, 1_{a=g(u)} p_U(u). \end{aligned} \tag{24}$$

Since A is a function of U, we have

$$H(A \mid Z) = H(U \mid Z). \tag{25}$$

Then, it is easy to see that

$$\begin{aligned} &\max_{f,g,p_U(u)} \min\{I(U;Y), H(A \mid Z), H(Y \mid Z)\} \\ &\quad = \max_{f,g,p_U(u)} \max_{f,g,p_U(u)} \min\{I(U;Y), H(U \mid Z), H(Y \mid Z)\}. \end{aligned} \tag{26}$$

Now it remains to calculate the characters $\max_{f,g,p_U(u)} H(Y \mid Z)$, $\max_{f,g,p_U(u)} H(U \mid Z)$, and $\max_{f,g,p_U(u)} I(U;Y)$; see the remaining of this section.

Let U take values in $\{0,1\}$. The probability of U is defined as follows: $p_U(0) = \alpha$ and $p_U(1) = 1-\alpha$.

In addition, there are 16 kinds of f and 4 kinds of g. Define the following:

$$f^{(1)}(u,s): \begin{cases} 00 \longrightarrow 0,\ 01 \longrightarrow 0, \\ 10 \longrightarrow 0,\ 11 \longrightarrow 0, \end{cases}$$

$$f^{(2)}(u,s): \begin{cases} 00 \longrightarrow 0,\ 01 \longrightarrow 0, \\ 10 \longrightarrow 0,\ 11 \longrightarrow 1, \end{cases}$$

$$f^{(3)}(u,s): \begin{cases} 00 \longrightarrow 0,\ 01 \longrightarrow 0, \\ 10 \longrightarrow 1,\ 11 \longrightarrow 0, \end{cases}$$

$$f^{(4)}(u,s): \begin{cases} 00 \longrightarrow 0,\ 01 \longrightarrow 0, \\ 10 \longrightarrow 1,\ 11 \longrightarrow 1, \end{cases}$$

$$f^{(5)}(u,s): \begin{cases} 00 \longrightarrow 0,\ 01 \longrightarrow 1, \\ 10 \longrightarrow 0,\ 11 \longrightarrow 0, \end{cases}$$

$$f^{(6)}(u,s): \begin{cases} 00 \longrightarrow 0,\ 01 \longrightarrow 1, \\ 10 \longrightarrow 0,\ 11 \longrightarrow 1, \end{cases}$$

$$f^{(7)}(u,s): \begin{cases} 00 \longrightarrow 0,\ 01 \longrightarrow 1, \\ 10 \longrightarrow 1,\ 11 \longrightarrow 0, \end{cases}$$

$$f^{(8)}(u,s): \begin{cases} 00 \longrightarrow 0,\ 01 \longrightarrow 1, \\ 10 \longrightarrow 1,\ 11 \longrightarrow 1, \end{cases}$$

$$f^{(9)}(u,s): \begin{cases} 00 \longrightarrow 1,\ 01 \longrightarrow 0, \\ 10 \longrightarrow 0,\ 11 \longrightarrow 0, \end{cases}$$

$$f^{(10)}(u,s): \begin{cases} 00 \longrightarrow 1,\ 01 \longrightarrow 0, \\ 10 \longrightarrow 0,\ 11 \longrightarrow 1, \end{cases} \tag{27}$$

$$f^{(11)}(u,s): \begin{cases} 00 \longrightarrow 1,\ 01 \longrightarrow 0, \\ 10 \longrightarrow 1,\ 11 \longrightarrow 0, \end{cases}$$

$$f^{(12)}(u,s): \begin{cases} 00 \longrightarrow 1,\ 01 \longrightarrow 0, \\ 10 \longrightarrow 1,\ 11 \longrightarrow 1, \end{cases}$$

$$f^{(13)}(u,s): \begin{cases} 00 \longrightarrow 1,\ 01 \longrightarrow 1, \\ 10 \longrightarrow 0,\ 11 \longrightarrow 0, \end{cases}$$

$$f^{(14)}(u,s): \begin{cases} 00 \longrightarrow 1,\ 01 \longrightarrow 1, \\ 10 \longrightarrow 0,\ 11 \longrightarrow 1, \end{cases}$$

$$f^{(15)}(u,s): \begin{cases} 00 \longrightarrow 1,\ 01 \longrightarrow 1, \\ 10 \longrightarrow 1,\ 11 \longrightarrow 0, \end{cases}$$

$$f^{(16)}(u,s): \begin{cases} 00 \longrightarrow 1,\ 01 \longrightarrow 1, \\ 10 \longrightarrow 1,\ 11 \longrightarrow 1, \end{cases}$$

$$g^{(1)}(u): \begin{cases} 0 \longrightarrow 0, \\ 1 \longrightarrow 0, \end{cases} \qquad g^{(2)}(u): \begin{cases} 0 \longrightarrow 0, \\ 1 \longrightarrow 1, \end{cases}$$

$$g^{(3)}(u): \begin{cases} 0 \longrightarrow 1, \\ 1 \longrightarrow 0, \end{cases} \qquad g^{(4)}(u): \begin{cases} 0 \longrightarrow 1, \\ 1 \longrightarrow 1. \end{cases}$$

The character $I(U;Y)$ depends on the joint probability mass functions $p_{UY}(u,y)$, and we have

$$\begin{aligned} p_{UY}(u,y) &= \sum_{x,s,a} p_{UYXSA}(u,y,x,s,a) \\ &= \sum_{x,s,a} p_{Y|XS}(y \mid x,s)\, p_{X|U,S}(x \mid u,s)\, p_U(u) \\ &\quad \times p_{A|U}(a \mid u)\, p_{S|A}(s \mid a). \end{aligned} \tag{28}$$

The character $H(U \mid Z)$ depends on the joint probability mass functions $p_{UZ}(u,z)$, and we have

$$\begin{aligned} p_{UZ}(u,z) &= \sum_{y} p_{UYZ}(u,y,z) \\ &= \sum_{y} p_{Z|Y}(z \mid y)\, p_{U,Y}(u,y). \end{aligned} \tag{29}$$

The character $H(Y \mid Z)$ depends on the joint probability mass functions $p_{YZ}(u,z)$, and we have

$$\begin{aligned} p_{YZ}(y,z) &= \sum_{u} p_{UYZ}(u,y,z) \\ &=^{(a)} \sum_{u} p_{Z|Y}(z \mid y)\, p_{U,Y}(u,y), \end{aligned} \tag{30}$$

where (a) is from $U \to Y \to Z$.

By choosing the above f, g, and α, we find that

$$\max_{f,g,p_U(u)} H(U \mid Z) = h(p+q-2pq), \tag{31}$$

where $h(p+q-2pq) = (p+q-2pq)\log(1/(p+q-2pq)) + (1-p-q+2pq)\log(1/(1-p-q+2pq))$. Moreover, $h(p+q-2pq)$ is achieved when $f = f^{(7)}$, $g = g^{(2)}$ and $\alpha = 1/2$.

On the other hand,

$$\max_{f,g,p_U(u)} I(U;Y) = 1 - h(p), \tag{32}$$

and "=" is achieved if $f = f^{(7)}$, $g = g^{(2)}$ and $\alpha = 1/2$.

Moreover,

$$\max_{f,g,p_U(u)} H(Y \mid Z) = h(q), \tag{33}$$

and "=" is achieved if $f = f^{(7)}$, $g = g^{(2)}$ and $\alpha = 1/2$.

Therefore, the secrecy capacity for the causal case is given by

$$C_s^c = \min\{h(p+q-2pq), 1-h(p), h(q)\}. \tag{34}$$

Figures 7, 8, and 9 give the secrecy capacity of the model of Figure 6 with causal channel state information for several values of q. It is easy to see that the secrecy capacity C_s^c is increasing while q is getting larger.

4. Conclusion

In this paper, we study the model of the wiretap channel with action-dependent channel state information and noiseless

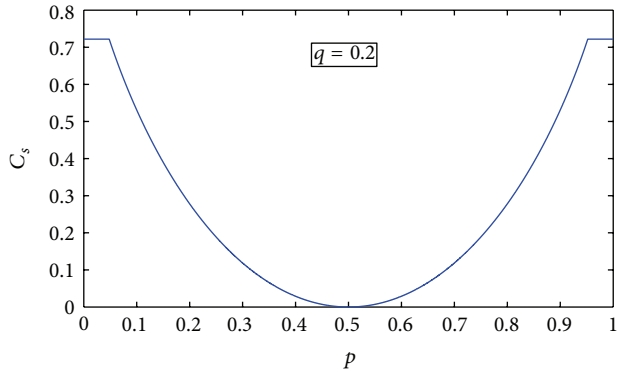

FIGURE 7: When $q = 0.2$, the secrecy capacity of the model of Figure 6 is with causal channel state information.

