

Article

Some New Results on the Wiretap Channel with Side Information

Bin Dai ^{1,2,3} and Yuan Luo ^{1,3*}

¹ Computer Science and Engineering Department, Shanghai Jiao Tong University, Dongchuan road 800, Shanghai 200240, China; E-Mail: daibin007@sjtu.edu.cn (B.D.)

² Institute for Experimental Mathematics, Duisburg-Essen University, Ellernstraße 29, Essen 45326, Germany

³ The State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an 710000, China

* Author to whom correspondence should be addressed; E-Mail: luoyuan@cs.sjtu.edu.cn; Tel.: +86 21 3420 5477.

Received: 17 July 2012; in revised form: 20 August 2012 / Accepted: 30 August 2012 /

Published: 7 September 2012

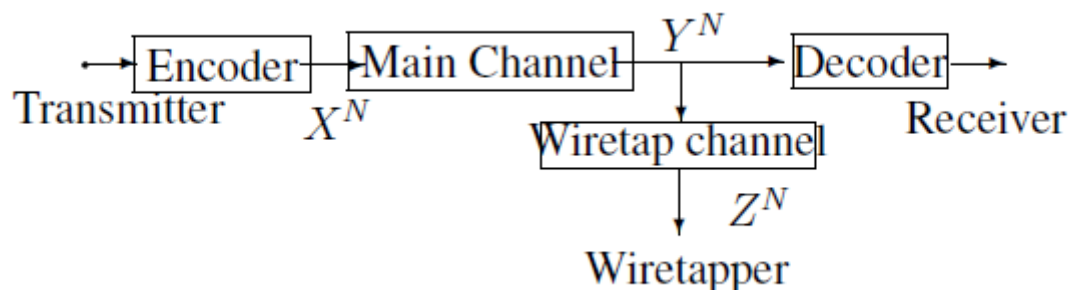
Abstract: In this paper, the model of wiretap channel has been reconsidered for the case that the main channel is controlled by channel state information (side information), and it is available at the transmitter in a noncausal manner (termed here noncausal side information) or causal manner (termed here causal side information). Inner and outer bounds are derived on the capacity-equivocation regions for the noncausal and causal manners, and the secrecy capacities for both manners are described and bounded, which provide the best transmission rate with perfect secrecy. Moreover, for the case that the side information is available at the transmitter in a memoryless manner (termed here memoryless side information), both the capacity-equivocation region and the secrecy capacity are determined. The results of this paper extend the previous work on wiretap channel with noncausal side information by providing an outer bound on the capacity-equivocation region. In addition, we find that the memoryless side information can not help to obtain the same secrecy capacity as that of the causal case, and this is different from the well known fact that the memoryless manner can achieve the capacity of the channel with causal side information.

Keywords: capacity-equivocation region; secrecy capacity; side information; wiretap channel

1. Introduction

The concept of the wiretap channel was first introduced by A.D. Wyner [1]. It is a kind of degraded broadcast channel. The wiretapper knows the encoding scheme used at the transmitter and the decoding scheme used at the legitimate receiver, see Figure 1. The object is to describe the rate of reliable communication from the transmitter to the legitimate receiver, subject to a constraint of the equivocation to the wiretapper.

Figure 1. The model of wiretap channel.



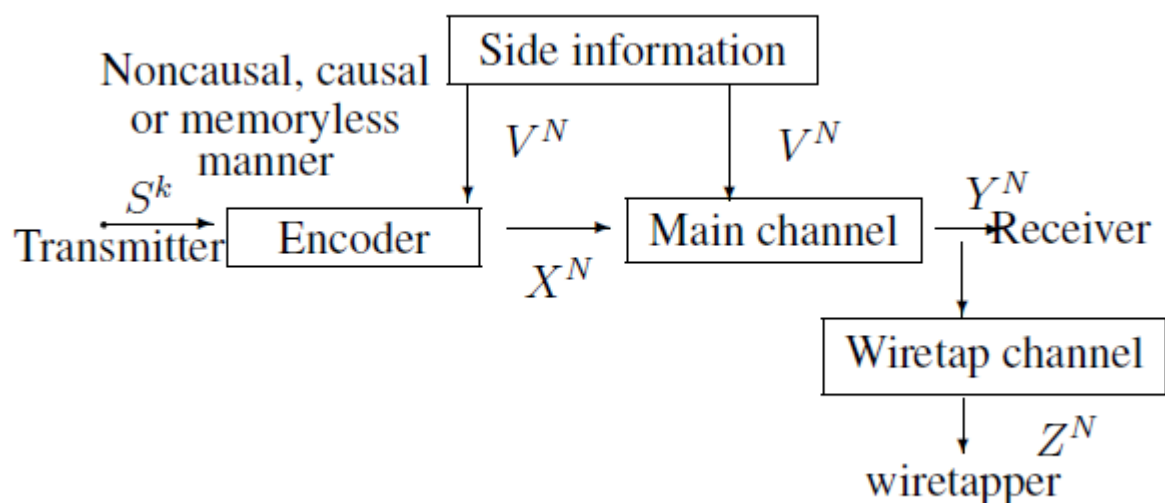
After the publication of A.D. Wyner's work, I. Csiszár and J. Körner [2] investigated a more general situation: the broadcast channels with confidential messages. It is clear that A.D. Wyner's wiretap channel is a special case of the model of I. Csiszár and J. Körner, in a manner that the main channel is less noisy than the wiretap channel. Furthermore, S.K. Leung-Yan-Cheong and M.E. Hellman studied the Gaussian wiretap channel (GWC) [3], and showed that its secrecy capacity was the difference between the main channel capacity and the overall wiretap channel capacity (the cascade of main channel and wiretap channel).

The coding for channels with causal (past and current) side information at the encoder was first investigated by C.E. Shannon [4] in 1958. After that, in order to solve the problem of coding for a computer memory with defective cells, N.V. Kuznetsov and B.S. Tsybakov [5] considered a channel in the presence of non-causal side information at the transmitter. They provided some coding techniques without determination of the capacity. The capacity was found in 1980 by S. I. Gel'fand and M. S. Pinsker [6]. Furthermore, Max H.M. Costa [7] 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. Based on the dirty paper channel, C. Mitrpant *et al.* [8] studied the Gaussian wiretap channel with side information, and provided an inner bound on the capacity-equivocation region. Furthermore, Y. Chen *et al.* [9] investigated the discrete memoryless wiretap channel with noncausal side information, and also provided an inner bound on the capacity-equivocation region. Note that the coding scheme of [9] is a combination of those in [1,6]. Chen *et al.* [9] generalize Mitrpant *et al.*'s work [8] by extending the Gaussian channel to the discrete memoryless channel (DMC), *i.e.*, the result of [8] can be obtained from that of [9]. Recently, N. Merhav [10] studied a variation of the wiretap channel, and obtained the capacity-equivocation 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 this paper, we study the model of wiretap channel with side information, see Figure 2. The transition probability distribution of the main channel depends on a channel state information V^N , which is available at the encoder in a noncausal or causal manner. The wiretapper can get a degraded version of the symbols Y^N via a wiretap channel. Both the main channel and the wiretap channel are discrete memoryless channels.

Figure 2. Wiretap channel with side information.



Inner and outer bounds are derived on the capacity-equivocation regions for the noncausal and causal manners (the inner bound for the noncausal manner is in fact equivalent to that of [9]), and the secrecy capacity for both manners is described and bounded. Moreover, for the case that the side information is available at the transmitter in a memoryless manner (at time i , the encoder is only allowed to use the side information V_i), both the capacity-equivocation region and the secrecy capacity are determined. In Shannon's well known paper [4], it shows that the optimal way to achieve the capacity of the channel with causal side information is to use V_i instead of V^i for the channel encoder. Then, it is natural to think about whether the memoryless side information can help to obtain the same secrecy capacity as that of the wiretap channel with causal side information, and this is also our motivation on the study of the memoryless model.

Compared with [9], the inner bound on the capacity-equivocation region for the noncausal manner of this paper, in fact, is equivalent to the achievable region in [9]. However, the region provided in this paper is easier to understand than that of [9].

The remainder of this paper is organized as follows. In Section 2, we present the basic definitions and the main results on the capacity-equivocation regions. In Section 3, we prove the outer bounds on the capacity-equivocation regions for noncausal and causal manners, and provide the converse proof of the capacity-equivocation region for the memoryless manner. The inner bound for causal manner and the

direct part of the capacity-equivocation region for the memoryless manner are proved in Section 4. Final conclusions are presented in Section 5.

2. Notations, Definitions and the Main Results

Throughout the paper, random variables, sample values and alphabets are denoted by capital letters, lower case 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 $(U_1, \dots, U_N)$, and $u^N = (u_1, \dots, u_N)$ is a specific vector value in $\mathcal{U}^N$ that is the N th Cartesian power of $\mathcal{U}$. U_i^N denotes a random $N - i + 1$ -vector $(U_i, \dots, U_N)$, and $u_i^N = (u_i, \dots, u_N)$ is a specific vector value of U_i^N . Let $p_V(v)$ denote the probability mass function $\Pr\{V = v\}$.

In this section, the model of Figure 2 is considered in three parts. The model of Figure 2 with noncausal side information is described in Section 2.1, the causal side information is described in Section 2.2, and the memoryless side information is described in Section 2.3, see the following.

2.1. The Model of Figure 2 with Noncausal Side Information

In this subsection, a description of the wiretap channel with noncausal side information is given by Definitions 1–4. The inner and outer bounds on the capacity-equivocation region $\mathcal{C}$ composed of all achievable (R, d) pairs are given in Theorem 1 and Theorem 2, respectively, where the achievable (R, d) pair is defined in Definition 5.

Definition 1 (encoder) The source S^k is defined as $(S_1, S_2, \dots, S_k)$, where $S_i (1 \leq i \leq k)$ are i.i.d. random variables that take values in the finite set $\mathcal{S}$. Then $H(S^k) = kH_S$, where $H_S = H(S_i)$ for $1 \leq i \leq k$. The side information V^N is the output of a discrete memoryless source $P_V(\cdot)$, and it is available at the encoder in a noncausal manner. V^N is independent of S^k .

The inputs of the encoder are S^k and V^N , while the output is X^N . The encoder f^N is a matrix of conditional probabilities $f^N(x^N|s^k, v^N)$, where $x^N \in \mathcal{X}^N$, $s^k \in \mathcal{S}^k$, $v^N \in \mathcal{V}^N$, $\sum_{x^N} f^N(x^N|s^k, v^N) = 1$, and $f^N(x^N|s^k, v^N)$ is the probability that the source s^k and the side information v^N are encoded as the channel input x^N .

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

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

$$\Delta = \frac{H(S^k|Z^N)}{H(S^k)} \quad (2.1)$$

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

$$Q_{MW}(z|x, v) = \sum_{y \in \mathcal{Y}} Q_W(z|y) Q_M(y|x, v) \quad (2.2)$$

Let C_{MW} be the capacity of the channel Q_{MW} .

Note that, $(S^k, V^N) \rightarrow (X^N, V^N) \rightarrow Y^N \rightarrow Z^N$ is a Markov chain in the model of Figure 2.

Definition 4 (decoder) The decoder is a mapping $f_D : \mathcal{Y}^N \rightarrow \mathcal{S}^k$, with input Y^N and output $\hat{S}^k = f_D(Y^N)$. Let P_e be the error probability, and it is defined as $\Pr\{S^k \neq \hat{S}^k\}$.

Definition 5 (achievable (R, d) pair in the model of Figure 2) A pair (R, d) (where $R, d > 0$) is called achievable if, for any $\epsilon > 0$, there exists an encoder-decoder (N, k, Δ, P_e) such that

$$\frac{H_{S^k}}{N} \geq R - \epsilon, \Delta \geq d - \epsilon, P_e \leq \epsilon \quad (2.3)$$

The capacity-equivocation region $\mathcal{C}$ is a set composed of all achievable (R, d) pairs. Inner and outer bounds on $\mathcal{C}$ are respectively provided in the following Theorem 1 and Theorem 2.

Theorem 1 The capacity-equivocation region $\mathcal{C}$ of the wiretap channel with noncausal side information satisfies $\mathcal{R}_i \subseteq \mathcal{C}$, where

$$\begin{aligned} \mathcal{R}_i &= \{(R, d) : 0 \leq d \leq 1 \\ 0 \leq R &\leq I(U; Y) - I(U; V) \\ Rd &\leq \min\{I(U; Y) - I(U; Z), I(U; Y) - I(U; V)\}\} \end{aligned}$$

where the random variables U, X, V, Y and Z satisfy the following Markov chain,

$$U \rightarrow (X, V) \rightarrow Y \rightarrow Z$$

Remark 1 There are some notes on Theorem 1, see the following.

- The range of the random variable U satisfies

$$\|U\| \leq \|X\| \|V\| + 3$$

The proof is similar to that of Theorem 2, and it is omitted here.

