entropy

MDPI

Article

Feeding Back the Output or Sharing the State: Which Is Better for the State-Dependent Wiretap Channel?

Bin Dai [1,2,*], Zheng Ma [1] and Linman Yu [3]

Received: 29 July 2015 / Accepted: 23 November 2015 / Published: 30 November 2015
Academic Edtior: Raúl Alcaraz Martínez

1 School of Information Science and Technology, Southwest JiaoTong University, Northbound Section Second Ring Road 111, Chengdu 610031, China; zma@home.swjtu.edu.cn
2 The National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China
3 School of Economics and Management, Chengdu Textile College, Chengdu 611731, China; yulinmanylm@163.com
* Correspondence: daibin@home.swjtu.edu.cn; Tel./Fax: +86-28-8763-4758

Abstract: In this paper, the general wiretap channel with channel state information (CSI) at the transmitter and noiseless feedback is investigated, where the feedback is from the legitimate receiver to the transmitter, and the CSI is available at the transmitter in a causal or noncausal manner. The capacity-equivocation regions are determined for this model in both causal and noncausal cases, and the results are further explained via Gaussian and binary examples. For the Gaussian model, we find that in some particular cases, the noiseless feedback performs better than Chia and El Gamal's CSI sharing scheme, *i.e.*, the secrecy capacity of this feedback scheme is larger than that of the CSI sharing scheme. For the degraded binary model, we find that the noiseless feedback performs no better than Chia and El Gamal's CSI sharing scheme. However, if the cross-over probability of the wiretap channel is large enough, we show that the two schemes perform the same.

Keywords: capacity-equivocation region; channel state information; noiseless feedback; secrecy capacity; wiretap channel

1. Introduction

It is known to all that the capacity of a point-to-point discrete memoryless channel (DMC) cannot be increased by using noiseless feedback. However, does the feedback (from the legitimate receiver to the transmitter) enhance the security of the wiretap channel? Ahlswede and Cai [1] and Dai *et al.* [2] studied this problem. Specifically, Ahlswede and Cai [1] showed that the secrecy capacity C_{sf} of the degraded wiretap channel with noiseless feedback is given by:

$$C_{sf} = \max_{p(x)} \min\{I(X;Y), I(X;Y) - I(X;Z) + H(Y|X,Z)\}, \tag{1}$$

where X, Y and Z are for the transmitter, legitimate receiver and wiretapper, respectively, and $X \to Y \to Z$ forms a Markov chain. Recall that the secrecy capacity C_s of the degraded wiretap channel is determined by Wyner [3], and it is given by:

$$C_s = \max_{p(x)} \min\{I(X;Y), I(X;Y) - I(X;Z)\}. \tag{2}$$

From (1) and (2), it is easy to see that the noiseless feedback increases the secrecy capacity of the wiretap channel. Based on the work of [1], Dai *et al.* [2] studied a special wiretap channel with feedback ($Y \to X \to Z$) and showed that the secrecy capacity of this model is larger than that of

the model without feedback, *i.e.*, the noiseless feedback helps to enhance the security of the special wiretap channel $Y \to X \to Z$. Here, note that in [1] and [2], the legitimate receiver just sends back the previous received symbols to the transmitter, and it is natural to ask: is it better for the legitimate receiver to send back purely random secret keys to the transmitter? Ardestanizadeh *et al.* [4] answered this question by considering the model of the wiretap channel with secure rate-limited feedback. Ardestanizadeh *et al.* [4] showed that if the limits (capacity) of the feedback channel are denoted by R_f, the secrecy capacity of the physically-degraded wiretap channel ($X \to Y \to Z$) with secure rate-limited feedback is given by:

$$C_{sf} = \max_{p(x)} \min\{I(X;Y), I(X;Y) - I(X;Z) + R_f\}. \tag{3}$$

Compared to (1), it is easy to see that if $R_f \leq H(Y|X,Z)$, sending purely random secret keys is no better than sending Y^{i-1} back. If $R_f > H(Y|X,Z)$, $I(X;Y) - I(X;Z) + R_f > H(Y|Z)$, sending purely random secret keys is better than sending Y^{i-1} back. Besides these works on the wiretap channel with feedback, Lai *et al.* [5] studied the wiretap channel with noisy feedback; He *et al.* [6] studied the Gaussian two-way wiretap channel and the Gaussian half-duplex two-way relay channel with an un-trusted relay; and Bassi *et al.* [7] studied the wiretap channel with generalized feedback. Bounds on the secrecy capacities of these feedback models are obtained in [5–7].