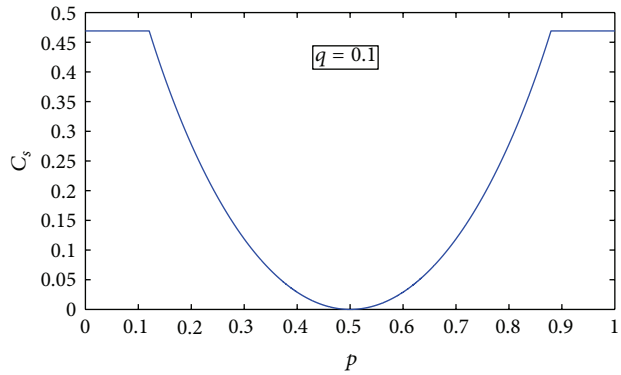

FIGURE 8: When $q = 0.1$, the secrecy capacity of the model of Figure 6 is with causal channel state information.

feedback. The capacity-equivocation regions are provided both for the case where the channel inputs are allowed to depend noncausally on the state sequence and the case where they are restricted to causal dependence. Furthermore, the secrecy capacities for both cases are formulated, which provide the best transmission rate with perfect secrecy. The result is further explained via a binary example.

The contribution of this paper is as follows.

(i) Compared with the existing model of wiretap channel with side information [18] (see Figure 5), the channel state information in [18] is a special case of that in our new model; that is, the model of [18] is included in the model of Figure 6. Therefore, our result generalizes the result of [18].

(ii) Our new model also extends the model of Figure 4 by considering an additional secrecy constraint, and therefore our result also generalizes the result of [16].

Appendices

A. Proof of the Direct Part of Theorem 6

In this section, we will show that any pair $(R, R_e) \in \mathscr{R}^n$ is achievable. Gel'efand-Pinsker's binning, block Markov

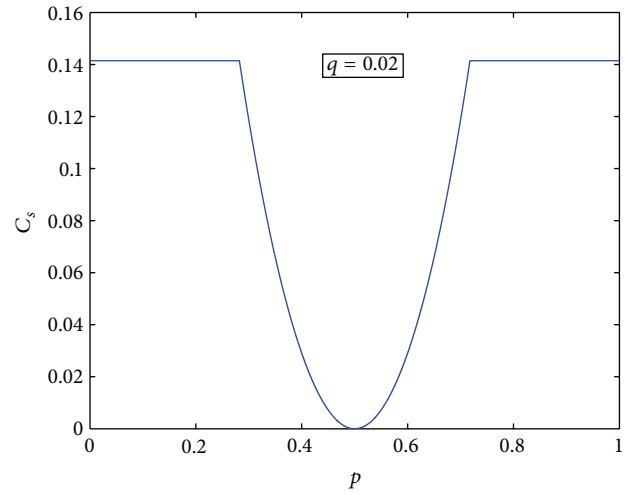

FIGURE 9: When $q = 0.02$, the secrecy capacity of the model of Figure 6 is with causal channel state information.

coding, and Ahlswede-Cai's secret key on the feedback system are used in the construction of the code book.

Now the remainder of this section is organized as follows. The code construction is in Appendix A.1. The proof of achievability is given in Appendix A.2.

A.1. Code Construction. Since $R_e \le H(Y \mid Z)$, $R_e \le H(A \mid Z)$ and $R_e \le R \le I(U;Y) - I(U;S \mid A)$, it is sufficient to show that the pair $(R, R_e = \min\{H(Y \mid Z), H(A \mid Z), I(U;Y) - I(U;S \mid A)\}$ is achievable, and note that this implies that $R \ge R_e = \min\{H(Y \mid Z), H(A \mid Z), I(U;Y) - I(U;S \mid A)\}$.

The construction of the code and the proof of achievability are considered into two cases.

(i) *Case 1*: $I(U;Y) - I(U;S \mid A) \ge \min\{H(Y \mid Z), H(A \mid Z)\}$.

(ii) *Case 2*: $I(U;Y) - I(U;S \mid A) \le \min\{H(Y \mid Z), H(A \mid Z)\}$.

We use the block Markov coding method. The random vectors U^N, A^N, S^N, X^N, Y^N, and Z^N consist of n blocks of length N. The message for n blocks is $W^n \triangleq (W_2, W_3, \ldots, W_n)$, where W_i $(2 \le i \le n)$ are i.i.d. random variables uniformly distributed over $\mathscr{W}$. Note that in the first block, there is no W_1.

Let $\widetilde{Z}_i$ $(1 \le i \le n)$ be the output of the wiretap channel for block i, $Z^n = (\widetilde{Z}_1, \ldots, \widetilde{Z}_n)$, $Z^{\tilde{j}} = (\widetilde{Z}_1, \ldots, \widetilde{Z}_{j-1}, \widetilde{Z}_{j+1}, \ldots, \widetilde{Z}_n)$ $(1 \le j \le n)$. Similarly, $Y^n = (\widetilde{Y}_1, \ldots, \widetilde{Y}_n)$, and $\widetilde{Y}_i$ $(1 \le i \le n)$ is the output of the main channel for block i. The specific values of the above random vectors are denoted by lowercase letters.

(i) Code Construction for Case 1. Given a pair (R, R_e), choose a joint probability mass function $p_{U,A,S,X,Y,Z}(u, a, s, x, y, z)$

such that

$$\begin{aligned} 0 \le R_e \le R, \\ R \le I(U;Y) - I(U;S \mid A), \\ R_e = \min\{H(Y \mid Z), H(A \mid Z)\}. \end{aligned} \tag{A.1}$$

The message set $\mathscr{W}$ satisfies the following condition:

$$\begin{aligned} \lim_{N\to\infty} \frac{1}{N} \log \|\mathscr{W}\| \\ = R = I(U;Y) - I(U;S \mid A) - \gamma, \end{aligned} \tag{A.2}$$

where γ is a fixed positive real numbers and

$$\begin{aligned} 0 \le \gamma \le^{(a)} I(U;Y) - I(U;S \mid A) \\ - \min\{H(Y \mid Z), H(A \mid Z)\}. \end{aligned} \tag{A.3}$$

Note that (a) is from $R \ge R_e = \min\{H(Y \mid Z), H(A \mid Z)\}$ and (A.2). Let $\mathscr{W} = \{1, 2, \ldots, 2^{NR}\}$.

Code-book generation:

(a) Construction of A^N and S^N. In the first block, generate a i.i.d. sequence a^N according to the probability mass function $p_A(a)$, and choose it as the output of the action encoder. Let s^N be the state sequence generated in response to the chosen action sequence a^N.

For the ith block $(2 \le i \le n)$, generate 2^{NR} i.i.d. sequences a^N, according to the probability mass function $p_A(a)$. Index each sequence by $i \in \{1, 2, \ldots, 2^{NR}\}$. For a given message w_i $(w_i \in \mathscr{W})$, choose a corresponding $a^N(w_i)$ as the output of the action encoder. Let s^N be the state sequence generated in response to the action sequence $a^N(w_i)$.

(b) Construction of the Secret Key. For the ith block $(2 \le i \le n)$, firstly we generate a mapping $g_f : \mathscr{Y}^N \to \mathscr{W}$ (note that $\|\mathscr{Y}\|^N \ge \|\mathscr{W}\|$). Define a random variable $K_i = g_f(\widetilde{Y}_{i-1})$ $(2 \le i \le n)$, which is uniformly distributed over $\mathscr{W}$, and K_i is independent of W_i. Reveal the mapping g_f to both receivers and the transmitter.

Then, when the transmitter receives the output $\widetilde{y}_{i-1}$ of the i-1th block, he computes $k_i = g_f(\widetilde{y}_{i-1}) \in \mathscr{W}$ as the secret key used in the ith block.