- The region $\mathcal{R}_i$, in fact, is equivalent to the achievable region in [9], however, it is easier to understand than that of [9]. The proof of Theorem 1 is a combination of Gel'fand–Pinsker's technique [6] and Wyner's random binning method [1], and we omit it here.
- Secrecy capacity

The points in $\mathcal{R}_i$ for which $d = 1$ are of considerable interest, which imply the perfect secrecy $H(S^k) = H(S^k|Z^N)$. Clearly, we can easily bound the secrecy capacity C_s of the model of Figure 2 with noncausal side information by

$$\max \min\{I(U; Y) - I(U; Z), I(U; Y) - I(U; V)\} \leq C_s \leq \max\{I(U; Y) - I(U; V)\} \quad (2.4)$$

Theorem 2 The capacity-equivocation region $\mathcal{C}$, as defined above, satisfies $\mathcal{C} \subseteq \mathcal{R}_o$, where

$$\begin{aligned} \mathcal{R}_o &= \{(R, d) : 0 \leq d \leq 1 \\ 0 \leq R &\leq I(U; Y) - I(U; V) \\ Rd &\leq I(U; Y) - I(K; Z|A)\} \end{aligned}$$

where the random variables U, K, A, X, V, Y and Z satisfy the following Markov chains,

$$(U, K, A) \rightarrow (X, V) \rightarrow Y \rightarrow Z$$

$$(K, A) \rightarrow U \rightarrow Y \rightarrow Z$$

and A may be assumed to be a (deterministic) function of K (these are directly from the definitions of the random variables U, K, A, X, V, Y and Z , see Equations (3.18), (3.19), (3.20) and (3.21)).

Remark 2 There are some notes on Theorem 2, see the following.

- The ranges of the random variables U, K and A satisfy

$$\|\mathcal{A}\| \leq \|\mathcal{X}\| \|\mathcal{V}\|$$

$$\|\mathcal{K}\| \leq \|\mathcal{X}\|^2 \|\mathcal{V}\|^2$$

$$\|\mathcal{U}\| \leq \|\mathcal{X}\|^2 \|\mathcal{V}\|^2 (\|\mathcal{X}\| \|\mathcal{V}\| + 1)$$

The proof is in Appendix 5.

- Observing the formula $Rd \leq I(U; Y) - I(K; Z|A)$ in Theorem 2, we have

$$\begin{aligned} I(U; Y) - I(K; Z|A) & \stackrel{(a)}{=} I(U; Y) - H(Z|A) + H(Z|K) \\ & \geq I(U; Y) - H(Z) + H(Z|K) \\ & \geq I(U; Y) - H(Z) + H(Z|K, U) \\ & \stackrel{(b)}{=} I(U; Y) - H(Z) + H(Z|U) = I(U; Y) - I(U; Z) \end{aligned} \quad (2.5)$$

where (a) is from the fact that A may be assumed to be a (deterministic) function of K , and (b) is from the Markov chain $K \rightarrow U \rightarrow Y \rightarrow Z$. Then it is easy to see that $\mathcal{R}_i \subseteq \mathcal{R}_o$.

2.2. The Model of Figure 2 with Causal Side Information

The model of Figure 2 with causal side information is similar to the model with noncausal side information in Section 2.1, except that the side information V^N in Definition 1 is known to the encoder in a causal manner, i.e., at the i -th time ($1 \leq i \leq N$), the output of the encoder $x_i = f_i(s^k, v^i)$, where $v^i = (v_1, v_2, \dots, v_i)$ and f_i is the probability that the source s^k and the side information v^i are encoded as the channel input x_i at time i . Define

$$f^N(x^N | s^k, v^N) = \prod_{i=1}^N f_i(x_i | s^k, v^i) \quad (2.6)$$

Inner and outer bounds on the capacity-equivocation region $\mathcal{C}_c$ for the model of Figure 2 with causal side information are respectively provided in the following Theorem 3 and Theorem 4.

Theorem 3 The capacity-equivocation region $\mathcal{C}_c$ satisfies $\mathcal{R}_{ci} \subseteq \mathcal{C}_c$, where

$$\mathcal{R}_{ci} = \{(R, d) : 0 \leq d \leq 1$$

$$0 \leq R \leq I(U; Y)$$

$$Rd \leq I(U; Y) - I(U; Z)\}$$

where the random variables U , X , V , Y and Z satisfy the following Markov chain,

$$U \rightarrow (X, V) \rightarrow Y \rightarrow Z$$

Remark 3 There are some notes on Theorem 3, see the following.

- The range of the random variable U satisfies

$$\|\mathcal{U}\| \leq \|\mathcal{X}\| \|\mathcal{V}\| + 1$$

The proof is similar to that in Theorem 2, and it is omitted here.

- Secrecy capacity

The points in $\mathcal{R}_{ci}$ for which $d = 1$ are of considerable interest, which imply the perfect secrecy $H(S^k) = H(S^k|Z^N)$. Clearly, we can easily bound the secrecy capacity C_s^c of the model of Figure 2 with causal side information by

$$\max(I(U; Y) - I(U; Z)) \leq C_s^c \leq \max I(U; Y) \quad (2.7)$$

Theorem 4 The capacity-equivocation region $\mathcal{C}_c$ satisfies $\mathcal{C}_c \subseteq \mathcal{R}_{co}$, where

$$\begin{aligned} \mathcal{R}_{co} &= \{(R, d) : 0 \leq d \leq 1 \\ 0 \leq R &\leq I(U; Y) \\ Rd &\leq I(U; Y) - I(K; Z|A)\} \end{aligned}$$

where the random variables U , K , A , X , V , Y and Z satisfy the following Markov chains,

$$(U, K, A) \rightarrow (X, V) \rightarrow Y \rightarrow Z$$

$$(K, A) \rightarrow U \rightarrow Y \rightarrow Z$$

and A may be assumed to be a (deterministic) function of K (these are directly from the definitions of the random variables U , K , A , X , V , Y and Z , see Equations (3.18), (3.19), (3.20) and (3.21)).

Remark 4 There are some notes on Theorem 4, see the following.

- The ranges of the random variables U , K and A satisfy

$$\|\mathcal{A}\| \leq \|\mathcal{X}\| \|\mathcal{V}\|$$

$$\|\mathcal{K}\| \leq \|\mathcal{X}\|^2 \|\mathcal{V}\|^2$$

$$\|\mathcal{U}\| \leq \|\mathcal{X}\|^3 \|\mathcal{V}\|^3$$

The proof is similar to that of Theorem 2, and it is omitted here.

- Since the causal side information is a special case of the noncausal manner, the outer bound $\mathcal{R}_{co}$ can be directly obtained from $\mathcal{R}_o$ by using the fact that U is independent of V .
- Note that $I(U; Y) - I(K; Z|A) \geq I(U; Y) - I(U; Z)$ (the proof is the same as that in Remark 2), and therefore, it is easy to see that $\mathcal{R}_{ci} \subseteq \mathcal{R}_{co}$

2.3. The Model of Figure 2 with Memoryless Side Information

The model of Figure 2 with memoryless side information is similar to the model with causal side information in Section 2.2, except that the side information V^N in Definition 1 is known to the encoder in a memoryless manner, *i.e.*, at the i -th time ($1 \leq i \leq N$), the output of the encoder $x_i = f_i(s^k, v_i)$, where f_i is the probability that the source s^k and the side information v_i are encoded as the channel input x_i at time i . Define

$$f^N(x^N|s^k, v^N) = \prod_{i=1}^N f_i(x_i|s^k, v_i) \quad (2.8)$$

The capacity of the main channel for the memoryless case is determined by C. E. Shannon [4],

$$C_M = \max_{p_{X|U,V}(x|u,v)p_U(u)} I(U; Y) \quad (2.9)$$

where $U \rightarrow (X, V) \rightarrow Y$ is a Markov chain and $\|\mathcal{U}\| \leq \|\mathcal{X}\|\|\mathcal{V}\| + 1$. The proof of $\|\mathcal{U}\| \leq \|\mathcal{X}\|\|\mathcal{V}\| + 1$ is similar to that of Theorem 2, and it is omitted here.

A function $\Gamma'(R)$ used for describing the capacity-equivocation region composed of all achievable (R, d) pairs in the model of Figure 2 with memoryless side information is defined in Definition 6.

Definition 6 (function $\Gamma'(R)$) For $R \geq 0$, let

$$\rho(R) = \{p_{X|U,V}(x|u,v)p_U(u) : I(U; Y) \geq R\} \quad (2.10)$$

It is easy to see that $\rho(R)$ is empty for $R > C_M$, where C_M is the capacity of the main channel, see Equation (2.9). For $0 \leq R \leq C_M$, denote

$$\Gamma'(R) = \sup_{p_{X|U,V}(x|u,v)p_U(u) \in \rho(R)} I(U; Y) - I(U; Z) \quad (2.11)$$

The following Lemma 1 provides some properties about $\Gamma'(R)$. The proof of Lemma 1 is in Appendix 5.

Lemma 1 The quantity $\Gamma'(R)$, where $0 \leq R \leq C_M$, satisfies the following properties:

(i) The “supremum” in the definition of $\Gamma'(R)$ is, in fact, a maximum, *i.e.*, for each R , there exists a mass function $p_{X|U,V}(x|u,v)p_U(u) \in \rho(R)$ such that $I(U; Y) - I(U; Z) = \Gamma'(R)$.

(ii) $\Gamma'(R)$ is a concave function of R .

(iii) $\Gamma'(R)$ is non-increasing in R .

(iv) $\Gamma'(R)$ is continuous in R .

Our problem in the model of Figure 2 with memoryless side information is to characterize the capacity-equivocation region $\mathcal{C}_m$ composed of all achievable (R, d) pairs. The following Theorem 5 gives a characterization of the capacity-equivocation region $\mathcal{C}_m$, which is proved in the remaining sections. The secrecy capacity is defined in Remark 5 (see Equation (2.12)), which is bounded by the Formula (2.14).

Theorem 5 The capacity-equivocation region $\mathcal{C}_m$ is equal to $\mathcal{R}^*$, where

$$\begin{aligned}\mathcal{R}^* &= \{(R, d) : 0 \leq d \leq 1 \\ 0 \leq R \leq C_M \\ Rd \leq \Gamma'(R)\}\end{aligned}$$

Remark 5 There are some notes on Theorem 5, see the following.

- Comparison with A. D. Wyner's wiretap channel [1]

The main channel capacity C_M denoted by Equation (2.9) in Theorem 1 is different from that of [1]. When the channel state information V is a constant, the model of Figure 2 reduces to A. D. Wyner's wiretap channel [1]. Substituting V by a constant and U by X into Equations (2.9), (2.10) and (2.11), the characters C_M , $\rho(R)$, $\Gamma'(R)$ and the region $\mathcal{R}^*$ are the same as those of [1].

- Secrecy capacity

A transmission rate C'_s denoted by

$$C'_s = \max_{(R,1) \in \mathcal{R}^*} R \quad (2.12)$$

is called the **secrecy capacity** in the model of Figure 2 with memoryless side information. Furthermore, C'_s is the unique solution of the equation

$$C'_s = \Gamma'(C'_s) \quad (2.13)$$

and satisfies

$$0 \leq \Gamma'(C_M) \leq C'_s \leq \Gamma'(0) \quad (2.14)$$

Proof 1 (Proof of Equations (2.13) and (2.14)) Firstly, since $\Gamma'(0) > 0$, $\Gamma'(C_M) - C_M \leq 0$ and $\Gamma'(R) - R$ is a non-increasing function of R , then there exists a unique $C^* \in (0, C_M]$ such that $\Gamma'(C^*) - C^* = 0$ and $(C^*, 1) \in \mathcal{R}^*$. Secondly, if $(R_1, 1) \in \mathcal{R}^*$, then $R_1 \leq \Gamma'(R_1)$, so that $\Gamma'(R_1) - R_1 \geq 0$. Since $\Gamma'(R) - R$ is a non-increasing function of R , we conclude that $R_1 \leq C^*$. Thus C^* is the maximum of those R_1 in which $(R_1, 1) \in \mathcal{R}^*$, i.e., C^* is the secrecy capacity C'_s in the model of Figure 2 with memoryless side information. By using the Formula (2.13), and the non-increasing property of $\Gamma'(\cdot)$ (see Lemma 1 (iii)), we get Equation (2.14). The proof is completed.

- Note that in Equation (2.14), we have $C'_s \leq \Gamma'(0)$, which implies that $C'_s \leq \max(I(U; Y) - I(U; Z))$. Also note that for the causal model, the secrecy capacity satisfies $\max(I(U; Y) - I(U; Z)) \leq C_s^c \leq \max I(U; Y)$ (see Equation 2.7).

Then, it is easy to see that the memoryless manner for the encoder can not help to obtain the same secrecy capacity as that of the wiretap channel with causal side information.