Recently, the wiretap channel with channel state information (CSI) has received much attention. The Gaussian wiretap channel with noncausal CSI at the transmitter was studied in [8,9], and an achievable rate-equivocation region was provided for this Gaussian model. Based on the work of [8], Chen *et al.* [10] studied the discrete memoryless wiretap channel with noncausal CSI at the transmitter and also provided an achievable rate-equivocation region for this model. The encoding-decoding scheme of [10] is a combination of the binning technique of Gel'fand and Pinsker's channel [11] and the random binning technique of Wyner's wiretap channel [3]. After that, Dai *et al.* [12] studied the outer bound on the capacity-equivocation region of [10] and also investigated the capacity results of the discrete memoryless wiretap channel with causal or memoryless CSI at the transmitter. Besides these works on the wiretap channel with CSI only available at the transmitter, Chia and El Gamal [13] investigated the wiretap channel with CSI causally or non-causally at both the transmitter and the legitimate receiver and provided an achievable secrecy rate, which was larger than that of [10]. In [13], since both the transmitter and the legitimate receiver have access to the CSI, the CSI serves as a secret key shared by them. Therefore, the encoding-decoding scheme of [13] is similar to that of the wiretap channel with rate-limited feedback [4]. Besides these works on the wiretap channel with CSI, Liu *et al.* [14] studied the block Rayleigh fading MIMO wiretap channel with no CSI available at the legitimate receiver, the wiretapper and the transmitter, and they showed that if the legitimate receiver had more antennas than the wiretapper, non-zero secure degrees of freedom (s.d.o.f) could also be achieved.

In this paper, we study the general wiretap channel with CSI (causally or non-causally at the transmitter) and noiseless feedback; see Figure 1. In Figure 1, the transition probability of the channel depends on a CSI sequence V^N, which is available at the channel encoder in a noncausal or causal manner. The inputs of the channel are X^N and V^N, while the outputs of the channel are Y^N and Z^N. Moreover, there exists a noiseless feedback from Y^N to the channel encoder. The motivation of this work is to find whether the noiseless feedback helps to enhance the secrecy rate of the wiretap channel with noncausal or causal CSI at the transmitter [10,12] and whether the noiseless feedback does better than the shared CSI between the legitimate receiver and the transmitter [13] in enhancing the secrecy rate of the state-dependent wiretap channel.

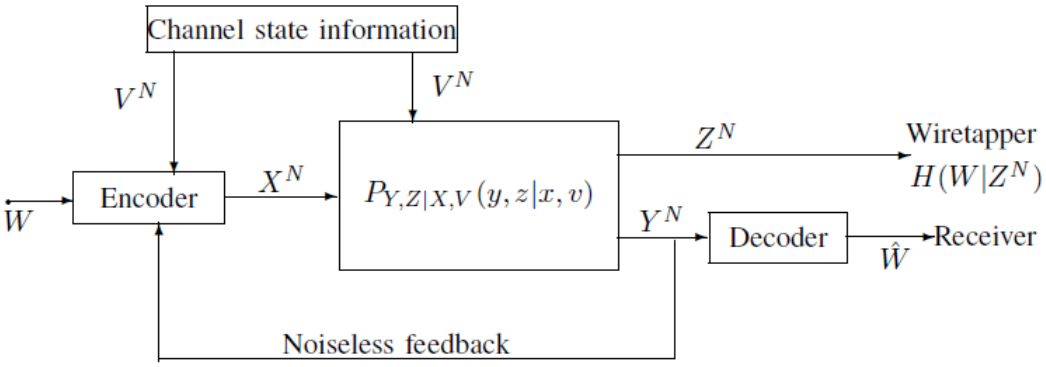

Figure 1. General wiretap channel with noncausal or causal channel state information (CSI) and noiseless feedback.

The capacity-equivocation region of the model of Figure 1 is determined for both the noncausal and causal cases, and the results are further explained via degraded binary and Gaussian examples. For the Gaussian example, we find that both the feedback scheme and the CSI sharing scheme [13] help to enhance the security of the wiretap channel with noncausal CSI at the transmitter [10,12], and moreover, we find that in some particular cases, the noiseless feedback performs even better than the shared CSI [13], *i.e.*, the secrecy capacity of the degraded Gaussian case of the model of Figure 1 is larger than that of the degraded Gaussian case of [13]. For the binary example, we also find that both the feedback scheme and the CSI sharing scheme [13] help to enhance the security of the wiretap channel with causal CSI at the transmitter. Unlike the Gaussian case, we find that the noiseless feedback performs no better than the shared CSI [13], *i.e.*, the secrecy capacity of the degraded binary case of the model of Figure 1 is not more than that of the degraded binary case of [13]. However, if the cross-over probability of the wiretap channel is large enough, we find that the two schemes perform the same.

The remainder of this paper is organized as follows. The capacity-equivocation region of the model of Figure 1 is provided in Section 2. Gaussian and binary examples of the model of Figure 1 are shown in Section 3. Section 4 is for the final conclusion.