(c) Construction of U^N. In the first block, for the transmitted action sequence a^N and the corresponding state sequence s^N, generate a i.i.d. sequence u^N according to the probability mass function $p_{U|A,S}(u_i \mid a_i, s_i)$. Choose u^N as a realization of U^N for the first block.

For the ith block $(2 \le i \le n)$, given the transmitted action sequence $a^N(w_i)$ and the corresponding state sequence s^N, generate $2^{N(I(U;Y)-\epsilon_{2,N})}$ $(\epsilon_{2,N} \to 0$ as $N \to \infty)$ i.i.d. sequences u^N, according to the probability mass function $p_{U|A,S}(u_i \mid a_i(w_i), s_i)$. Distribute these sequences at random into $2^{NR} = 2^{N(I(U;Y)-I(U;S|A)-\gamma)}$ bins such that each bin contains $2^{N(I(U;S|A)+\gamma-\epsilon_{2,N})}$ sequences. Index each bin by $i \in \{1, 2, \ldots, 2^{NR}\}$.

For a given w_i, $a^N(w_i)$, s^N, and a secret key k_i, the transmitter chooses a sequence $u^N(w_i \oplus k_i, j^*)$ from the bin $w_i \oplus k_i$ (where $\oplus$ is the modulo addition over $\mathscr{W}$) such that $(u^N(w_i \oplus k_i, j^*), a^N(w_i), s^N) \in T^N_{US|A}(\epsilon)$. If such multiple sequences in bin $w_i \oplus k_i$ exist, choose the one with the smallest index in the bin. If no such sequence exists, declare an encoding error.

(d) Construction of X^N. For each block, the x^N is generated according to a new discrete memoryless channel (DMC) with inputs u^N, s^N, and output x^N. The transition probability of this new DMC is $p_{X|U,S}(x \mid u, s)$, which is obtained from the joint probability mass function $p_{U,A,S,X,Y,Z}(u, a, s, x, y, z)$.

The probability $p_{X^N|U^N,S^N}(x^N \mid u^N, s^N)$ is calculated as follows:

$$\&p_{X^N|U^N,S^N}\left(x^N \mid u^N, s^N\right) = \prod_{i=1}^{N} p_{X|U,S}\left(x_i \mid u_i, s_i\right). \tag{A.4}$$

Decoding. For the ith block $(2 \le i \le n)$, given a vector $y^N \in \mathscr{Y}^N$ and a secret key k_i (k_i is known by the receiver), try to find a sequence $u^N(\widehat{w}_i \oplus k_i, \widehat{j})$ such that $(u^N(\widehat{w}_i \oplus k_i, \widehat{j}), y^N) \in T^N_{UY}(\epsilon_3)$. If there exist sequences with the same $\widehat{w}_i \oplus k_i$, by using the secret key k_i, put out the corresponding $\widehat{w}_i$. Otherwise, that is, if no such sequence exists or multiple sequences have different message indices, declare a decoding error.

(ii) Code Construction for Case 2. Given a pair (R, R_e), choose a joint probability mass function $p_{U,A,S,X,Y,Z}(u, a, s, x, y, z)$ such that

$$\begin{aligned} 0 \le R_e \le R, \\ R = I(U;Y) - I(U;S \mid A), \\ R_e = I(U;Y) - I(U;S \mid A). \end{aligned} \tag{A.5}$$

The message set $\mathscr{W}$ satisfies the following condition:

$$\begin{aligned} \lim_{N\to\infty} \frac{1}{N} \log \|\mathscr{W}\| = R \\ = I(U;Y) - I(U;S \mid A). \end{aligned} \tag{A.6}$$

Let $\mathscr{W} = \{1, 2, \ldots, 2^{NR}\}$.

Code-book generation:

(a) Construction of A^N and S^N. In the first block, generate a i.i.d. sequence a^N according to the probability mass function $p_A(a)$, and choose it as the output of the action encoder. Let s^N be the state sequence generated in response to the chosen action sequence a^N.

For the ith block $(2 \le i \le n)$, generate 2^{NR} i.i.d. sequences a^N, according to the probability mass function $p_A(a)$. Index each sequence by $i \in \{1, 2, \ldots, 2^{NR}\}$. For a given message w_i $(w_i \in \mathscr{W})$, choose a corresponding $a^N(w_i)$ as the output of

the action encoder. Let s^N be the state sequence generated in response to the action sequence $a^N(w_i)$.

(b) Construction of the Secret Key. For the ith block ($2 \le i \le n$), firstly we generate a mapping $g_f : \mathcal{Y}^N \to \mathcal{W}$ (note that $\|\mathcal{Y}\|^N \ge \|\mathcal{W}\|$). Define a random variable $K_i = g_f(\widetilde{Y}_{i-1})$ ($2 \le i \le n$), which is uniformly distributed over $\mathcal{W}$, and K_i is independent of W_i. Reveal the mapping g_f to both receivers and the transmitter.

Then, when the transmitter receives the output $\widetilde{y}_{i-1}$ of the i-1th block, he computes $k_i = g_f(\widetilde{y}_{i-1}) \in \mathcal{W}$ as the secret key used in the ith block.

(c) Construction of U^N. In the first block, for the transmitted action sequence a^N and the corresponding state sequence s^N, generate a i.i.d. sequence u^N according to the probability mass function $p_{U|A,S}(u_i \mid a_i, s_i)$. Choose u^N as a realization of U^N for the first block.

For the ith block ($2 \le i \le n$), given the transmitted action sequence $a^N(w_i)$ and the corresponding state sequence s^N, generate $2^{N(I(U;Y)-\epsilon_{2,N})}$ ($\epsilon_{2,N} \to 0$ as $N \to \infty$) i.i.d. sequences u^N, according to the probability mass function $p_{U|A,S}(u_i \mid a_i(w_i), s_i)$. Distribute these sequences at random into $2^{NR} = 2^{N(I(U;Y)-I(U;S|A))}$ bins such that each bin contains $2^{N(I(U;S|A)-\epsilon_{2,N})}$ sequences. Index each bin by $i \in \{1, 2, \ldots, 2^{NR}\}$.

For a given w_i, $a^N(w_i)$, s^N, and a secret key k_i, the transmitter chooses a sequence $u^N(w_i \oplus k_i, j^*)$ from the bin $w_i \oplus k_i$ (where $\oplus$ is the modulo addition over $\mathcal{W}$) such that $(u^N(w_i \oplus k_i, j^*), a^N(w_i), s^N) \in T^N_{US|A}(\epsilon)$. If such multiple sequences in bin $w_i \oplus k_i$ exist, choose the one with the smallest index in the bin. If no such sequence exists, declare an encoding error.

(d) Construction of X^N. The x^N is generated the same as that for the case 1, and it is omitted here.

Decoding. For the ith block ($2 \le i \le n$), given a vector $y^N \in \mathcal{Y}^N$ and a secret key k_i (k_i is known by the receiver), try to find a sequence $u^N(\widehat{w}_i \oplus k_i, \widehat{j})$ such that $(u^N(\widehat{w}_i \oplus k_i, \widehat{j}), y^N) \in T^N_{UY}(\epsilon_3)$. If there exist sequences with the same $\widehat{w}_i \oplus k_i$, by using the secret key k_i, put out the corresponding $\widehat{w}_i$. Otherwise, that is, if no such sequence exists or multiple sequences have different message indices, declare a decoding error.

A.2. Proof of Achievability. The rate of the message W^n is defined as

$$R^* \triangleq \lim_{N\to\infty} \lim_{n\to\infty} \frac{H(W^n)}{nN}, \tag{A.7}$$

and it satisfies

$$\begin{aligned} R^* &= \lim_{N\to\infty} \lim_{n\to\infty} \frac{H(W^n)}{nN} \\ &= \lim_{N\to\infty} \lim_{n\to\infty} \frac{\sum_{i=2}^{n} H(W_i)}{nN} \\ &= \lim_{N\to\infty} \lim_{n\to\infty} \frac{(n-1)H(W)}{nN} \\ &= R. \end{aligned} \tag{A.8}$$

In addition, note that the encoding and decoding scheme for Theorem 6 is exactly the same as that in [16], except that the transmitted message for the legitimate receiver is $w \oplus k$. Since the legitimate receiver knows k, the decoding scheme for Theorem 6 is in fact the same as that in [16]. Hence, we omit the proof of $P_e \le \epsilon$ here.

It remains to show that $\lim_{N\to\infty}\Delta \ge R_e$, see the following.