3. Proof of Theorem 2, Theorem 4 and Converse Half of Theorem 5

3.1. Proof of Theorem 2

Suppose (R, d) is achievable, i.e., for any given $\epsilon > 0$, there exists an encoder-decoder (N, k, Δ, P_e) such that

$$\frac{H_{S^k}}{N} \geq R - \epsilon, \Delta \geq d - \epsilon, P_e \leq \epsilon$$

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

$$0 \leq d \leq 1 \quad (3.1)$$

$$0 \leq R \leq I(U; Y) - I(U; V) \quad (3.2)$$

$$Rd \leq I(U; Y) - I(K; Z|A) \quad (3.3)$$

3.1.1. Proof of Equation (3.1)

$$d - \epsilon \leq \Delta = \frac{H(S^k|Z^N)}{H(S^k)} \leq \frac{H(S^k)}{H(S^k)} = 1$$

Letting $\epsilon \rightarrow 0$, we have $d \leq 1$.

3.1.2. Proof of Equations (3.2) and (3.3)

The Formulas (3.2) and (3.3) are proved by Lemma 2, see the following.

Lemma 2 The random vectors S^k , Y^N , Z^N and the random variables U , K , A , Y , Z of Theorem 2 satisfy:

$$\frac{1}{N}H(S^k) \leq I(U; Y) - I(U; V) + \delta(P_e) \quad (3.4)$$

$$\frac{1}{N}H(S^k|Z^N) \leq I(U; Y) - I(K; Z|A) + \delta(P_e) \quad (3.5)$$

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

Substituting $H(S^k) = kH_S$, $\frac{H(S^k|Z^N)}{H(S^k)} = \Delta$ and Equation (2.3) into Equations (3.4) and (3.5), it is easy to see that

$$R - \epsilon \leq I(U; Y) - I(U; V) + \delta(\epsilon) \quad (3.6)$$

$$(R - \epsilon)(d - \epsilon) \leq I(U; Y) - I(K; Z|A) + \delta(\epsilon) \quad (3.7)$$

Letting $\epsilon \rightarrow 0$ and using the fact that $\delta(\epsilon) \rightarrow 0$ as $\epsilon \rightarrow 0$, the Formulas (3.2) and (3.3) are obtained.

It remains to prove Lemma 2, see the following.

Proof 2 (Proof of Lemma 2) The Formula (3.4) is from Equations (3.8), (3.10) and (3.22). The Formula (3.5) is proved by Equations (3.9), (3.10), (3.14), (3.22) and (3.26).

<Part i> We begin with the left parts of the inequalities Equations (3.4) and (3.5), see the following.

$$\begin{aligned}\frac{1}{N}H(S^k) &= \frac{1}{N}(I(S^k; Y^N) + H(S^k|Y^N)) \\ &\stackrel{(1)}{\leq} \frac{1}{N}(I(S^k; Y^N) + k\delta(P_e)) \\ &\leq \frac{1}{N}I(S^k; Y^N) + \delta(P_e)\end{aligned}\quad (3.8)$$

$$\begin{aligned}\frac{1}{N}H(S^k|Z^N) &\stackrel{(2)}{\leq} \frac{1}{N}(H(S^k|Z^N) + k\delta(P_e) - H(S^k|Y^N)) \\ &= \frac{1}{N}(H(S^k|Z^N) + k\delta(P_e) - H(S^k|Y^N) + H(S^k) - H(S^k)) \\ &= \frac{1}{N}(I(S^k; Y^N) - I(S^k; Z^N)) + \delta(P_e)\end{aligned}\quad (3.9)$$

where (1) and (2) follow from the Fano's inequality.

<Part ii> The character $\frac{1}{N}I(S^k; Y^N)$ in Formulas (3.8) and (3.9) can be bounded by Equation (3.10), see the following.

$$\begin{aligned}\frac{1}{N}I(S^k; Y^N) &\stackrel{(a)}{=} \frac{1}{N}(I(S^k; Y^N) - I(S^k; V^N)) \\ &= \frac{1}{N} \sum_{i=1}^N (I(S^k; Y_i|Y^{i-1}) - I(S^k; V_i|V_{i+1}^N)) \\ &= \frac{1}{N} \sum_{i=1}^N (H(Y_i|Y^{i-1}) - H(Y_i|Y^{i-1}, S^k) - H(Y_i|Y^{i-1}, S^k, V_{i+1}^N) + H(Y_i|Y^{i-1}, S^k, V_{i+1}^N) \\ &\quad - I(S^k; V_i|V_{i+1}^N)) \\ &= \frac{1}{N} \sum_{i=1}^N (I(Y_i; S^k, V_{i+1}^N|Y^{i-1}) - I(Y_i; V_{i+1}^N|Y^{i-1}, S^k) - I(S^k; V_i|V_{i+1}^N)) \\ &= \frac{1}{N} \sum_{i=1}^N (I(Y_i; S^k, V_{i+1}^N|Y^{i-1}) - I(Y_i; V_{i+1}^N|Y^{i-1}, S^k) - H(V_i|V_{i+1}^N) \\ &\quad + H(V_i|V_{i+1}^N, S^k) - H(V_i|V_{i+1}^N, S^k, Y^{i-1}) + H(V_i|V_{i+1}^N, S^k, Y^{i-1})) \\ &= \frac{1}{N} \sum_{i=1}^N (I(Y_i; S^k, V_{i+1}^N|Y^{i-1}) - I(Y_i; V_{i+1}^N|Y^{i-1}, S^k) - I(V_i; S^k, Y^{i-1}|V_{i+1}^N) \\ &\quad + I(V_i; Y^{i-1}|V_{i+1}^N, S^k)) \\ &\stackrel{(b)}{=} \frac{1}{N} \sum_{i=1}^N (I(Y_i; S^k, V_{i+1}^N|Y^{i-1}) - I(V_i; S^k, Y^{i-1}|V_{i+1}^N)) \\ &= \frac{1}{N} \sum_{i=1}^N (H(Y_i|Y^{i-1}) - H(Y_i|Y^{i-1}, S^k, V_{i+1}^N) - H(V_i|V_{i+1}^N) + H(V_i|V_{i+1}^N, S^k, Y^{i-1})) \\ &\stackrel{(c)}{\leq} \frac{1}{N} \sum_{i=1}^N (H(Y_i) - H(Y_i|Y^{i-1}, S^k, V_{i+1}^N) - H(V_i) + H(V_i|V_{i+1}^N, S^k, Y^{i-1})) \\ &= \frac{1}{N} \sum_{i=1}^N (I(Y_i; S^k, V_{i+1}^N, Y^{i-1}) - I(V_i; S^k, Y^{i-1}, V_{i+1}^N))\end{aligned}\quad (3.10)$$

Formula (a) follows from the fact that S^k is independent of V^N .

Formula (b) is from

$$\sum_{i=1}^N I(Y_i; V_{i+1}^N | Y^{i-1}, S^k) = \sum_{i=1}^N I(V_i; Y^{i-1} | V_{i+1}^N, S^k) \quad (3.11)$$

Formula (c) follows from that V^N is composed of N i.i.d. random variables.

Proof 3 (Proof of Equation (3.11)) Since the left part of Equation (3.11) is equal to

$$\sum_{i=1}^N I(Y_i; V_{i+1}^N | Y^{i-1}, S^k) = \sum_{i=1}^N \sum_{j=i+1}^N I(Y_i; V_j | Y^{i-1}, S^k, V_{j+1}^N) \quad (3.12)$$

and the right part of Equation (3.11) is equal to

$$\begin{aligned} \sum_{i=1}^N I(V_i; Y^{i-1} | V_{i+1}^N, S^k) &= \sum_{i=1}^N \sum_{j=1}^{i-1} I(V_i; Y_j | V_{i+1}^N, S^k, Y^{j-1}) \\ &= \sum_{j=1}^N \sum_{i=1}^{j-1} I(V_j; Y_i | S^k, V_{j+1}^N, Y^{i-1}) \\ &= \sum_{i=1}^N \sum_{j=i+1}^N I(V_j; Y_i | S^k, V_{j+1}^N, Y^{i-1}) \end{aligned} \quad (3.13)$$

The Formula (3.11) is verified by Equations (3.12) and (3.13).

<Part iii> The character $\frac{1}{N}I(S^k; Z^N)$ in Formula (3.9) can be bounded by the following Equation (3.14).

$$\begin{aligned} \frac{1}{N}I(S^k; Z^N) &\stackrel{(1)}{=} \frac{1}{N}(I(S^k; Z^N) - I(S^k; V^N)) \\ &= \frac{1}{N} \sum_{i=1}^N (I(S^k; Z_i | Z^{i-1}) - I(S^k; V_i | V_{i+1}^N)) \\ &= \frac{1}{N} \sum_{i=1}^N (H(Z_i | Z^{i-1}) - H(Z_i | Z^{i-1}, S^k) - H(Z_i | Z^{i-1}, S^k, V_{i+1}^N) + H(Z_i | Z^{i-1}, S^k, V_{i+1}^N) \\ &\quad - I(S^k; V_i | V_{i+1}^N)) \\ &= \frac{1}{N} \sum_{i=1}^N (I(Z_i; S^k, V_{i+1}^N | Z^{i-1}) - I(Z_i; V_{i+1}^N | Z^{i-1}, S^k) - I(S^k; V_i | V_{i+1}^N)) \\ &= \frac{1}{N} \sum_{i=1}^N (I(Z_i; S^k, V_{i+1}^N | Z^{i-1}) - I(Z_i; V_{i+1}^N | Z^{i-1}, S^k) - H(V_i | V_{i+1}^N) \\ &\quad + H(V_i | V_{i+1}^N, S^k) - H(V_i | V_{i+1}^N, S^k, Z^{i-1}) + H(V_i | V_{i+1}^N, S^k, Z^{i-1})) \\ &= \frac{1}{N} \sum_{i=1}^N (I(Z_i; S^k, V_{i+1}^N | Z^{i-1}) - I(Z_i; V_{i+1}^N | Z^{i-1}, S^k) - I(V_i; S^k, Z^{i-1} | V_{i+1}^N) \\ &\quad + I(V_i; Z^{i-1} | V_{i+1}^N, S^k)) \end{aligned}$$

$$\begin{aligned}
&\stackrel{(2)}{=} \frac{1}{N} \sum_{i=1}^N (I(Z_i; S^k, V_{i+1}^N | Z^{i-1}) - I(V_i; S^k, Z^{i-1} | V_{i+1}^N)) \\
&= \frac{1}{N} \sum_{i=1}^N (H(Z_i | Z^{i-1}) - H(Z_i | Z^{i-1}, S^k, V_{i+1}^N) - H(V_i | V_{i+1}^N) + H(V_i | V_{i+1}^N, S^k, Z^{i-1})) \\
&\stackrel{(3)}{=} \frac{1}{N} \sum_{i=1}^N (H(Z_i | Z^{i-1}) - H(Z_i | Z^{i-1}, S^k, V_{i+1}^N) - H(V_i) + H(V_i | V_{i+1}^N, S^k, Z^{i-1})) \\
&\geq \frac{1}{N} \sum_{i=1}^N (H(Z_i | Z^{i-1}) - H(Z_i | Z^{i-1}, S^k, V_{i+1}^N) - H(V_i) + H(V_i | V_{i+1}^N, S^k, Z^{i-1}, Y^{i-1})) \\
&\stackrel{(4)}{=} \frac{1}{N} \sum_{i=1}^N (H(Z_i | Z^{i-1}) - H(Z_i | Z^{i-1}, S^k, V_{i+1}^N) - H(V_i) + H(V_i | V_{i+1}^N, S^k, Y^{i-1})) \quad (3.14)
\end{aligned}$$

Formula (1) is from the fact that S^k is independent of V^N .

Formula (2) follows from

$$\sum_{i=1}^N I(Z_i; V_{i+1}^N | Z^{i-1}, S^k) = \sum_{i=1}^N I(V_i; Z^{i-1} | V_{i+1}^N, S^k) \quad (3.15)$$

Formula (3) is from the fact that V_i is independent of V_{i+1}^N .

Formula (4) is from the Markov chain $V_i \rightarrow (V_{i+1}^N, S^k, Y^{i-1}) \rightarrow Z^{i-1}$.

Proof 4 (Proof of Equation (3.15)) Since the left part of Equation (3.15) is equal to

$$\sum_{i=1}^N I(Z_i; V_{i+1}^N | Z^{i-1}, S^k) = \sum_{i=1}^N \sum_{j=i+1}^N I(Z_i; V_j | Z^{i-1}, S^k, V_{j+1}^N) \quad (3.16)$$

and the right part of Equation (3.15) is equal to

$$\begin{aligned}
\sum_{i=1}^N I(V_i; Z^{i-1} | V_{i+1}^N, S^k) &= \sum_{i=1}^N \sum_{j=1}^{i-1} I(V_i; Z_j | V_{i+1}^N, S^k, Z^{j-1}) \\
&= \sum_{j=1}^N \sum_{i=1}^{j-1} I(V_j; Z_i | S^k, V_{j+1}^N, Z^{i-1}) \\
&= \sum_{i=1}^N \sum_{j=i+1}^N I(V_j; Z_i | S^k, V_{j+1}^N, Z^{i-1}) \quad (3.17)
\end{aligned}$$

The Formula (3.15) is verified by Equations (3.16) and (3.17).

<Part iv>(single letter) To complete the proof, we introduce an random variable J , which is independent of S^k , X^N , V^N , Y^N and Z^N . Furthermore, J is uniformly distributed over $\{1, 2, \dots, N\}$.

Define

$$A = (Z^{J-1}, J) \quad (3.18)$$

$$K = (Z^{J-1}, V_{J+1}^N, S^k, J) \quad (3.19)$$

$$U = (Y^{J-1}, V_{J+1}^N, S^k, J) \quad (3.20)$$

$$X = X_J, Y = Y_J, Z = Z_J, V = V_J \quad (3.21)$$

<Part v> Then Equation (3.10) can be rewritten as

$$\begin{aligned} \frac{1}{N} I(S^k; Y^N) &\leq \frac{1}{N} \sum_{i=1}^N (I(Y_i; S^k, V_{i+1}^N, Y^{i-1}) - I(V_i; S^k, Y^{i-1}, V_{i+1}^N)) \\ &= \frac{1}{N} \sum_{i=1}^N (I(Y_i; S^k, V_{i+1}^N, Y^{i-1} | J = i) - I(V_i; S^k, Y^{i-1}, V_{i+1}^N | J = i)) \\ &= I(Y_J; S^k, V_{J+1}^N, Y^{J-1} | J) - I(V_J; S^k, Y^{J-1}, V_{J+1}^N | J) \\ &\stackrel{(a)}{\leq} I(Y_J; S^k, V_{J+1}^N, Y^{J-1}, J) - I(V_J; S^k, Y^{J-1}, V_{J+1}^N, J) \\ &\leq I(U; Y) - I(U; V) \end{aligned} \quad (3.22)$$

where (a) is from the fact that V_J is independent of J , i.e., $p(V_J = v, J = i) = p(V_J = v)p(J = i)$.

Proof 5 (Proof of $p(V_J = v, J = i) = p(V_J = v)p(J = i)$) Since V^N is the output of a discrete memoryless source $p_V(v)$, then we have

$$p(V_i = v) = p(V = v) \quad (3.23)$$

From <Part iv>, we know that the random variable J is independent of V^N , and therefore,

$$\begin{aligned} p(V_J = v, J = i) &= p(V_i = v, J = i) \\ &= p(V_i = v)p(J = i) \\ &\stackrel{(1)}{=} p(V = v)p(J = i) \end{aligned} \quad (3.24)$$

where (1) follows from Equation (3.23).

On the other hand, the probability $p(V_J = v)$ can be calculated as follows,

$$\begin{aligned} p(V_J = v) &= \sum_{i=1}^N p(V_J = v, J = i) = \sum_{i=1}^N p(V_i = v, J = i) \\ &\stackrel{(a)}{=} \sum_{i=1}^N p(V_i = v)p(J = i) \\ &\stackrel{(b)}{=} \sum_{i=1}^N p(V = v)p(J = i) \\ &= p(V = v) \sum_{i=1}^N p(J = i) = p(V = v) \end{aligned} \quad (3.25)$$

where (a) is from the fact that J is independent of V^N , the Formula (b) is from Equation (3.23).

By using Equations (3.24) and (3.25), it is easy to verify that V_J is independent of J , completing the proof.

<Part vi> Analogously, Equation (3.14) can be rewritten as

$$\begin{aligned}
 \frac{1}{N}I(S^k; Z^N) &= \frac{1}{N} \sum_{i=1}^N (H(Z_i|Z^{i-1}) - H(Z_i|Z^{i-1}, S^k, V_{i+1}^N) - H(V_i) + H(V_i|V_{i+1}^N, S^k, Y^{i-1})) \\
 &= \frac{1}{N} \sum_{i=1}^N (H(Z_i|Z^{i-1}, J=i) - \\
 &\quad H(Z_i|Z^{i-1}, S^k, V_{i+1}^N, J=i) - H(V_i|J=i) + H(V_i|V_{i+1}^N, S^k, Y^{i-1}, J=i)) \\
 &= {}^{(a)} H(Z_J|Z^{J-1}, J) - H(Z_J|Z^{J-1}, S^k, V_{J+1}^N, J) - H(V_J) + H(V_J|V_{J+1}^N, S^k, Y^{J-1}, J) \\
 &= H(Z|A) - H(Z|K, A) - H(V) + H(V|U) \\
 &= I(K; Z|A) - I(U; V)
 \end{aligned} \tag{3.26}$$

where (a) follows from the fact that V_J is independent of J .

Substituting Equations (3.22) and (3.26) into Equations (3.8) and (3.9), Lemma 2 is proved.

The Markov chains $(U, K, A) \rightarrow (X, V) \rightarrow Y \rightarrow Z$ and $(K, A) \rightarrow U \rightarrow Y \rightarrow Z$ are easily verified by Equations (3.18), (3.19), (3.20) and (3.21).

The proof of Theorem 2 is completed.

3.2. Proof of Theorem 4

Suppose (R, d) is achievable, i.e., for any given $\epsilon > 0$, there exists an encoder-decoder (N, k, Δ, P_e) such that

$$\frac{H_{S^k}}{N} \geq R - \epsilon, \Delta \geq d - \epsilon P_e \leq \epsilon$$

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

$$0 \leq d \leq 1 \tag{3.27}$$

$$0 \leq R \leq I(U; Y) \tag{3.28}$$

$$Rd \leq I(U; Y) - I(K; Z|A) \tag{3.29}$$

The Formula (3.27) is from

$$d - \epsilon \leq \Delta = \frac{H(S^k|Z^N)}{H(S^k)} \leq \frac{H(S^k)}{H(S^k)} = 1$$

. Letting $\epsilon \rightarrow 0$, we have $d \leq 1$.

Since the model of Figure 2 with causal side information is a special case of the model of Figure 2 with noncausal side information, the Formulas (3.28) and (3.29) are obtained from Equations (3.2) and (3.3), respectively, see the following.

Proof 6 (Proof of Equation (3.28)) The parameter R of Equation (3.28) can be written as follows,

$$\begin{aligned}
 R - \epsilon &\leq \frac{H(S^k)}{N} \\
 &\leq {}^{(a)} \delta(P_e) + \frac{1}{N} \sum_{i=1}^N (H(Y_i) - H(Y_i|Y^{i-1}, S^k, V_{i+1}^N) - H(V_i) + H(V_i|Y^{i-1}, S^k, V_{i+1}^N)) \\
 &\leq {}^{(b)} \delta(\epsilon) + H(Y) - H(Y|U) \\
 &= I(U; Y) + \delta(\epsilon)
 \end{aligned} \tag{3.30}$$

where (a) follows from Equations (3.8) and (3.10), and the Formula (b) is from the definitions of Y , U , see Equations (3.20) and (3.21), and V_i is independent of $(Y^{i-1}, S^k, V_{i+1}^N)$.

Letting $\epsilon \rightarrow 0$, the proof of Equation (3.28) is completed.

Proof 7 (Proof of Equation (3.29)) The parameter Rd of Equation (3.29) satisfies

$$\begin{aligned}
 (R - \epsilon)(d - \epsilon) &\leq \frac{H(S^k|Z^N)}{N} \\
 &\stackrel{(a)}{\leq} \frac{1}{N} I(S^k; Y^N) - \frac{1}{N} I(S^k; Z^N) + \delta(P_e) \\
 &\stackrel{(b)}{\leq} \frac{1}{N} \sum_{i=1}^N (I(Y_i; S^k, V_{i+1}^N, Y^{i-1}) - I(V_i; S^k, Y^{i-1}, V_{i+1}^N) - \\
 &\quad H(Z_i|Z^{i-1}) + H(Z_i|Z^{i-1}, S^k, V_{i+1}^N) + H(V_i) - H(V_i|V_{i+1}^N, S^k, Y^{i-1})) + \delta(P_e) \\
 &\stackrel{(c)}{=} \frac{1}{N} \sum_{i=1}^N (I(Y_i; S^k, V_{i+1}^N, Y^{i-1}) - H(Z_i|Z^{i-1}) + H(Z_i|Z^{i-1}, S^k, V_{i+1}^N) + \delta(P_e)) \\
 &\stackrel{(d)}{\leq} I(U; Y) - I(K; Z|A) + \delta(\epsilon) \tag{3.31}
 \end{aligned}$$

where (a) follows from Equation (3.9), the Formula (b) is from Equations (3.10) and (3.14), the Formula (c) is from the fact that V_i is independent of $(Z^{i-1}, Y^{i-1}, S^k, V_{i+1}^N)$, the Formula (d) is from the definitions of Y , Z , U , K , A , see Equations (3.18), (3.19), (3.20) and (3.21). Letting $\epsilon \rightarrow 0$, the proof of Equation (3.29) is completed.

The proof of Theorem 4 is completed.

3.3. Converse Half of Theorem 5

In this subsection, we establish the converse theorem of Theorem 5: the region $\mathcal{C}_m$ which is composed of all achievable (R, d) pairs is contained in the set $\mathcal{R}^*$, i.e., $\mathcal{C}_m \subseteq \mathcal{R}^*$.

Suppose $(R, d) \in \mathcal{C}_m$, i.e., for any given $\epsilon > 0$, there exists an encoder-decoder (N, k, Δ, P_e) such that

$$\frac{H_{Sk}}{N} \geq R - \epsilon, \Delta \geq d - \epsilon, P_e \leq \epsilon$$

Then we will show that $(R, d) \in \mathcal{R}^*$, i.e., (R, d) satisfies the following conditions

$$0 \leq R \leq C_M, 0 \leq d \leq 1 \text{ and } Rd \leq \Gamma'(R)$$

The proof of $R \leq C_M$ and $d \leq 1$ is obvious, and it is omitted here. It only needs to prove $Rd \leq \Gamma'(R)$, see the following.

The following Lemma 3 provides a Markov chain used in the remaining of this subsection. The proof of Lemma 3 is in Appendix 5.

Lemma 3 In the model of Figure 2, the random variable Z_i and the random vectors S^k and Y^{i-1} ($1 \leq i \leq N$) form the following Markov chain:

$$Y^{i-1} \rightarrow S^k \rightarrow Z_i$$

The proof of $Rd \leq \Gamma'(R)$ is considered in the following five steps:

(i) Show that

$$H(S^k)\Delta \leq I(S^k; Y^N | Z^N) + k\delta(P_e)$$

(ii) In the right part of step (i), show that

$$I(S^k; Y^N | Z^N) \leq \sum_{n=1}^N (I(U_n; Y_n | Y^{n-1}) - I(U_n; Z_n | Y^{n-1}))$$

(iii) In the right part of step (ii), show that

$$\frac{1}{N} \sum_{n=1}^N (I(U_n; Y_n | Y^{n-1}) - I(U_n; Z_n | Y^{n-1})) \leq \Gamma'(\frac{1}{N} \sum_{n=1}^N I(U_n; Y_n | Y^{n-1}))$$

(iv) A property about the variable of the function $\Gamma'(\cdot)$ in step (iii) is

$$\frac{k}{N} (H_S - \delta(P_e)) \leq \frac{1}{N} \sum_{n=1}^N I(U_n; Y_n | Y^{n-1})$$

(v) Substituting step (ii), step (iii) and step (iv) into step (i), we have

$$Rd \leq \Gamma'(R)$$

3.3.1. Proof of Step (i)

By using Fano's inequality,

$$H(S^k | Z^N, Y^N) \leq H(S^k | Y^N) \leq k\delta(P_e) \quad (3.32)$$

where $\delta(P_e) = h(P_e) + P_e \log(|\mathcal{S}| - 1)$.

Then we have

$$\begin{aligned} H(S^k)\Delta &= H(S^k | Z^N) \\ &\leq H(S^k | Z^N) + k\delta(P_e) - H(S^k | Z^N, Y^N) \\ &= k\delta(P_e) + I(S^k; Y^N | Z^N) \end{aligned} \quad (3.33)$$

Thus, the proof of step (i) is completed.