Since the message W is encrypted by $W \oplus K$, the equivocation about W is equivalent to the equivocation about the secret key K. There are two ways for the wiretapper to obtain the secret key k. One way is that he tries to guess the k from its alphabet $\mathcal{W}$. The other way is that he tries to guess the feedback y^N (y^N is the output of the main channel for the previous block, and $k = g_f(y^N)$) from the conditional typical set $T^N_{[Y|Z]}(\delta)$, and this is because for a given z^N and sufficiently large N, $\Pr\{(y^N \notin T^N_{[Y|Z]}(\delta)\} \to 0$. Note that there are $2^{NH(Y|Z)}$ sequences $y^N \in T^N_{[Y|Z]}(\delta)$ when $N \to \infty$ and $\delta \to 0$. Therefore, the equivocation about W is $\min\{(\log\|\mathcal{W}\|)/N = R, H(Y \mid Z)\}$, and note that $R \ge R_e$ and $H(Y \mid Z) \ge R_e$, and then $\lim_{N\to\infty}\Delta \ge R_e$ is obtained.

The details about the proof are as follows.

First, we will show that $K_i \oplus W_i$ is independent of K_i and W_i, and this is used in the proof of $\lim_{N\to\infty}\Delta \ge R_e$.

Since K_i is independent of W_i ($2 \le i \le n$), and all of them are uniformly distributed over $\mathcal{W}$, the fact that $K_i \oplus W_i$ is independent of K_i and W_i is proved by the following (A.9):

$$\begin{aligned} \Pr\{K_i \oplus W_i = a\} &= \sum_{k_i \in \mathcal{W}} \Pr\{K_i \oplus W_i = a, K_i = k_i\} \\ &= \sum_{k_i \in \mathcal{W}} \Pr\{W_i = a \ominus k_i, K_i = k_i\} \\ &= \sum_{k_i \in \mathcal{W}} \Pr\{W_i = a \ominus k_i\} \Pr\{K_i = k_i\} \\ &= \frac{1}{\|\mathcal{W}\|}, \\ \Pr\{K_i \oplus W_i = a, K_i = k_i\} &= \Pr\{W_i = a \ominus k_i, K_i = k_i\} \\ &= \Pr\{W_i = a \ominus k_i\} \Pr\{K_i = k_i\} \\ &= \frac{1}{\|\mathcal{W}\|^2}. \end{aligned} \tag{A.9}$$

Then, $\lim_{N\to\infty}\Delta \ge R_e$ for both cases is proved by the following (A.10):

$$\begin{aligned} \lim_{N\to\infty}\Delta &\triangleq \lim_{N\to\infty} \lim_{n\to\infty} \frac{H(W^n \mid Z^n)}{nN} \\ &= \lim_{N\to\infty} \lim_{n\to\infty} \frac{\sum_{i=2}^{n} H\left(W_i \mid W^{i-1}, Z^n\right)}{nN} \end{aligned}$$

$$
\begin{aligned}
&=^{(a)} \lim_{N\to\infty}\lim_{n\to\infty}\frac{\sum_{i=2}^{n} H\left(W_i \mid \widetilde{Z}_i, \widetilde{Z}_{i-1}\right)}{nN}\\
&\geq \lim_{N\to\infty}\lim_{n\to\infty}\frac{\sum_{i=2}^{n} H\left(W_i \mid \widetilde{Z}_i, \widetilde{Z}_{i-1}, W_i\oplus K_i\right)}{nN}\\
&=^{(b)} \lim_{N\to\infty}\lim_{n\to\infty}\frac{\sum_{i=2}^{n} H\left(W_i \mid \widetilde{Z}_{i-1}, W_i\oplus K_i\right)}{nN}\\
&= \lim_{N\to\infty}\lim_{n\to\infty}\frac{\sum_{i=2}^{n} H\left(K_i \mid \widetilde{Z}_{i-1}, W_i\oplus K_i\right)}{nN}\\
&=^{(c)} \lim_{N\to\infty}\lim_{n\to\infty}\frac{\sum_{i=2}^{n} H\left(K_i \mid \widetilde{Z}_{i-1}\right)}{nN}\\
&=^{(d)} \lim_{N\to\infty}\lim_{n\to\infty}\frac{\sum_{i=2}^{n} \min\left\{NH\left(Y \mid Z\right), NR\right\}}{nN}\\
&= \lim_{N\to\infty}\lim_{n\to\infty}\frac{(n-1)\min\left\{NR, NH\left(Y \mid Z\right)\right\}}{nN}\\
&= \min\left\{R, H\left(Y \mid Z\right)\right\}\\
&\geq R_e,
\end{aligned}
\tag{A.10}
$$

where (a) is from $W_i \to (\widetilde{Z}_i, \widetilde{Z}_{i-1}) \to (W^{i-1}, \widetilde{Z}^{i-2}, \widetilde{Z}_{i+1}^{n})$ that is a Markov chain, (b) is from $W_i \to (W_i \oplus K_i, \widetilde{Z}_{i-1}) \to \widetilde{Z}_i$ that is a Markov chain, (c) follows from the fact that $K_i \oplus W_i$ is independent of K_i and $\widetilde{Z}_{i-1}$, and (d) is from the fact that the wiretapper can guess the specific vector $\widetilde{Y}_{i-1}$ (corresponding to the key K_i) from the conditional typical set $T^N_{[Y|Z]}(\delta)$, and K_i is uniformly distributed over $\mathscr{W}$ (K_i is the key used in block i).

On the other hand,

$$
\begin{aligned}
\lim_{N\to\infty}\Delta &\triangleq \lim_{N\to\infty}\lim_{n\to\infty}\frac{H\left(W^n \mid Z^n\right)}{nN}\\
&=^{(1)} \lim_{N\to\infty}\lim_{n\to\infty}\frac{1}{nN}H\left(A_2^n \mid Z_2^n\right)\\
&=^{(2)} \lim_{N\to\infty}\lim_{n\to\infty}\frac{1}{nN}\left((n-1)NH\left(A \mid Z\right)\right)\\
&= H\left(A \mid Z\right) \geq R_e,
\end{aligned}
\tag{A.11}
$$

where (1) is from the fact that $W^n \triangleq (W_2, W_3, \ldots, W_n)$, $\widetilde{A}_i$ $(2 \leq i \leq n)$ is the state sequence for block i, $A_2^n = (\widetilde{A}_2, \ldots, \widetilde{A}_n)$ and $\widetilde{A}_i$ is a function of W_i, (2) is from A^n and X^n that are i.i.d. generated random vectors, and the channels are discrete memoryless.

Therefore, it is easy to see that, for the case 1, $\lim_{N\to\infty}\Delta \geq R_e$ is proved by (A.10) and (A.11). For the case 2, $\lim_{N\to\infty}\Delta \geq R_e$ is proved by the formula (A.10).

Thus, $\lim_{N\to\infty}\Delta \geq R_e$ for both cases is proved. The proof of Theorem 6 is completed.

B. Proof of the Converse Part of Theorem 6

In this section, we prove the converse part of Theorem 6: all the achievable (R, R_e) pairs are contained in the set $\mathscr{R}^{(n)}$. Suppose that (R, R_e) is achievable; that is, for any given $\epsilon > 0$, there exists a channel encoder-decoder (N, Δ, P_e) such that

$$
\lim_{N\to\infty}\frac{\log\|\mathscr{W}\|}{N} = R, \qquad \lim_{N\to\infty}\Delta \geq R_e, \quad P_e \leq \epsilon. \tag{B.1}
$$

Then we will show the existence of random variables $(A, U) \to (X, S) \to Y \to Z$ such that

$$
0 \leq R_e \leq R, \tag{B.2}
$$

$$
R \leq I\left(U; Y\right) - I\left(U; S \mid A\right), \tag{B.3}
$$

$$
R_e \leq H\left(Y \mid Z\right), \tag{B.4}
$$

$$
R_e \leq H\left(A \mid Z\right). \tag{B.5}
$$

Since W is uniformly distributed over $\mathscr{W}$, we have $H(W) = \log\|\mathscr{W}\|$. The formulas (B.3), (B.4), and (B.5) are proved by Lemma B.1; see the following.