3.3.2. Proof of Step (ii)

$$\begin{aligned} I(S^k; Y^N | Z^N) &= H(S^k | Z^N) - H(S^k | Y^N, Z^N) \\ &\stackrel{(a)}{=} H(S^k | Z^N) - H(S^k | Y^N) \\ &= I(S^k; Y^N) - I(S^k; Z^N) \\ &= H(Y^N) - H(Y^N | S^k) - H(Z^N) + H(Z^N | S^k) \\ &= \sum_{n=1}^N (H(Y_n | Y^{n-1}) - H(Y_n | Y^{n-1}, S^k) - H(Z_n | Z^{n-1}) + H(Z_n | Z^{n-1}, S^k)) \end{aligned}$$

$$\begin{aligned}
&\leq^{(b)} \sum_{n=1}^N (H(Y_n|Y^{n-1}) - H(Y_n|Y^{n-1}, S^k, V_{n+1}^N) - H(Z_n|Z^{n-1}, Y^{n-1}) + H(Z_n|S^k)) \\
&=^{(c)} \sum_{n=1}^N (H(Y_n|Y^{n-1}) - H(Y_n|Y^{n-1}, S^k, V_{n+1}^N) - H(Z_n|Y^{n-1}) + H(Z_n|S^k, Y^{n-1})) \\
&=^{(d)} \sum_{n=1}^N (H(Y_n|Y^{n-1}) - H(Y_n|Y^{n-1}, S^k, V_{n+1}^N) - H(Z_n|Y^{n-1}) + H(Z_n|S^k, Y^{n-1}, V_{n+1}^N)) \\
&=^{(e)} \sum_{n=1}^N (H(Y_n|Y^{n-1}) - H(Y_n|Y^{n-1}, U_n) - H(Z_n|Y^{n-1}) + H(Z_n|U_n, Y^{n-1})) \\
&= \sum_{n=1}^N (I(U_n; Y_n|Y^{n-1}) - I(U_n; Z_n|Y^{n-1})) \tag{3.34}
\end{aligned}$$

where Formula (a) follows from $S^k \rightarrow Y^N \rightarrow Z^N$, see Lemma 3 in Appendix 5. Formula (b) follows from the fact that V_{n+1}^N is independent of Y_n, Y^{n-1}, S^k . Formula (c) follows from $Z^{n-1} \rightarrow Y^{n-1} \rightarrow Z_n$ and $Y^{n-1} \rightarrow S^k \rightarrow Z_n$ (see Lemma 2). Formula (d) follows from the fact that V_{n+1}^N is independent of Z_n, Y^{n-1}, S^k . Formula (e) follows from the definition that $U_n = (S^k, Y^{n-1}, V_{n+1}^N)$, and this is coincident with the definition of U used in the converse proof of Equation (2.9).

The proof of step (ii) is completed.

3.3.3. Proof of Step (iii)

The proof of step (iii) is considered in two parts. The first part is for some definitions, and the second part is for the main proof.

- For $n = 2, 3, \dots, N$, and any $y^{n-1} \in \mathcal{Y}^{n-1}$, let

$$\alpha_n(y^{n-1}) = I(U_n; Y_n|Y^{n-1} = y^{n-1}) \tag{3.35}$$

Denote

$$\alpha_1 = I(U_1; Y_1) \tag{3.36}$$

It follows from the definition of $\rho(R)$ in Equation (2.10) that the distribution p_1 , defined by

$$p_1 = Pr\{X_1 = x|U_1 = u, V_1 = v\}Pr\{U_1 = u\}, u \in \mathcal{U}, x \in \mathcal{X}, v \in \mathcal{V} \tag{3.37}$$

belongs to $\rho(\alpha_1)$. Similarly, for $2 \leq n \leq N$, let

$$p_{n,y^{n-1}} = Pr\{X_n = x|U_n = u, V_n = v, Y^{n-1} = y^{n-1}\}Pr\{U_n = u|Y^{n-1} = y^{n-1}\} \tag{3.38}$$

where $u \in \mathcal{U}, x \in \mathcal{X}, v \in \mathcal{V}, y^{n-1} \in \mathcal{Y}^{n-1}$. Then it is easy to see that $p_{n,y^{n-1}} \in \rho(\alpha_n(y^{n-1}))$. Thus, from the definition of $\Gamma'(R)$ in Equation (2.11),

$$\Gamma'(\alpha_1) \geq I(U_1; Y_1) - I(U_1; Z_1) \tag{3.39}$$

and for $2 \leq n \leq N, y^{n-1} \in \mathcal{Y}^{n-1}$,

$$\Gamma'(\alpha_n(y^{n-1})) \geq I(U_n; Y_n|Y^{n-1} = y^{n-1}) - I(U_n; Z_n|Y^{n-1} = y^{n-1}) \tag{3.40}$$

- By using the Formulas (3.35) and (3.40), the proof of step (iii) is as follows,

$$\begin{aligned}
& \frac{1}{N} \sum_{n=1}^N (I(U_n; Y_n | Y^{n-1}) - I(U_n; Z_n | Y^{n-1})) \\
&= \frac{1}{N} \sum_{n=1}^N \sum_{y^{n-1} \in \mathcal{Y}^{n-1}} Pr\{Y^{n-1} = y^{n-1}\} (I(U_n; Y_n | Y^{n-1} = y^{n-1}) - I(U_n; Z_n | Y^{n-1} = y^{n-1})) \\
&\stackrel{(a)}{\leq} \frac{1}{N} \sum_{n=1}^N \sum_{y^{n-1} \in \mathcal{Y}^{n-1}} Pr\{Y^{n-1} = y^{n-1}\} \Gamma'(\alpha_n(y^{n-1})) \\
&\stackrel{(b)}{\leq} \Gamma' \left(\frac{1}{N} \sum_{n=1}^N \sum_{y^{n-1} \in \mathcal{Y}^{n-1}} Pr\{Y^{n-1} = y^{n-1}\} \alpha_n(y^{n-1}) \right) \\
&\stackrel{(c)}{=} \Gamma' \left(\frac{1}{N} \sum_{n=1}^N I(U_n; Y_n | Y^{n-1}) \right)
\end{aligned}$$

where Formula (a) follows from the inequality Equation (3.40). Formula (b) follows from the concavity of $\Gamma'(R)$ [Lemma 1 (ii)]. Formula (c) follows from the definition Equation (3.35).

The proof of step (iii) is completed.

3.3.4. Proof of Step (iv)

$$\begin{aligned}
\frac{1}{N} \sum_{n=1}^N I(U_n; Y_n | Y^{n-1}) &= \frac{1}{N} \sum_{n=1}^N (H(U_n | Y^{n-1}) - H(U_n | Y^{n-1}, Y_n)) \\
&\stackrel{(a)}{=} \frac{1}{N} \sum_{n=1}^N (H(V_{n+1}^N, S^k | Y^{n-1}) - H(V_{n+1}^N, S^k | Y^{n-1}, Y_n)) \\
&\stackrel{(b)}{=} \frac{1}{N} \sum_{n=1}^N (H(V_{n+1}^N) + H(S^k | V_{n+1}^N, Y^{n-1}) - H(V_{n+1}^N) - H(S^k | V_{n+1}^N, Y^{n-1}, Y_n)) \\
&= \frac{1}{N} \sum_{n=1}^N I(S^k; Y_n | V_{n+1}^N, Y^{n-1}) \\
&= \frac{1}{N} \sum_{n=1}^N (H(Y_n | V_{n+1}^N, Y^{n-1}) - H(Y_n | V_{n+1}^N, Y^{n-1}, S^k)) \\
&\stackrel{(c)}{=} \frac{1}{N} \sum_{n=1}^N (H(Y_n | Y^{n-1}) - H(Y_n | Y^{n-1}, S^k)) \\
&= \frac{1}{N} (H(Y^N) - H(Y^N | S^k)) \\
&= \frac{1}{N} I(S^k; Y^N) \\
&= \frac{1}{N} (H(S^k) - H(S^k | Y^N)) \\
&\stackrel{(d)}{\geq} \frac{1}{N} (kH_S - k\delta(P_e))
\end{aligned}$$

where Formula (a) follows from the definition that $U_n = (S^k, Y^{n-1}, V_{n+1}^N)$. Formulas (b) and (c) follow from the fact that V_{n+1}^N is independent of Y_n, Y^{n-1}, S^k, V_n . Formula (d) follows from $H(S^k) = kH_S$ and the Fano's inequality.

The proof of step (iv) is completed.

3.3.5. Proof of Step (v)

Substituting step (ii), step (iii), step (iv) into step (i), and using the non-increasing property of $\Gamma'(R)$ [Lemma 1 (iii)], it is easy to see that

$$\frac{kH_S\Delta - k\delta(P_e)}{N} \leq \Gamma'\left(\frac{kH_S - k\delta(P_e)}{N}\right) \quad (3.41)$$

By using the definition of achievable (R, d) pair, i.e., $\Delta \geq d - \epsilon$, $\frac{H_S k}{N} \geq R - \epsilon$, $P_e \leq \epsilon$, and the fact that $\delta(P_e) \leq \delta(\epsilon)$, we know from Equation (3.41) that

$$\begin{aligned} (R - \epsilon)(d - \epsilon) - \delta(\epsilon) &\leq \frac{kH_S\Delta - k\delta(P_e)}{N} \\ &\leq \Gamma'\left(\frac{kH_S - k\delta(P_e)}{N}\right) \\ &\stackrel{(a)}{\leq} \Gamma'(R - \epsilon - \delta(\epsilon)) \end{aligned} \quad (3.42)$$

where the formula (a) follows from the non-increasing property of $\Gamma'(R)$ [Lemma 1 (iii)]. In Equation (3.42), letting $\epsilon \rightarrow 0$ and invoking the continuity of $\Gamma'(R)$ [Lemma 1 (iv)] yield $Rd \leq \Gamma'(R)$. The proof of step (v) is completed.

The converse part of Theorem 5 is proved.

4. Proof of Theorem 3 and Direct Half of Theorem 5

In this section, all logarithms are taken to the base 2.

4.1. Proof of Theorem 3

In this subsection, we will show the achievability of the region $\mathcal{R}_{ci}$, and we only need to prove that the pair $(R, d = \frac{I(U;Y) - I(U;Z)}{R})$ is achievable.

4.1.1. Coding Construction

Given the pair $(R, d = \frac{I(U;Y) - I(U;Z)}{R})$, let k and N satisfy $\frac{H_S k}{N} = R = I(U;Y) - \gamma$, where γ satisfies $0 \leq \gamma \leq^{(a)} I(U;Z)$, and (a) is from $d = \frac{I(U;Y) - I(U;Z)}{R} \leq 1$ and $R = I(U;Y) - \gamma$.

A separated source-channel coding method is provided. The source encoder is a mapping

$$S^k \rightarrow \mathcal{W} = \{1, 2, \dots, 2^{kH_S(1+k^{-\frac{1}{4}})}\}$$

with the input S^k and the output W .

Generate a random code-book composed of $2^{N(I(U;Y) - \gamma_1)}$ codewords of u^N (γ_1 is a small fixed positive number), and each of them is i.i.d. generated according to $p_U(u)$. Divide the code-book into $2^{kH_S(1+k^{-\frac{1}{4}})}$

bins, and each bin corresponds to a specific value in $\mathcal{W}$. There are $2^{NI(U;Y) - N\gamma_1 - kH_S(1+k^{-\frac{1}{4}})}$ codewords in each bin. Note that

$$\begin{aligned}
 & NI(U;Y) - N\gamma_1 - kH_S(1+k^{-\frac{1}{4}}) = NI(U;Y) - N\gamma_1 - NR - NRk^{-\frac{1}{4}} \\
 &= NI(U;Y) - N\gamma_1 - N(I(U;Y) - \gamma) - NRk^{-\frac{1}{4}} \\
 &= N\gamma - N\gamma_1 - NRk^{-\frac{1}{4}} \\
 &\stackrel{(1)}{\leq} NI(U;Z) - N\gamma_1 - NRk^{-\frac{1}{4}} \leq NI(U;Z)
 \end{aligned} \tag{4.1}$$

where (1) is from $0 \leq \gamma \leq I(U;Z)$. For a given w , randomly choose a codeword in bin w to transmit.

The x^N is generated according to a new discrete memoryless channel (DMC) with inputs u^N, v^N , and output x^N . The transition probability of this new DMC is $p_{X|U,V}(x|u, v)$. Furthermore, we have

$$p_{X^N|U^N, V^N}(x^N|u^N, v^N) = \prod_{i=1}^N p_{X|U,V}(x_i|u_i, v_i) \tag{4.2}$$

For given y^N , the legitimate receiver tries to find a sequence u^N such that $(u^N, y^N) \in T_{UY}^N(\epsilon^{**})$. If there exists one sequence, put out the corresponding index $\hat{w}$ of the bin, else declare a decoding error. Then, by using the mapping $\mathcal{S}^k \rightarrow \mathcal{W}$, put out the corresponding source $\hat{s}^k$.

4.1.2. Proof of $\frac{H_{Sk}}{N} \geq R - \epsilon$, $P_e \leq \epsilon$, and $\Delta \geq d - \epsilon$

By using the above definitions, it is easy to verify that $\frac{H_{Sk}}{N} = R \geq R - \epsilon$.

Then, observing the construction of U^N , it is easy to see that the codewords of U^N is upper-bounded by $2^{NI(U;Y)}$. Since the main channel can be viewed as an ordinary DMC with input U^N and output Y^N , from the standard channel coding theorem, we have $Pr\{W \neq \hat{W}\} \rightarrow 0$ as the coding length $N \rightarrow \infty$. From the source coding theorem, we have $Pr\{S^k \neq \hat{S}^k\} \rightarrow 0$ as $k = \frac{NR}{H_S} \rightarrow \infty$. So we can choose sufficiently large N to satisfy $Pr\{S^k \neq \hat{S}^k\} + Pr\{W \neq \hat{W}\} \leq \epsilon$, thus $P_e \leq \epsilon$ is proved.

It remains to show that $\Delta \geq d - \epsilon$, see the following.

4.1.3. Proof of $\Delta \geq \frac{I(U;Y) - I(U;Z)}{R} - \epsilon$

Since

$$\begin{aligned}
 \frac{H(S^k)\Delta}{N} &= \frac{1}{N}H(S^k|Z^N) = \frac{1}{N}(H(S^k, Z^N) - H(Z^N)) \\
 &= \frac{1}{N}(H(S^k, Z^N, U^N) - H(U^N|S^k, Z^N) - H(Z^N)) \\
 &= \frac{1}{N}(H(Z^N|S^k, U^N) + H(S^k) + H(U^N|S^k) - H(U^N|S^k, Z^N) - H(Z^N)) \\
 &\stackrel{(a)}{=} \frac{1}{N}(H(Z^N|U^N) + H(S^k) + H(U^N|S^k) - H(U^N|S^k, Z^N) - H(Z^N)) \\
 &= \frac{1}{N}(H(S^k) + I(U^N; Z^N|S^k) - I(U^N; Z^N)) \\
 &= \frac{1}{N}(H(S^k) + H(Z^N|S^k) - H(Z^N|S^k, U^N) - I(U^N; Z^N)) \\
 &\stackrel{(b)}{\geq} \frac{1}{N}(H(S^k) + H(Z^N|S^k, W) - H(Z^N|U^N) - I(U^N; Z^N))
 \end{aligned}$$

$$\begin{aligned}
&=^{(c)} \frac{1}{N} (H(S^k) + H(Z^N|W) - H(Z^N|U^N, W) - I(U^N; Z^N)) \\
&= \frac{1}{N} (H(S^k) + I(U^N; Z^N|W) - I(U^N; Z^N)) \\
&\geq^{(d)} I(U; Y) - \gamma_1 - Rk^{-\frac{1}{4}} - \delta(\epsilon^{**}) - I(U; Z)
\end{aligned} \tag{4.3}$$

where (a), (b) and (c) follow from $S^k \rightarrow W \rightarrow U^N \rightarrow Z^N$, (d) is from $\frac{H(S^k)}{N} = R = I(U; Y) - \gamma$,

$$\begin{aligned}
H(U^N|W) &= \sum_{i=1}^M Pr\{W = i\} H(U^N|W = i) \\
&= \sum_{i=1}^M Pr\{W = i\} \log 2^{NI(U; Y) - N\gamma_1 - kH_S(1+k^{-\frac{1}{4}})} \\
&= NI(U; Y) - N\gamma_1 - kH_S(1+k^{-\frac{1}{4}}) \\
&= NI(U; Y) - N\gamma_1 - kH_S - NRk^{-\frac{1}{4}}
\end{aligned} \tag{4.4}$$

and the fact that given W and Z^N , there are $2^{NI(U; Y) - N\gamma_1 - kH_S(1+k^{-\frac{1}{4}})}$ codewords left for the wiretapper, and therefore, by using Equation (4.1) and the standard channel coding theorem, we have $H(U^N|W, Z^N) \leq N\delta(\epsilon^{**})$, where ϵ^{**} is an arbitrary small positive number. Then for sufficiently large N , choosing ϵ^{**} , γ_1 , $Rk^{-\frac{1}{4}}$ such that $\gamma_1 + Rk^{-\frac{1}{4}} + \delta(\epsilon^{**}) \leq \epsilon R$, and using $\frac{H(S^k)}{N} = R$, we have $\Delta \geq \frac{I(U; Y) - I(U; Z)}{R} - \epsilon$. The proof for $\Delta \geq d - \epsilon$ is completed.

The proof of Theorem 3 is completed.

4.2. Proof of the Direct Half of Theorem 5

In this subsection we establish the direct part of Theorem 5 (about existence), i.e., $\mathcal{R}^* \subseteq \mathcal{C}_m$. Suppose $(R, d) \in \mathcal{R}^*$, i.e., (R, d) satisfies the following conditions:

$$0 \leq R \leq C_M, 0 \leq d \leq 1 \text{ and } Rd \leq \Gamma'(R)$$

We will show that $(R, d) \in \mathcal{R}$, that is to say, (R, d) is achievable, i.e., for any given $\epsilon > 0$, there exists an encoder-decoder (N, k, Δ, P_e) such that

$$\frac{H_S k}{N} \geq R - \epsilon, \Delta \geq d - \epsilon, P_e \leq \epsilon$$

A sufficient condition of the corresponding proof is to show that the (R, d) pair satisfying

$$Rd = \Gamma'(R) \tag{4.5}$$

is achievable, see the remaining of this section. The construction of the code is introduced in Section 4.2.1. For any given $\epsilon > 0$, the proofs of $\frac{H_S k}{N} \geq R - \epsilon$ and $\Delta \geq d - \epsilon$ are given in Section 4.2.2. Section 4.2.3 is about $P_e \leq \epsilon$.

4.2.1. Code Construction

The existence of the encoder-decoder is under the sufficient condition $Rd = \Gamma'(R)$. Let k and N satisfy $\frac{H_S k}{N} = R$. Choose a probability mass function $Pr\{X^* = x|U^* = u, V = v\}Pr\{U^* = u\}$

such that $I(U^*; Y^*) \geq R$ and $I(U^*; Y^*) - I(U^*; Z^*) = \Gamma'(R)$, where U^* and V are the inputs of the channel encoder, while X^* is the output, Y^* and Z^* are the respective outputs of the main channel and the wiretap channel.

A separated source-channel coding method is provided. The source encoder is a mapping $\mathcal{S}^k \rightarrow \{1, 2, \dots, M\}$, with the input S^k and the output W . According to the specific value of W , generate M_1 codewords $\{u^N(w, m) : 1 \leq w \leq M, 1 \leq m \leq M_2\}$ i.i.d. according to $Pr\{U^* = u\}$, where $M_1 = 2^{NI(U^*; Y^*)}$, $M = 2^{kH_S(1+k^{-\frac{1}{4}})}$ and $M_2 = \frac{M_1}{M} = 2^{NI(U^*; Y^*) - kH_S(1+k^{-\frac{1}{4}})}$ (note that $\frac{1}{N} \log M_2 \leq I(U^*; Z^*) - \epsilon$, and this is from the similar argument in [1], p. 1377).

For a given w , there is a corresponding subcode $C_w = \{u^N(w, 1), \dots, u^N(w, M_2)\}$. Randomly choose a codeword $u^N(w, m)$ from C_w to transmit.

The x^N is generated according to a new discrete memoryless channel (DMC) with inputs $u^N(w, m)$, v^N , and output x^N . The transition probability of this new DMC is $p_{X^*|U^*, V}(x|u, v)$. Furthermore, we have

$$p_{X^N|U^N, V^N}(x^N|u^N(w, m), v^N) = \prod_{i=1}^N p_{X^*|U^*, V}(x_i|u_i, v_i) \quad (4.6)$$

The inputs of the main channel are x^N and v^N , while the output is y^N . In the decoding scheme, for given y^N , try to find a codeword $u^N(\hat{w}, \hat{m})$ such that $(u^N(\hat{w}, \hat{m}), y^N) \in T_{UY}^N(\epsilon^{****})$. If there is one or more such codeword, choose one and put out the corresponding $\hat{w}$. According to $\hat{w}$ and the mapping $F_D : \{1, 2, \dots, M\} \rightarrow \mathcal{S}^k$, put out the corresponding $\hat{s}^k$.

4.2.2. Proofs of $\frac{H_{S^k}}{N} \geq R - \epsilon$ and $\Delta \geq d - \epsilon$

Since $\frac{H_{S^k}}{N} = R$, it is easy to see that $\frac{H_{S^k}}{N} \geq R - \epsilon$ for any $\epsilon > 0$. It remains to show that $\Delta \geq d - \epsilon$, see the Formulas (4.7), (4.8), (4.12), (4.16) and (4.18).

$$\begin{aligned} \frac{H(S^k)\Delta}{N} &= \frac{1}{N}H(S^k|Z^N) = \frac{1}{N}(H(S^k, Z^N) - H(Z^N)) \\ &= \frac{1}{N}(H(S^k, Z^N, U^N) - H(U^N|S^k, Z^N) - H(Z^N)) \\ &= \frac{1}{N}(H(Z^N|S^k, U^N) + H(S^k) + H(U^N|S^k) - H(U^N|S^k, Z^N) - H(Z^N)) \\ &\stackrel{(a)}{=} \frac{1}{N}(H(Z^N|U^N) + H(S^k) + H(U^N|S^k) - H(U^N|S^k, Z^N) - H(Z^N)) \\ &= \frac{1}{N}(H(S^k) + I(U^N; Z^N|S^k) - I(U^N; Z^N)) \\ &= \frac{1}{N}(H(S^k) + H(Z^N|S^k) - H(Z^N|S^k, U^N) - I(U^N; Z^N)) \\ &\geq^{(b)} \frac{1}{N}(H(S^k) + H(Z^N|S^k, W) - H(Z^N|U^N) - I(U^N; Z^N)) \\ &\stackrel{(c)}{=} \frac{1}{N}(H(S^k) + H(Z^N|W) - H(Z^N|U^N, W) - I(U^N; Z^N)) \\ &= \frac{1}{N}(H(S^k) + I(U^N; Z^N|W) - I(U^N; Z^N)) \end{aligned} \quad (4.7)$$

where (a), (b) and (c) follow from $S^k \rightarrow W \rightarrow U^N \rightarrow Z^N$. Then, we will estimate the two characters $I(U^N; Z^N|W)$ and $I(U^N; Z^N)$, respectively.

•

$$I(U^N; Z^N | W) \geq \log M_2 - h(\bar{\lambda}) - \bar{\lambda} \log M_2 \quad (4.8)$$

where $\bar{\lambda} = \sum_{i=1}^M \Pr\{W = i\} \lambda_i$, and λ_i is the resulting error probability of a code $C_i = \{c^N(i, 1), \dots, c^N(i, M_2)\}$ used on the channel Q_{MW} .

Proof 8 (Proof of Equation (4.8))

$$I(U^N; Z^N | W) = H(U^N | W) - H(U^N | W, Z^N) \quad (4.9)$$

$$\begin{aligned} H(U^N | W) &= \sum_{i=1}^M \Pr\{W = i\} H(U^N | W = i) \\ &= \sum_{i=1}^M \Pr\{W = i\} \log M_2 \\ &= \log M_2 \end{aligned} \quad (4.10)$$

$$\begin{aligned} H(U^N | W, Z^N) &= \sum_{i=1}^M \Pr\{W = i\} H(U^N | W = i, Z^N) \\ &\stackrel{(a)}{\leq} \sum_{i=1}^M \Pr\{W = i\} (h(\lambda_i) + \lambda_i \log M_2) \\ &\stackrel{(b)}{\leq} h\left(\sum_{i=1}^M \Pr\{W = i\} \lambda_i\right) + \sum_{i=1}^M \Pr\{W = i\} \lambda_i \log M_2 \\ &\stackrel{(c)}{=} h(\bar{\lambda}) + \bar{\lambda} \log M_2 \end{aligned} \quad (4.11)$$

Formula (a) follows from the Fano's inequality. Formula (b) follows from the concavity of $h(\cdot)$.

Formula (c) follows from the definition of $\bar{\lambda} = \sum_{i=1}^M \Pr\{W = i\} \lambda_i$.

Substituting Equations (4.10) and (4.11) into Equation (4.9), we get Equation (4.8).

- Since U^N is composed of N i.i.d. random variables with probability mass function $\Pr\{U^* = u\}, u \in \mathcal{U}$, and V^N is available at the encoder in a memoryless case, we have

$$\begin{aligned} \frac{1}{N} I(U^N; Z^N) &= \frac{1}{N} \sum_{n=1}^N I(U_n; Z_n) \\ &= I(U^*; Z^*) \end{aligned} \quad (4.12)$$

Substituting $\frac{H(S^k)}{N} = R$, Equations (4.8) and (4.12) into Equation (4.7),

$$R\Delta \geq \frac{H(S^k)}{N} + \frac{\log M_2}{N} - \frac{h(\bar{\lambda})}{N} - \frac{\bar{\lambda} \log M_2}{N} - I(U^*; Z^*) \quad (4.13)$$

For $H(S^k) = kH_S$ and $\log M_2 = NI(U^*; Y^*) - kH_S(1 + k^{-\frac{1}{4}})$, the Formula (4.13) can be written as follows,

$$R\Delta \geq I(U^*; Y^*) - \frac{kH_S}{N} k^{-\frac{1}{4}} - \frac{h(\bar{\lambda})}{N} - \frac{\bar{\lambda} \log M_2}{N} - I(U^*; Z^*) \quad (4.14)$$

Since $\frac{kH_S}{N} = R$ and $I(U^*; Y^*) - I(U^*; Z^*) = \Gamma'(R)$, we know from Equation (4.14) that

$$R\Delta \geq \Gamma'(R) - Rk^{-\frac{1}{4}} - \frac{h(\bar{\lambda})}{N} - \frac{\bar{\lambda} \log M_2}{N} \quad (4.15)$$

$Rk^{-\frac{1}{4}} + \frac{h(\bar{\lambda})}{N} + \frac{\bar{\lambda} \log M_2}{N}$ in the right part of Equation (4.15) is estimated as follows. Since the channel Q_{MW} can be viewed as an ordinary DMC with input U^N and output Z^N , by using the similar argument in [1], p. 1377, $\bar{\lambda} \rightarrow 0$ as the coding length $N \rightarrow \infty$. So with sufficiently large N , we choose N to satisfy

$$Rk^{-\frac{1}{4}} + \frac{h(\bar{\lambda})}{N} + \frac{\bar{\lambda} \log M_2}{N} \leq \epsilon R \quad (4.16)$$

Substituting Equation (4.16) into Equation (4.15),

$$R\Delta \geq \Gamma'(R) - \epsilon R \quad (4.17)$$

Then by using Equation (4.5), Formula (4.17) can be rewritten as

$$\Delta \geq \frac{\Gamma'(R)}{R} - \epsilon = d - \epsilon \quad (4.18)$$

The proof of $\Delta \geq d - \epsilon$ is completed.