Lemma B.1. The random vectors Y^N and Z^N and the random variables W, U, A, S, X, Y, and Z of Theorem 6 satisfy

$$
\frac{1}{N}H\left(W\right) \leq I\left(U; Y\right) - I\left(U; S \mid A\right) + \frac{1}{N}\delta\left(P_e\right), \tag{B.6}
$$

$$
\frac{1}{N}H\left(W \mid Z^N\right) \leq H\left(Y \mid Z\right) + \frac{1}{N}\delta\left(P_e\right), \tag{B.7}
$$

$$
\frac{1}{N}H\left(W \mid Z^N\right) \leq H\left(A \mid Z\right), \tag{B.8}
$$

where $\delta(P_e) = h(P_e) + P_e\log(|\mathscr{W}| - 1)$. Note that $h(P_e) = -P_e\log P_e - (1 - P_e)\log(1 - P_e)$.

Substituting $H(W) = \log\|\mathscr{W}\|$ and (5) into (B.6), (B.7), and (B.8) and using the fact that $\epsilon \to 0$, the formulas (B.3), (B.4), and (B.5) are obtained. The formula (B.2) is from

$$
R_e \leq \lim_{N\to\infty}\Delta = \lim_{N\to\infty}\frac{1}{N}H\left(W \mid Z^N\right) \leq \lim_{N\to\infty}\frac{1}{N}H\left(W\right) = R. \tag{B.9}
$$

It remains to prove Lemma B.1; see the following.

Proof of Lemma B.1. The formula (B.6) follows from (B.10), (B.13), and (B.21). The formula (B.7) is from (B.11), (B.17) and (B.22). The formula (B.8) is proved by (B.12), (B.18), and (B.23).

Part (i). We begin with the left parts of the inequalities (B.6), (B.7), and (B.8); see the following.

Since $W \to Y^N \to Z^N$ is a Markov chain, for the message W, we have

$$
\begin{aligned}
\frac{1}{N}H\left(W\right) &= \frac{1}{N}H\left(W \mid Y^N\right) + \frac{1}{N}I\left(Y^N; W\right)\\
&\leq^{(a)} \frac{1}{N}\delta\left(P_e\right) + \frac{1}{N}I\left(Y^N; W\right).
\end{aligned}
\tag{B.10}
$$

For the equivocation to the wiretapper, we have

$$\begin{aligned}\frac{1}{N}H\left(W \mid Z^N\right) &=^{(b)} \frac{1}{N}\left(H\left(W \mid Z^N\right)+\delta\left(P_e\right)\right. \\ &\qquad \left.-H\left(W \mid Z^N, Y^N\right)\right) \\ &= \frac{1}{N}\left(I\left(W; Y^N \mid Z^N\right)+\delta\left(P_e\right)\right) \\ &\le \frac{1}{N}\left(H\left(Y^N \mid Z^N\right)+\delta\left(P_e\right)\right).\end{aligned} \tag{B.11}$$

Moreover,

$$\frac{1}{N}H\left(W \mid Z^N\right) = \frac{1}{N}H\left(A^N \mid Z^N\right). \tag{B.12}$$

Note that (a) and (b) follow from Fano's inequality, and (B.12) is from the fact that A^N is a deterministic function of W.

Part (ii). By using chain rule, the character $I(Y^N; W)$ in formulas (B.10) and (B.11) can be bounded as follows:

$$\begin{aligned}&\frac{1}{N}I\left(Y^N; W\right) \\ &= \frac{1}{N}\sum_{i=1}^{N} I\left(Y_i; W \mid Y^{i-1}\right) \\ &=^{(1)} \frac{1}{N}\sum_{i=1}^{N}\left(I\left(Y_i; W \mid Y^{i-1}\right)-I\left(S_i; W \mid S_{i+1}^N, A^N\right)\right) \\ &= \frac{1}{N}\sum_{i=1}^{N}\left(I\left(Y_i; W, S_{i+1}^N, A^N \mid Y^{i-1}\right)\right. \\ &\qquad -I\left(Y_i; S_{i+1}^N, A^N \mid W, Y^{i-1}\right) \\ &\qquad -I\left(S_i; W, Y^{i-1} \mid S_{i+1}^N, A^N\right) \\ &\qquad \left.+I\left(S_i; Y^{i-1} \mid W, S_{i+1}^N, A^N\right)\right) \\ &=^{(2)} \frac{1}{N}\sum_{i=1}^{N}\left(I\left(Y_i; W, S_{i+1}^N, A^N \mid Y^{i-1}\right)\right. \\ &\qquad \left.-I\left(S_i; W, Y^{i-1} \mid S_{i+1}^N, A^N\right)\right) \\ &= \frac{1}{N}\sum_{i=1}^{N}\left(H\left(Y_i \mid Y^{i-1}\right)-H\left(Y_i \mid Y^{i-1}, W, S_{i+1}^N, A^N\right)\right. \\ &\qquad \left.-H\left(S_i \mid S_{i+1}^N, A^N\right)+H\left(S_i \mid S_{i+1}^N, A^N, W, Y^{i-1}\right)\right) \\ &\le^{(3)} \frac{1}{N}\sum_{i=1}^{N}\left(H\left(Y_i\right)-H\left(Y_i \mid Y^{i-1}, W, S_{i+1}^N, A^N\right)\right. \\ &\qquad \left.-H\left(S_i \mid A_i\right)+H\left(S_i \mid S_{i+1}^N, A^N, W, Y^{i-1}\right)\right),\end{aligned} \tag{B.13}$$

where formula (1) follows from that $W \to A^N \to S^N$, and formula (2) follows from that

$$\sum_{i=1}^{N} I\left(Y_i; S_{i+1}^N, A^N \mid W, Y^{i-1}\right) = \sum_{i=1}^{N} I\left(S_i; Y^{i-1} \mid W, S_{i+1}^N, A^N\right), \tag{B.14}$$

and formula (3) follows from that $S_i \to A_i \to (S_{i+1}^N, A^{i-1}, A_{i+1}^N)$.

Proof of (B.14). The left part of (B.14) can be rewritten as

$$\begin{aligned}&\sum_{i=1}^{N} I\left(Y_i; S_{i+1}^N, A^N \mid W, Y^{i-1}\right) \\ &\quad = \sum_{i=1}^{N}\left(H\left(Y_i \mid W, Y^{i-1}\right)-H\left(Y_i \mid W, Y^{i-1}, S_{i+1}^N, A^N\right)\right) \\ &\quad =^{(1)} \sum_{i=1}^{N}\left(H\left(Y_i \mid A^N, Y^{i-1}\right)-H\left(Y_i \mid Y^{i-1}, S_{i+1}^N, A^N\right)\right) \\ &\quad = \sum_{i=1}^{N} I\left(Y_i; S_{i+1}^N \mid A^N, Y^{i-1}\right) \\ &\quad = \sum_{i=1}^{N}\sum_{j=i+1}^{N} I\left(Y_i; S_j \mid A^N, Y^{i-1}, S_{j+1}^N\right) \\ &\quad = \sum_{j=1}^{N}\sum_{i=j+1}^{N} I\left(Y_j; S_i \mid A^N, Y^{j-1}, S_{i+1}^N\right) \\ &\quad = \sum_{i=1}^{N}\sum_{j=1}^{i+1} I\left(Y_j; S_i \mid A^N, Y^{j-1}, S_{i+1}^N\right),\end{aligned} \tag{B.15}$$

where (1) is from the fact that A^N is a deterministic function of W.

The right part of (B.14) can be rewritten as

$$\begin{aligned}&\sum_{i=1}^{N} I\left(S_i; Y^{i-1} \mid W, S_{i+1}^N, A^N\right) \\ &\quad =^{(2)} \sum_{i=1}^{N} I\left(S_i; Y^{i-1} \mid S_{i+1}^N, A^N\right) \\ &\quad = \sum_{i=1}^{N}\sum_{j=1}^{i-1} I\left(Y_j; S_i \mid A^N, Y^{j-1}, S_{i+1}^N\right),\end{aligned} \tag{B.16}$$

where (2) is from the fact that A^N is a deterministic function of W.

The formula (B.14) is proved by (B.15) and (B.16). The proof is completed.