4.2.3. Proof of $P_e \leq \epsilon$

$P_e = \Pr\{S^k \neq \hat{S}^k\} \leq \Pr\{S^k \neq F_D(W)\} + \Pr\{W \neq \hat{W}\}$. Since the main channel can be viewed as an ordinary DMC with input U^N and output Y^N , from the standard channel coding theorem, we have $\Pr\{W \neq \hat{W}\} \rightarrow 0$ as the coding length $N \rightarrow \infty$. From the source coding theorem, we have $\Pr\{S^k \neq F_D(W)\} \rightarrow 0$ as $k = \frac{NR}{H_S} \rightarrow \infty$. So we can choose sufficiently large N to satisfy $\Pr\{S^k \neq F_D(W)\} + \Pr\{W \neq \hat{W}\} \leq \epsilon$, thus $P_e \leq \epsilon$. The proof is completed.

5. Conclusions

In this paper, we study the model of wiretap channel with side information. Inner and outer bounds are derived on the capacity-equivocation regions for the noncausal and causal manners (the inner bound for the noncausal manner is in fact equivalent to that of [9]), and the secrecy capacities for both manners are described and bounded. Moreover, for the case that the side information is available at the transmitter in a memoryless manner, both the capacity-equivocation region and the secrecy capacity are determined.

Acknowledgment

This work was supported by the German Research Foundation DFG, the National Basic Research Program of China under Grant 2007CB310900, and the National Natural Science Foundation of China under Grants 61271222, 60972033 and 60832001. The authors would like to thank N. Cai for his help to improve this paper, and also are grateful to the anonymous reviewers for their helpful suggestions.

Appendix

Size Constraints of the Auxiliary Random Variables in Theorem 2

By using the support lemma (see [11], p. 310), it suffices to show that the random variables U , A and K can be replaced by new ones, preserving the Markovity $(U, A, K) \rightarrow (X, V) \rightarrow Y \rightarrow Z$ and the characters $I(U; Y)$, $I(U; V)$, $H(Z|A)$, $H(Z|K)$, and furthermore, the range of the new U , A and K satisfies: $\|A\| \leq \|\mathcal{X}\|\|\mathcal{V}\|$, $\|K\| \leq \|\mathcal{X}\|^2\|\mathcal{V}\|^2$, $\|U\| \leq \|\mathcal{X}\|^2\|\mathcal{V}\|^2(\|\mathcal{X}\|\|\mathcal{V}\| + 1)$. The proof of which is in the reminder of this section.

- **(Proof of $\|A\| \leq \|\mathcal{X}\|\|\mathcal{V}\|$)**

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

$$f_{XV}(\bar{p}) = p_{XV}(x, v), f_Z(\bar{p}) = H(Z)$$

Since there are $\|\mathcal{X}\|\|\mathcal{V}\| - 1$ functions of $f_{XV}(\bar{p})$, the total number of the continuous scalar functions of $\bar{p}$ is $\|\mathcal{X}\|\|\mathcal{V}\|$.

Let $\bar{p}_{XV|A} = Pr\{X = x, V = v|A = a\}$. With these distributions $\bar{p}_{XV|A}$, we have

$$p_{XV}(x, v) = \sum_{a \in \mathcal{A}} p(A = a) f_{XV}(\bar{p}_{XV|A}) \quad (1)$$

$$H(Z|A) = \sum_{a \in \mathcal{A}} p(A = a) f_Z(\bar{p}_{XV|A}) \quad (2)$$

According to the support lemma ([11], p. 310), the random variable A can be replaced by new ones such that the new A takes at most $\|\mathcal{X}\|\|\mathcal{V}\|$ different values and the expressions in Equations (1) and (2) are preserved.

- **(Proof of $\|K\| \leq \|\mathcal{X}\|^2\|\mathcal{V}\|^2$)**

Once the alphabet of A is fixed, we apply similar arguments to bound the alphabet of K , see the following. Let $\bar{p} = p_{XV}(x, v)$, define the following continuous scalar functions of $\bar{p}$:

$$f_{XV}(\bar{p}) = p_{XV}(x, v), f_Z(\bar{p}) = H(Z)$$

Since there are $\|\mathcal{X}\|\|\mathcal{V}\| - 1$ functions of $f_{XV}(\bar{p})$, the total number of the continuous scalar functions of $\bar{p}$ is $\|\mathcal{X}\|\|\mathcal{V}\|$.

Let $\bar{p}_{XV|K} = Pr\{X = x, V = v|K = k\}$. With these distributions $\bar{p}_{XV|K}$, we have

$$p_{XV|A}(x, v|a) = \sum_{k \in \mathcal{K}} p(K = k|A = a) f_{XV}(\bar{p}_{XV|K}) \quad (3)$$

$$H(Z|K, A) = \sum_{k \in \mathcal{K}} p(K = k|A = a) f_Z(\bar{p}_{XV|K}) \quad (4)$$

According to the support lemma ([11], p. 310), for every fixed a , the random variable K can be replaced by new ones such that the new K takes at most $\|\mathcal{X}\|\|\mathcal{V}\| + 1$ different values and the expressions Equations (3) and (4) are preserved. Therefore, $\|K\| \leq \|\mathcal{X}\|^2\|\mathcal{V}\|^2$ is proved.

- **(Proof of $\|\mathcal{U}\| \leq \|\mathcal{X}\|^2 \|\mathcal{V}\|^2 (\|\mathcal{X}\| \|\mathcal{V}\| + 1)$)**

Once the the alphabet of K is fixed, we apply similar arguments to bound the alphabet of U , see the following. Define the following continuous scalar functions of $\bar{p}$:

$$f_{XV}(\bar{p}) = p_{XV}(x, v), f_Y(\bar{p}) = H(Y), f_V(\bar{p}) = H(V)$$

Since there are $\|\mathcal{X}\| \|\mathcal{V}\| - 1$ functions of $f_{XV}(\bar{p})$, the total number of the continuous scalar functions of $\bar{p}$ is $\|\mathcal{X}\| \|\mathcal{V}\| + 1$.

Let $\bar{p}_{XV|U} = Pr\{X = x, V = v | U = u\}$. With these distributions $\bar{p}_{XV|U}$, we have

$$p_{XV|K}(x, v | k) = \sum_{u \in \mathcal{U}} p(U = u | K = k) f_{XV}(\bar{p}_{XV|U}) \quad (5)$$

$$I(U; Y) = f_Y(\bar{p}) - \sum_{u \in \mathcal{U}} p(U = u | K = k) f_Y(\bar{p}_{XV|U}) \quad (6)$$

$$I(U; V) = f_V(\bar{p}) - \sum_{u \in \mathcal{U}} p(U = u | K = k) f_V(\bar{p}_{XV|U}) \quad (7)$$

According to the support lemma ([11], p. 310), for every fixed k , the random variable U can be replaced by new ones such that the new U takes at most $\|\mathcal{X}\| \|\mathcal{V}\| + 1$ different values and the expressions in Equations (5–7) are preserved. Therefore, $\|\mathcal{U}\| \leq \|\mathcal{X}\|^2 \|\mathcal{V}\|^2 (\|\mathcal{X}\| \|\mathcal{V}\| + 1)$ is proved.

Proof of Lemma 1

Proof of (i)

Since $I(U; Y) - I(U; Z)$ and $I(U; Y)$ are continuous functions of $Pr\{X = x, U = u | V = v\}$, using similar argument of [1], p. 1382, we conclude that $I(U; Y) - I(U; Z)$ has a maximum on $\rho(R)$.

Proof of (ii)

Let $0 \leq R_1, R_2 \leq C_M$ and $0 \leq \theta \leq 1$, we will show that $\Gamma'(\theta R_1 + (1 - \theta)R_2) \geq \theta \Gamma'(R_1) + (1 - \theta) \Gamma'(R_2)$, see the Formulas (3) and (8).

Let $Pr\{U' = u, X' = x | V = v\} \in \rho(R_1)$ achieve $\Gamma'(R_1)$, i.e., $I(U'; Y') \geq R_1$ and $I(U'; Y') - I(U'; Z') = \Gamma'(R_1)$. Also let $Pr\{U'' = u, X'' = x | V = v\} \in \rho(R_2)$ achieve $\Gamma'(R_2)$, i.e., $I(U''; Y'') \geq R_2$ and $I(U''; Y'') - I(U''; Z'') = \Gamma'(R_2)$. Note that $U' \rightarrow Y' \rightarrow Z'$ and $U'' \rightarrow Y'' \rightarrow Z''$ are two Markov chains. The coefficient θ is determined by a random variable Q in Figure 3, such that $Pr\{Q = 1\} = \theta$ and $Pr\{Q = 2\} = 1 - \theta$. Q is independent of V, U', U'', Y', Y'', Z' and Z'' . By using $U' \rightarrow Y' \rightarrow Z', U'' \rightarrow Y'' \rightarrow Z''$ and the fact that Q is independent of $U', U'', Y', Y'', Z', Z''$, we conclude that $Q \rightarrow U \rightarrow Y \rightarrow Z$. Then we have,

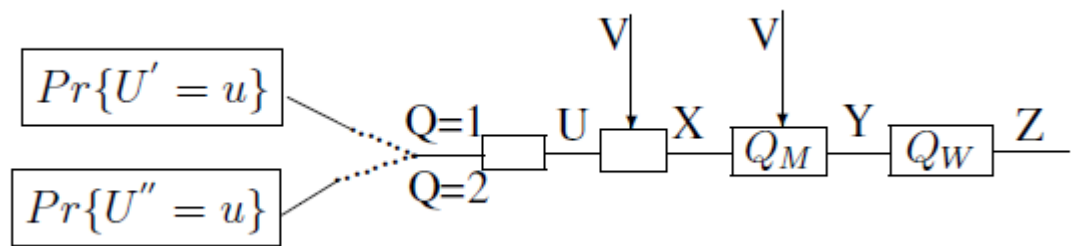
$$\begin{aligned} I(U; Y) &= H(Y) - H(Y|U) \\ &= H(Y) - H(Y|U, Q) \\ &\geq H(Y|Q) - H(Y|U, Q) \\ &= I(U; Y|Q) \end{aligned} \quad (1)$$

$$\begin{aligned}
&= \Pr\{Q = 1\}I(U; Y|Q = 1) + \Pr\{Q = 2\}I(U; Y|Q = 2) \\
&= \theta I(U'; Y') + (1 - \theta)I(U''; Y'')
\end{aligned} \tag{2}$$

$$\geq \theta R_1 + (1 - \theta)R_2 \tag{3}$$

Formula (1) follows from $Q \rightarrow U \rightarrow Y$. Formula (2) follows from that Q is independent of U', U'', Y', Y'' .

Figure 3. The definition of the random variable Q .



From the definition of $\Gamma'(R)$ and the Formula (3),

$$\begin{aligned}
\Gamma'(\theta R_1 + (1 - \theta)R_2) &\geq I(U; Y) - I(U; Z) \\
&= H(U|Z) - H(U|Y) \\
&= H(U|Z) - H(U|Y, Z) \\
&= I(U; Y|Z)
\end{aligned} \tag{4}$$

$$\begin{aligned}
&= H(Y|Z) - H(Y|U, Z) \\
&= H(Y|Z) - H(Y|U, Z, Q) \\
&\geq H(Y|Z, Q) - H(Y|U, Z, Q) \\
&= I(U; Y|Z, Q)
\end{aligned} \tag{5}$$

$$\begin{aligned}
&= \Pr\{Q = 1\}I(U; Y|Z, Q = 1) + \Pr\{Q = 2\}I(U; Y|Z, Q = 2) \\
&= \theta I(U'; Y'|Z') + (1 - \theta)I(U''; Y''|Z'')
\end{aligned} \tag{6}$$

$$= \theta(I(U'; Y') - I(U'; Z')) + (1 - \theta)(I(U''; Y'') - I(U''; Z'')) \tag{7}$$

$$= \theta \Gamma'(R_1) + (1 - \theta) \Gamma'(R_2) \tag{8}$$

Formula (4) follows from $U \rightarrow Y \rightarrow Z$. Formula (5) follows from $H(Y|U, Z) = H(Y|U, Z, Q)$. Formula (6) follows from the fact that Q is independent of $U', U'', Y', Y'', Z', Z''$. Formula (7) follows from $U' \rightarrow Y' \rightarrow Z'$ and $U'' \rightarrow Y'' \rightarrow Z''$.

Proof 9 (Proof of Equation (5))

$$\begin{aligned}
H(Y|U, Z) - H(Y|U, Z, Q) &= I(Y, Q|U, Z) \\
&= H(Q|U, Z) - H(Q|U, Z, Y) \\
&= {}^{(a)}H(Q|U) - H(Q|U) \\
&= 0
\end{aligned}$$

where (a) follows from $Q \rightarrow U \rightarrow Y \rightarrow Z$, completing the proof.

The proof of (ii) is completed.

Proof of (iii)

This part is from the definition of $\Gamma'(R)$, since $\rho(R)$ is a non-increasing set.

Proof of (iv)

The proof of (iv) is similar to [1], p. 1383.

The proof of Lemma 1 is completed.

Proof of Lemma 3

Lemma 4 Note that in the model of Figure 2, $(S^k, V^N) \rightarrow (X^N, V^N) \rightarrow Y^N \rightarrow Z^N$ is assumed to be a Markov chain, then we have $S^k \rightarrow (X^N, V^N) \rightarrow Y^N \rightarrow Z^N$.

Proof 10 $S^k \rightarrow (X^N, V^N) \rightarrow Y^N \rightarrow Z^N$ is a Markov chain if and only if $S^k \rightarrow (X^N, V^N) \rightarrow Y^N$ and $(S^k, X^N, V^N) \rightarrow Y^N \rightarrow Z^N$ are Markov chains, see [12], p. 10.

Since $(S^k, V^N) \rightarrow (X^N, V^N) \rightarrow Y^N \rightarrow Z^N$ is a Markov chain, then we have

$$H(Y^N | X^N, V^N, S^k, V^N) = H(Y^N | X^N, V^N, S^k) = H(Y^N | X^N, V^N),$$

$$H(Z^N | X^N, V^N, S^k, V^N, Y^N) = H(Z^N | X^N, V^N, S^k, Y^N) = H(Z^N | Y^N)$$

which imply that $S^k \rightarrow (X^N, V^N) \rightarrow Y^N$ and $(S^k, X^N, V^N) \rightarrow Y^N \rightarrow Z^N$. Lemma 4 is proved.

The proof of the Markov chain in Lemma 3 in the form of probability mass functions depends on the joint probability distribution $P(Z^N = z^N, Y^N = y^N, X^N = x^N, V^N = v^N, S^k = s^k)$, see the following.

$$\begin{aligned} & P(Z^N = z^N, Y^N = y^N, X^N = x^N, V^N = v^N, S^k = s^k) \\ = & P(Z^N = z^N | Y^N = y^N) P(Y^N = y^N | X^N = x^N, V^N = v^N) \cdot \\ & P(X^N = x^N | V^N = v^N, S^k = s^k) P(S^k = s^k) P(V^N = v^N) \end{aligned} \quad (9)$$

$$\begin{aligned} = & P(S^k = s^k) \prod_{n=1}^N P(Z_n = z_n | Y_n = y_n) P(Y_n = y_n | X_n = x_n, V_n = v_n) \cdot \\ & P(X_n = x_n | V_n = v_n, S^k = s^k) P(V_n = v_n) \end{aligned} \quad (10)$$

Formula (9) is from $S^k \rightarrow (X^N, V^N) \rightarrow Y^N \rightarrow Z^N$ (see Lemma 3) and the fact that S^k is independent of V^N . Formula (10) is from Equation (2.8), the properties of the discrete memoryless channel and the fact that V^N is composed of N i.i.d. random variables.

Proof of Lemma 3

The proof is considered in two parts. The first part is to calculate $P(Y^{i-1} = y^{i-1}, Z_i = z_i, S^k = s^k)$, which is obtained from the joint probability distribution $P(Z^N = z^N, Y^N = y^N, X^N = x^N, V^N = v^N, S^k = s^k)$. The second part is to prove $Y^{i-1} \rightarrow S^k \rightarrow Z_i$ by Equations (15) and (17).

First part. $P(Z_i = z_i, S^k = s^k, Y^{i-1} = y^{i-1})$ is calculated as follows, where z_i, y^{i-1}, s^k are fixed. Note that $z^{N*} = (z_1, z_2, \dots, z_{i-1}, z_{i+1}, \dots, z_N)$.

$$\begin{aligned}
 & P(Z_i = z_i, S^k = s^k, Y^{i-1} = y^{i-1}) \\
 = & \sum_{\substack{z^{N*}: N^*=I_N \setminus \{i\} \\ y_i^N, v^N, x^N}} P(Z^N = z^N, Y^N = y^N, X^N = x^N, V^N = v^N, S^k = s^k) \\
 = & \sum_{\substack{z^{N*}: N^*=I_N \setminus \{i\} \\ y_i^N, v^N, x^N}} P(S^k = s^k) \prod_{n=1}^N \{P(Z_n = z_n | Y_n = y_n) P(Y_n = y_n | X_n = x_n, V_n = v_n) \cdot \\
 & P(X_n = x_n | V_n = v_n, S^k = s^k) P(V_n = v_n)\} \quad (11)
 \end{aligned}$$

$$\begin{aligned}
 = & \sum_{\substack{z^{N*}: N^*=I_N \setminus \{i\} \\ y_i^N, v^N, x^N}} P(S^k = s^k) \prod_{n=1}^N \{P(Z_n = z_n | Y_n = y_n) P(Y_n = y_n | X_n = x_n, V_n = v_n, S^k = s^k) \cdot \\
 & P(X_n = x_n | V_n = v_n, S^k = s^k) P(V_n = v_n)\} \quad (12)
 \end{aligned}$$

$$\begin{aligned}
 = & \sum_{\substack{z^{N*}: N^*=I_N \setminus \{i\} \\ y_i^N, v^N, x^N}} P(S^k = s^k) \prod_{n=1}^N \{P(Z_n = z_n | Y_n = y_n) P(Y_n = y_n, X_n = x_n | V_n = v_n, S^k = s^k) \cdot \\
 & P(V_n = v_n)\}
 \end{aligned}$$

$$\begin{aligned}
 = & \sum_{\substack{z^{N*}: N^*=I_N \setminus \{i\} \\ y_i^N, v^N, x^N}} P(S^k = s^k) \prod_{n=1}^N P(Z_n = z_n | Y_n = y_n) P(Y_n = y_n | V_n = v_n, S^k = s^k) P(V_n = v_n)
 \end{aligned}$$

$$\begin{aligned}
 = & P(S^k = s^k) \sum_{y_i, v_i} P(Z_i = z_i | Y_i = y_i) P(Y_i = y_i | V_i = v_i, S^k = s^k) P(V_i = v_i) \cdot
 \end{aligned}$$

$$\sum_{v^{i-1}, z^{i-1}} \prod_{n=1}^{i-1} \{P(Z_n = z_n | Y_n = y_n) P(Y_n = y_n | V_n = v_n, S^k = s^k) P(V_n = v_n)\} \cdot$$

$$\sum_{v_{i+1}^N, y_{i+1}^N, z_{i+1}^N} \prod_{n=i+1}^N \{P(Z_n = z_n | Y_n = y_n) P(Y_n = y_n | V_n = v_n, S^k = s^k) P(V_n = v_n)\}$$

$$\begin{aligned}
 = & P(S^k = s^k) \sum_{y_i, v_i} P(Z_i = z_i | Y_i = y_i) P(Y_i = y_i | V_i = v_i, S^k = s^k) P(V_i = v_i) \cdot
 \end{aligned}$$

$$\sum_{v^{i-1}} \prod_{n=1}^{i-1} \{P(Y_n = y_n | V_n = v_n, S^k = s^k) P(V_n = v_n)\}$$

$$\begin{aligned}
 = & P(S^k = s^k) \sum_{y_i, v_i} P(Z_i = z_i | Y_i = y_i, V_i = v_i, S^k = s^k) P(Y_i = y_i | V_i = v_i, S^k = s^k) P(V_i = v_i) \cdot
 \end{aligned}$$

$$\sum_{v^{i-1}} \prod_{n=1}^{i-1} \{P(Y_n = y_n | V_n = v_n, S^k = s^k) P(V_n = v_n)\} \quad (13)$$

$$= P(S^k = s^k) \sum_{y_i, v_i} P(Z_i = z_i, Y_i = y_i | V_i = v_i, S^k = s^k) P(V_i = v_i) \cdot$$

$$\sum_{v^{i-1}} \prod_{n=1}^{i-1} \{P(Y_n = y_n | V_n = v_n, S^k = s^k) P(V_n = v_n)\} \\ = \sum_{v_i} P(Z_i = z_i | V_i = v_i, S^k = s^k) P(V_i = v_i) P(S^k = s^k) \cdot$$

$$\sum_{v^{i-1}} \prod_{n=1}^{i-1} \{P(Y_n = y_n | V_n = v_n, S^k = s^k) P(V_n = v_n)\} \\ = P(Z_i = z_i, S^k = s^k) \sum_{v^{i-1}} \prod_{n=1}^{i-1} \{P(Y_n = y_n | V_n = v_n, S^k = s^k) P(V_n = v_n)\} \quad (14)$$

$$= P(Z_i = z_i | S^k = s^k) \sum_{v^{i-1}} \prod_{n=1}^{i-1} \{P(Y_n = y_n | V_n = v_n, S^k = s^k) P(V_n = v_n) P(S^k = s^k)\}$$

$$= P(Z_i = z_i | S^k = s^k) \prod_{n=1}^{i-1} P(Y_n = y_n, S^k = s^k) \quad (15)$$

Formula (11) is from Equation (10). Formula (12) is from the Markov chain $S^k \rightarrow (X_n, V_n) \rightarrow Y_n$. Formula (13) is from the Markov chain $(S^k, V_i) \rightarrow Y_i \rightarrow Z_i$. Formulas (14) and (15) are from the fact that S^k is independent of V^N .

Second part. By the definition of Markov chain in the form of probability mass function and Equation (15),

$$P(Z_i = z_i, S^k = s^k, Y^{i-1} = y^{i-1}) P(S^k = s^k) \\ = P(S^k = s^k) P(Z_i = z_i | S^k = s^k) \prod_{n=1}^{i-1} P(Y_n = y_n, S^k = s^k) \quad (16)$$

$$= P(Z_i = z_i, S^k = s^k) \sum_{z_i} P(Z_i = z_i, S^k = s^k, Y^{i-1} = y^{i-1}) \quad (17) \\ = P(Z_i = z_i, S^k = s^k) P(S^k = s^k, Y^{i-1} = y^{i-1})$$

where Equations (16) and (17) are from Equation (15). Thus, the proof of $Y^{i-1} \rightarrow S^k \rightarrow Z_i$ is completed.

The proof of Lemma 3 is completed.

References

1. Wyner, A.D. The wire-tap channel. *Bell Syst. Tech. J.* **1975**, *54*, 1355–1387.
2. Csiszár, I.; Körner, J. Broadcast channels with confidential messages. *IEEE Trans. Inf. Theory* **1978**, *24*, 339–348.
3. Leung-Yan-Cheong, S.K.; Hellman, M.E. The Gaussian wire-tap channel. *IEEE Trans. Inf. Theory* **1978**, *24*, 451–456.

4. Shannon, C.E. Channels with side information at the transmitter. *IBM J. Res. Dev.* **1958**, *2*, 289–293.
5. Kuznetsov, N.V.; Tsybakov, B.S. Coding in memories with defective cells. *Probl. Peredachi Informatsii* **1974**, *10*, 52–60.
6. Gel'fand, S.I.; Pinsker, M.S. Coding for channel with random parameters. *Problems. Control Inf. Theory* **1980**, *9*, 19–31.
7. Costa, M.H.M. Writing on dirty paper. *IEEE Trans. Inf. Theory* **1983**, *29*, 439–441.
8. Mitrpant, C.; Han Vinck, A.J.; Luo, Y. An achievable region for the gaussian wiretap channel with side information. *IEEE Trans. Inf. Theory* **2006**, *52*, 2181–2190.
9. Chen, Y.; Han Vinck, A.J. Wiretap channel with side information. *IEEE Trans. Inf. Theory* **2008**, *54*, 395–402.
10. Merhav, N. Shannon's secrecy system with informed receivers and its application to systematic coding for wiretapped channels. *IEEE Trans. Inf. Theory, special issue on Inf.-Secur.* **2008**, *54*, 2723–2734.
11. Csiszár, I.; Körner, J. *Information Theory: Coding Theorems for Discrete Memoryless Systems*; Academic: London, UK, 1981; pp. 123–124.
12. Yeung, R.W. *Information Theory and Network Coding*; Springer: New York, NY, USA, 2008; pp. 325–326.

© 2012 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).