Part (iii). The character $H(Y^N \mid Z^N)$ in formula (B.11) can be rewritten as follows:

$$\begin{aligned}\frac{1}{N}H\left(Y^N \mid Z^N\right) &= \frac{1}{N}\sum_{i=1}^{N}H\left(Y_i \mid Y^{i-1}, Z^N\right)\\ &\le \frac{1}{N}\sum_{i=1}^{N}H\left(Y_i \mid Z_i\right).\end{aligned} \tag{B.17}$$

Part (iv). The character $H(A^N \mid Z^N)$ in formula (B.12) can be rewritten as follows:

$$\frac{1}{N}H\left(A^N \mid Z^N\right) \le \frac{1}{N}\sum_{i=1}^{N}H\left(A_i \mid Z_i\right). \tag{B.18}$$

Part (v) (single letter). To complete the proof, we introduce a random variable J, which is independent of W, A^N, X^N, S^N, Y^N and Z^N. Furthermore, J is uniformly distributed over $\{1, 2, \ldots, N\}$. Define

$$U = \left(W, Y^{J-1}, S_{J+1}^N, A^N, J\right), \tag{B.19}$$

$$X = X_J, \qquad Y = Y_J, \qquad Z = Z_J, \qquad S = S_J, \qquad A = A_J. \tag{B.20}$$

Part (vi). Then (B.13) can be rewritten as

$$\begin{aligned}&\frac{1}{N}I\left(W; Y^N\right)\\ &\le \frac{1}{N}\sum_{i=1}^{N}\left(H\left(Y_i\right) - H\left(Y_i \mid Y^{i-1}, W, S_{i+1}^N, A^N\right)\right.\\ &\qquad - H\left(S_i \mid A_i\right)\\ &\qquad \left.+H\left(S_i \mid S_{i+1}^N, A^N, W, Y^{i-1}\right)\right)\\ &= \frac{1}{N}\sum_{i=1}^{N}\left(H\left(Y_i \mid J = i\right) - H\left(Y_i \mid Y^{i-1}, W, S_{i+1}^N, A^N, J = i\right)\right.\\ &\qquad - H\left(S_i \mid A_i, J = i\right)\\ &\qquad \left.+ H\left(S_i \mid S_{i+1}^N, A^N, W, Y^{i-1}, A_i, J = i\right)\right)\\ &= H\left(Y_J \mid J\right) - H\left(Y_J \mid Y^{J-1}, W, S_{J+1}^N, A^N, J\right)\\ &\quad - H\left(S_J \mid A_J, J\right) + H\left(S_J \mid S_{J+1}^N, A^N, W, Y^{J-1}, A_J, J\right)\\ &\le H\left(Y_J\right) - H\left(Y_J \mid Y^{J-1}, W, S_{J+1}^N, A^N, J\right) - H\left(S_J \mid A_J, J\right)\\ &\quad + H\left(S_J \mid S_{J+1}^N, A^N, W, Y^{J-1}, A_J, J\right)\\ &= H\left(Y\right) - H\left(Y \mid U\right) - H\left(S \mid A\right) + H\left(S \mid U, A\right)\\ &= I\left(U; Y\right) - I\left(U; S \mid A\right).\end{aligned} \tag{B.21}$$

Analogously, (B.17) is rewritten as follows:

$$\begin{aligned}\frac{1}{N}H\left(Y^N \mid Z^N\right) &\le \frac{1}{N}\sum_{i=1}^{N}H\left(Y_i \mid Z_i\right)\\ &= \frac{1}{N}\sum_{i=1}^{N}\left(H\left(Y_i \mid Z_i, J = i\right)\right)\\ &= H\left(Y_J \mid Z_J, J\right)\\ &\le H\left(Y \mid Z\right).\end{aligned} \tag{B.22}$$

Similarly, (B.18) can be rewritten as follows:

$$\begin{aligned}\frac{1}{N}H\left(A^N \mid Z^N\right) &\le \frac{1}{N}\sum_{i=1}^{N}H\left(A_i \mid Z_i\right)\\ &= \frac{1}{N}\sum_{i=1}^{N}H\left(A_i \mid Z_i, J = i\right)\\ &= H\left(A_J \mid Z_J, J\right)\\ &\le H\left(A_J, J \mid Z_J\right)\\ &= H\left(A \mid Z\right).\end{aligned} \tag{B.23}$$

Substituting (B.21), (B.22), (B.23) into (B.10), (B.11), and (B.12), Lemma B.1 is proved.

The proof of Theorem 6 is completed.

C. Size Constraint of the Auxiliary Random Variable in Theorem 6

By using the support lemma (see [20, page 310]), it suffices to show that the random variable U can be replaced by new one, preserving the Markovity $(U, A) \to (X, S) \to Y \to Z$ and the mutual information $I(U; Z)$, $I(U; Y)$, $I(U; S \mid A)$, and furthermore, the range of the new U satisfies, $\|\mathcal{U}\| \le \|\mathcal{X}\|\|\mathcal{S}\|\|\mathcal{A}\| + 1$. The proof is in the reminder of this section.

Let

$$\overline{p} = p_{XSA}\left(x, s, a\right). \tag{C.1}$$

Define the following continuous scalar functions of $\overline{p}$:

$$\begin{aligned}f_{XSA}\left(\overline{p}\right) &= p_{XSA}\left(x, s, a\right), f_Y\left(\overline{p}\right)\\ &= H\left(Y\right), f_{S|A}\left(\overline{p}\right) = H\left(S \mid A\right).\end{aligned} \tag{C.2}$$

Since there are $\|\mathcal{X}\|\|\mathcal{S}\|\|\mathcal{A}\| - 1$ functions of $f_{XSA}(\overline{p})$, the total number of the continuous scalar functions of $\overline{p}$ is $\|\mathcal{X}\|\|\mathcal{S}\|\|\mathcal{A}\| + 1$.

Let $\overline{p}_{XSA|U} = \Pr\{X = x, S = s, A = a \mid U = u\}$. With these distributions $\overline{p}_{XSA|U} = \Pr\{X = x, S = s, A = a \mid U = u\}$, we have

$$p_{XSA}(x,s,a) = \sum_{u\in\mathcal{U}} p(U=u)\, f_{XSA}\left(\overline{p}_{XSA|U}\right), \tag{C.3}$$

$$I(U;S \mid A) = f_{S|A}\left(\overline{p}\right) - \sum_{u\in\mathcal{U}} p(U=u)\, f_{S|A}\left(\overline{p}_{XSA|U}\right), \tag{C.4}$$

$$H(Y \mid U) = \sum_{u\in\mathcal{U}} p(U=u)\, f_{Y}\left(\overline{p}_{XSA|U}\right). \tag{C.5}$$

According to the support lemma ([20, page 310]), the random variable U can be replaced by new ones such that the new U takes at most $\|\mathcal{X}\|\|\mathcal{S}\|\|\mathcal{A}\| + 1$ different values and the expressions (C.3), (C.4), and (C.5) are preserved.

D. Proof of the Direct Part of Theorem 9

In this section, we will show that any pair $(R, R_e) \in \mathcal{R}^c$ is achievable. Block Markov coding and Ahlswede-Cai's secret key on the feedback system are used in the construction of the code-book.

Now the remainder of this section is organized as follows. The code construction is in Appendix D.1. The proof of achievability is given in Appendix D.2.

D.1. Code Construction. Since $R_e \le H(Y \mid Z)$, $R_e \le H(A \mid Z)$, and $R_e \le R \le I(U;Y)$, it is sufficient to show that the pair $(R, R_e = \min\{H(Y \mid Z), H(A \mid Z), I(U;Y)\}$ is achievable, and note that this implies that $R \ge R_e = \min\{H(Y \mid Z), H(A \mid Z), I(U;Y)\}$.

We use the block Markov coding method. The random vectors U^N, A^N, S^N, X^N, Y^N, and Z^N consist of n blocks of length N. The message for n blocks is $W^n \triangleq (W_2, W_3, \ldots, W_n)$, where W_i $(2 \le i \le n)$ are i.i.d. random variables uniformly distributed over $\mathcal{W}$. Note that, in the first block, there is no W_1.

Let $\widetilde{Z}_i$ $(1 \le i \le n)$ be the output of the wiretap channel for block i, $Z^n = (\widetilde{Z}_1, \ldots, \widetilde{Z}_n)$, $Z^{\overline{j}} = (\widetilde{Z}_1, \ldots, \widetilde{Z}_{j-1}, \widetilde{Z}_{j+1}, \ldots, \widetilde{Z}_n)$ $(1 \le j \le n)$. Similarly, $Y^n = (\widetilde{Y}_1, \ldots, \widetilde{Y}_n)$ and $\widetilde{Y}_i$ $(1 \le i \le n)$ are the output of the main channel for block i. The specific values of the above random vectors are denoted by lowercase letters.

Given a pair (R, R_e), choose a joint probability mass function $p_{U,A,S,X,Y,Z}(u,a,s,x,y,z)$ such that

$$\begin{aligned} &0 \le R_e \le R, \\ &R \le I(U;Y), \\ &R_e = \min\{H(Y \mid Z), H(A \mid Z)\}. \end{aligned} \tag{D.1}$$

The message set $\mathcal{W}$ satisfies the following condition:

$$\lim_{N\to\infty} \frac{1}{N} \log \|\mathcal{W}\| = R = I(U;Y) - \gamma, \tag{D.2}$$

where γ is a fixed positive real numbers.

Code-book generation:

(i) Construction of A^N and S^N. In the first block, generate a i.i.d. sequence a^N according to the probability mass function $p_A(a)$, and choose it as the output of the action encoder. Let s^N be the state sequence generated in response to the chosen action sequence a^N.

For the ith block $(2 \le i \le n)$, generate 2^{NR} i.i.d. sequences a^N, according to the probability mass function $p_A(a)$. Index each sequence by $i \in \{1, 2, \ldots, 2^{NR}\}$. For a given message w_i $(w_i \in \mathcal{W})$, choose a corresponding $a^N(w_i)$ as the output of the action encoder. Let s^N be the state sequence generated in response to the action sequence $a^N(w_i)$.

(ii) Construction of the Secret Key. For the ith block $(2 \le i \le n)$, firstly we generate a mapping $g_f : \mathcal{Y}^N \to \mathcal{W}$ (note that $\|\mathcal{Y}\|^N \ge \|\mathcal{W}\|$). Define a random variable $K_i = g_f(\widetilde{Y}_{i-1})$ $(2 \le i \le n)$, which is uniformly distributed over $\mathcal{W}$, and K_i is independent of W_i. Reveal the mapping g_f to both receivers and the transmitter.

Then, when the transmitter receives the output $\widetilde{y}_{i-1}$ of the i-1th block, he computes $k_i = g_f(\widetilde{y}_{i-1}) \in \mathcal{W}$ as the secret key used in the ith block.

(iii) Construction of U^N. In the first block, for the transmitted action sequence a^N and the corresponding state sequence s^N, generate a i.i.d. sequence u^N according to the probability mass function $p_{U|A,S}(u_i \mid a_i, s_i)$. Choose u^N as a realization of U^N for the first block.

For the ith block $(2 \le i \le n)$, given the transmitted action sequence $a^N(w_i)$ and the corresponding state sequence s^N, generate 2^{NR} i.i.d. sequences u^N, according to the probability mass function $p_{U|A,S}(u_i \mid a_i(w_i), s_i)$. Index each sequence by $i \in \{1, 2, \ldots, 2^{NR}\}$.

For a given w_i and a secret key k_i, the transmitter chooses a sequence $u^N(w_i \oplus k_i)$ (where $\oplus$ is the modulo addition over $\mathcal{W}$) to transmit.

(iv) Construction of X^N. For each block, the x^N is generated according to a new discrete memoryless channel (DMC) with inputs u^N, s^N, and output x^N. The transition probability of this new DMC is $p_{X|U,S}(x \mid u, s)$, which is obtained from the joint probability mass function $p_{U,A,S,X,Y,Z}(u,a,s,x,y,z)$.

The probability $p_{X^N|U^N,S^N}(x^N \mid u^N, s^N)$ is calculated as follows:

$$p_{X^N|U^N,S^N}\left(x^N \mid u^N, s^N\right) = \prod_{i=1}^{N} p_{X|U,S}\left(x_i \mid u_i, s_i\right). \tag{D.3}$$

Decoding. For the ith block $(2 \le i \le n)$, given a vector $y^N \in \mathcal{Y}^N$ and a secret key k_i (k_i is known by the receiver), try to find a sequence $u^N(\widehat{w}_i \oplus k_i)$ such that $(u^N(\widehat{w}_i \oplus k_i), y^N) \in T^N_{UY}(\epsilon_3)$. If there exist sequences with the same $\widehat{w}_i \oplus k_i$, by using the secret key k_i, put out the corresponding $\widehat{w}_i$. Otherwise, that is, if no such sequence exists or multiple sequences have different message indices, declare a decoding error.

D.2. Proof of Achievability. The proof of achievability for Theorem 9 is along the lines of that for Theorem 6, and, therefore, we omit the proof here.

The proof of Theorem 9 is completed.

E. Proof of the Converse Part of Theorem 9

In this section, we prove Theorem 9: all the achievable (R, R_e) pairs are contained in the set $\mathscr{R}^{(c)}$. Suppose that (R, R_e) is achievable; that is, for any given $\epsilon > 0$, there exists a channel encoder-decoder (N, Δ, P_e) such that

$$\lim_{N\to\infty} \frac{\log \|\mathscr{W}\|}{N} = R, \qquad \lim_{N\to\infty} \Delta \geq R_e, P_e \leq \epsilon. \tag{E.1}$$

Then we will show the existence of random variables $(A, U) \to (X, S) \to Y \to Z$ such that

$$0 \leq R_e \leq R, \tag{E.2}$$

$$R \leq I(U; Y), \tag{E.3}$$

$$R_e \leq H(Y \mid Z), \tag{E.4}$$

$$R_e \leq H(A \mid Z). \tag{E.5}$$

Since W is uniformly distributed over $\mathscr{W}$, we have $H(W) = \log \|\mathscr{W}\|$. The formulas (E.3), (E.4), and (E.5) are proved by Lemma E.1; see the following.

Lemma E.1. The random vectors Y^N and Z^N and the random variables W, X, U, S, A, Y, and Z of Theorem 9 satisfy

$$\frac{1}{N}H(W) \leq I(U; Y) + \frac{1}{N}\delta(P_e), \tag{E.6}$$

$$\frac{1}{N}H\left(W \mid Z^N\right) \leq H(Y \mid Z) + \frac{1}{N}\delta(P_e), \tag{E.7}$$

$$\frac{1}{N}H\left(W \mid Z^N\right) \leq H(A \mid Z), \tag{E.8}$$

where $\delta(P_e) = h(P_e) + P_e \log(|\mathscr{W}| - 1)$. Note that $h(P_e) = -P_e \log P_e - (1 - P_e)\log(1 - P_e)$.

Substituting $H(W) = \log \|\mathscr{W}\|$ and (5) into (E.6), (E.7), and (E.8) and using the fact that $\epsilon \to 0$, the formulas (E.3), (E.4), and (E.5) are obtained. The formula (E.2) is from

$$R_e \leq \lim_{N\to\infty} \Delta = \lim_{N\to\infty} \frac{1}{N}H\left(W \mid Z^N\right) \leq \lim_{N\to\infty} \frac{1}{N}H(W) = R. \tag{E.9}$$

It remains to prove Lemma E.1; see the following.

Proof of Lemma E.1. The formula (E.6) follows from (E.10), (E.13), and (E.17). The formula (E.7) is from (E.11), (E.14), and (E.18). The formula (E.8) is proved by (E.12), (E.15), and (E.19).

Part (i). We begin with the left parts of the inequalities (E.6), (E.7), and (E.8); see the following.

Since $W \to Y^N \to Z^N$ is a Markov chain, for the message W, we have

$$\begin{aligned} \frac{1}{N}H(W) &= \frac{1}{N}H\left(W \mid Y^N\right) + \frac{1}{N}I\left(Y^N; W\right) \\ &\leq^{(a)} \frac{1}{N}\delta(P_e) + \frac{1}{N}I\left(Y^N; W\right). \end{aligned} \tag{E.10}$$

For the equivocation to the wiretapper, we have

$$\begin{aligned} \frac{1}{N}H\left(W \mid Z^N\right) &=^{(b)} \frac{1}{N}\left(H\left(W \mid Z^N\right) + \delta(P_e)\right. \\ &\qquad \left. - H\left(W \mid Z^N, Y^N\right)\right) \\ &= \frac{1}{N}\left(I\left(W; Y^N \mid Z^N\right) + \delta(P_e)\right) \\ &\leq \frac{1}{N}\left(H\left(Y^N \mid Z^N\right) + \delta(P_e)\right), \end{aligned} \tag{E.11}$$

$$\frac{1}{N}H\left(W \mid Z^N\right) = \frac{1}{N}H\left(A^N \mid Z^N\right). \tag{E.12}$$

Note that (a) and (b) follow from Fano's inequality, and (E.12) is from the fact that A^N is a deterministic function of W.

Part (ii). By using chain rule, the character $I(Y^N; W)$ in formulas (E.10) and (E.11) can be bounded as follows:

$$\begin{aligned} \frac{1}{N}I\left(Y^N; W\right) &= \frac{1}{N}\sum_{i=1}^{N} I\left(Y_i; W \mid Y^{i-1}\right) \\ &\leq \frac{1}{N}\sum_{i=1}^{N} I\left(Y_i; W, Y^{i-1}\right) \\ &\leq \frac{1}{N}\sum_{i=1}^{N} I\left(Y_i; W, Y^{i-1}, S^{i-1}\right). \end{aligned} \tag{E.13}$$

Part (iii). Similar to (E.13), the character $I(W; Z^N)$ in formula (E.11) can be rewritten as follows:

$$\begin{aligned} \frac{1}{N}H\left(Y^N \mid Z^N\right) &= \frac{1}{N}\sum_{i=1}^{N} H\left(Y_i \mid Y^{i-1}, Z^N\right) \\ &\leq \frac{1}{N}\sum_{i=1}^{N} H\left(Y_i \mid Z_i\right). \end{aligned} \tag{E.14}$$

Part (iv). The character $H(A^N \mid Z^N)$ in formula (E.12) can be rewritten as follows:

$$\frac{1}{N}H\left(A^N \mid Z^N\right) \leq \frac{1}{N}\sum_{i=1}^{N} H\left(A_i \mid Z_i\right). \tag{E.15}$$

Part (v) (single letter). To complete the proof, we introduce a random variable J, which is independent of W, A^N, X^N,

S^N, Y^N, and Z^N. Furthermore, J is uniformly distributed over $\{1, 2, \ldots, N\}$. Define

$$U = \left(W, Y^{J-1}, S^{J-1}, J\right),$$
$$X = X_J, \quad Y = Y_J, \quad Z = Z_J, \quad S = S_J, \quad A = A_J. \tag{E.16}$$

Part (vi). Then (E.13) can be rewritten as

$$\begin{aligned}\frac{1}{N} I\left(W; Y^N\right) &\le \frac{1}{N}\sum_{i=1}^{N} I\left(Y_i; W, Y^{i-1}, S^{i-1}\right)\\ &= \frac{1}{N}\sum_{i=1}^{N} I\left(Y_i; W, Y^{i-1}, S^{i-1} \mid J = i\right)\\ &= I\left(Y_J; W, Y^{J-1}, S^{J-1} \mid J\right)\\ &\le I\left(Y_J; W, Y^{J-1}, S^{J-1}, J\right)\\ &= I\left(U; Y\right).\end{aligned} \tag{E.17}$$

Analogously, (E.14) is rewritten as follows:

$$\begin{aligned}\frac{1}{N} H\left(Y^N \mid Z^N\right) &\le \frac{1}{N}\sum_{i=1}^{N} H\left(Y_i \mid Z_i\right)\\ &= \frac{1}{N}\sum_{i=1}^{N} \left(H\left(Y_i \mid Z_i, J = i\right)\right)\\ &= H\left(Y_J \mid Z_J, J\right)\\ &\le H\left(Y \mid Z\right).\end{aligned} \tag{E.18}$$

Similarly, (E.15) can be rewritten as follows,

$$\begin{aligned}\frac{1}{N} H\left(A^N \mid Z^N\right) &\le \frac{1}{N}\sum_{i=1}^{N} H\left(A_i \mid Z_i\right)\\ &= \frac{1}{N}\sum_{i=1}^{N} H\left(A_i \mid Z_i, J = i\right)\\ &= H\left(A_J \mid Z_J, J\right)\\ &\le H\left(A_J \mid Z_J\right)\\ &= H\left(A \mid Z\right).\end{aligned} \tag{E.19}$$

Substituting (E.17), (E.18), and (E.19) into (E.10), (E.11), and (E.12), Lemma E.1 is proved.

The proof of Theorem 9 is completed.

Acknowledgments

The authors would like to thank Professor N. Cai for his help to improve this paper. This work was supported by a subproject in National Basic Research Program of China under Grant 2012CB316100 on Broadband Mobile Communications at High Speeds, the National Natural Science Foundation of China under Grant 61271222, and the Research Fund for the Doctoral Program of Higher Education of China (no. 20100073110016).

References

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

[2] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Transactions on Information Theory*, vol. 24, no. 3, pp. 339–348, 1978.

[3] J. Körner and K. Marton, "General broadcast channels with degraded message sets," *IEEE Transactions on Information Theory*, vol. 23, no. 1, pp. 60–64, 1977.

[4] N. Merhav, "Shannon's secrecy system with informed receivers and its application to systematic coding for wiretapped channels," *IEEE Transactions on Information Theory*, vol. 54, no. 6, pp. 2723–2734, 2008.

[5] R. Ahlswede and N. Cai, "Transmission, identification and common randomness capacities for wire-tape channels with secure feedback from the decoder," in *General Theory of Information Transfer and Combinatorics*, vol. 4123 of *Lecture Notes in Computer Science*, pp. 258–275, Springer, Berlin, Germany, 2006.

[6] B. Dai, A. J. Han Vinck, Y. Luo, and Z. Zhuang, "Capacity region of non-degraded wiretap channel with noiseless feedback," in *Proceedings of the IEEE International Symposium on Information Theory*, Cambridge, Mass, USA, July 2012.

[7] L. Lai, H. El Gamal, and H. V. Poor, "The wiretap channel with feedback: encryption over the channel," *IEEE Transactions on Information Theory*, vol. 54, no. 11, pp. 5059–5067, 2008.

[8] E. Ardestanizadeh, M. Franceschetti, T. Javidi, and Y.-H. Kim, "Wiretap channel with secure rate-limited feedback," *IEEE Transactions on Information Theory*, vol. 55, no. 12, pp. 5353–5361, 2009.

[9] G. Chen and C. Chang, "A construction for secret sharing scheme with general access structure," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 4, no. 1, pp. 1–8, 2013.

[10] R. Nishimura, S. Abe, N. Fujita, and Y. Suzuki, "Reinforcement of VoIP security with multipath routing and secret sharing scheme," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 1, no. 3, pp. 204–219, 2010.

[11] Z. Wang, C. Chang, H. N. Tu, and M. Li, "Sharing a secret image in binary images with verification," *Journal of Information Hiding and Multimedia Signal Processing*, vol. 2, no. 1, pp. 78–90, 2011.

[12] C. E. Shannon, "Channels with side information at the transmitter," *IBM Journal of Research and Development*, vol. 2, pp. 289–293, 1958.

[13] A. V. Kuznecov and B. S. Cybakov, "Coding in a memory with imperfect cells," *Problemy Peredachi Informatsii*, vol. 10, no. 2, pp. 52–60, 1974.

[14] S. I. Gel'efand and M. S. Pinsker, "Coding for channel with random parameters," *Problems of Control and Information Theory*, vol. 9, no. 1, pp. 19–31, 1980.

[15] M. H. M. Costa, "Writing on dirty paper," *IEEE Transactions on Information Theory*, vol. 29, no. 3, pp. 439–441, 1983.

[16] T. Weissman, "Capacity of channels with action-dependent states," *IEEE Transactions on Information Theory*, vol. 56, no. 11, pp. 5396–5411, 2010.

[17] C. Mitrpant, A. J. Han Vinck, and Y. Luo, "An achievable region for the Gaussian wiretap channel with side information," *IEEE Transactions on Information Theory*, vol. 52, no. 5, pp. 2181–2190, 2006.

[18] Y. Chen and A. J. H. Vinck, "Wiretap channel with side information," *IEEE Transactions on Information Theory*, vol. 54, no. 1, pp. 395–402, 2008.

[19] B. Dai and Y. Luo, "Some new results on wiretap channel with side information," *Entropy*, vol. 14, pp. 1671–1702, 2012.

[20] I. Csiszár and J. Körner, *Information Theory*, Probability and Mathematical Statistics, Academic Press, London, UK, 1981.