2. Capacity-Equivocation Region of the Model of Figure 1

In this 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 $(U_1, ..., U_N)$, and $u^N = (u_1, ..., u_N)$ 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 $(U_i, ..., U_N)$, and $u_i^N = (u_i, ..., u_N)$ 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 to the base two.

2.1. Definitions of the Model of Figure 1

Let W, uniformly distributed over the alphabet $\mathcal{W}$, be the message sent by the transmitter. The components of the channel state sequence V^N are independent and identically distributed. The probability of each component is $P_V(v)$. V^N is independent of W. Let Y^{i-1} $(2 \leq i \leq N)$ be the i-th time feedback from the legitimate receiver to the transmitter. For the noncausal case, the i-th time channel encoder f_i is a (stochastic) mapping:

$$f_i : \mathcal{W} \times \mathcal{Y}^{i-1} \times \mathcal{V}^N \to \mathcal{X}_i, \tag{4}$$

where $f_i(w, y^{i-1}, v^N) = x_i \in \mathcal{X}$, $w \in \mathcal{W}$, $y^{i-1} \in \mathcal{Y}^{i-1}$ and $v^N \in \mathcal{V}^N$. For the causal case, the i-th time channel encoder f_i is a (stochastic) mapping:

$$f_i : \mathcal{W} \times \mathcal{Y}^{i-1} \times \mathcal{V}^i \to \mathcal{X}_i, \tag{5}$$

where $f_i(w, y^{i-1}, v^i) = x_i \in \mathcal{X}$, $w \in \mathcal{W}$, $y^{i-1} \in \mathcal{Y}^{i-1}$ and $v^i \in \mathcal{V}^i$. Here, note that for the causal case, V_i is independent of $(Y^{i-1}, W, V_{i+1}^N, Z^{i-1})$.

The channel is discrete memoryless, and its transition probability is given by:

$$P_{Z^N,Y^N|X^N,V^N}(z^N, y^N|x^N, v^N) = \prod_{i=1}^{N} P_{Z,Y|X,V}(z_i, y_i|x_i, v_i), \tag{6}$$

where $x_i \in \mathcal{X}$, $v_i \in \mathcal{V}$, $y_i \in \mathcal{Y}$ and $z_i \in \mathcal{Z}$.

The wiretapper's equivocation about the message W is denoted by:

$$\Delta = \frac{1}{N} H(W|Z^N). \tag{7}$$

The decoder f_D is a function that maps a received sequence of N channel outputs to the messages set:

$$f_D : \mathcal{Y}^N \to \mathcal{W}. \tag{8}$$

We denote the probability of error P_e by $Pr\{W \neq \hat{W}\}$.

Given a pair (R, R_e) $(R, R_e > 0)$, it is said to be achievable if, for arbitrary small positive ϵ, there exists an encoding-decoding scheme, such that:

$$\lim_{N\to\infty} \frac{\log \parallel \mathcal{W} \parallel}{N} = R, \lim_{N\to\infty} \Delta \geq R_e, P_e \leq \epsilon. \tag{9}$$

The set $\mathcal{R}^{(nf)}$, which is composed of all achievable (R, R_e) pairs, is called the capacity-equivocation region of the model of Figure 1 with noncausal CSI at the transmitter. An achievable rare $C_s^{(nf)}$, which is denoted by:

$$C_s^{(nf)} = \max_{(R,R_e=R)\in\mathcal{R}^{(nf)}} R, \tag{10}$$

is called the secrecy capacity of the model of Figure 1 with noncausal CSI at the transmitter.

Analogously, let $\mathcal{R}^{(cf)}$ be the capacity-equivocation region of the model of Figure 1 with causal CSI at the transmitter and $C_s^{(cf)}$, which is denoted by:

$$C_s^{(cf)} = \max_{(R,R_e=R)\in\mathcal{R}^{(cf)}} R, \tag{11}$$

be the secrecy capacity of the model of Figure 1 with causal CSI at the transmitter.

2.2. Main Result of the Model of Figure 1

The following Theorem 1 characterizes the capacity-equivocation region $\mathcal{R}^{(nf)}$ of the model of Figure 1 with noncausal CSI at the transmitter; see the following.

Theorem 1. *A single-letter characterization of the region $\mathcal{R}^{(nf)}$ is as follows,*

$$\begin{aligned}\mathcal{R}^{(nf)} = \{&(R, R_e) : 0 \leq R_e \leq R,\\ &0 \leq R \leq I(K;Y) - I(K;V),\\ &R_e \leq H(Y|Z)\},\end{aligned}$$

for some distribution:

$$P_{KVXYZ}(k,v,x,y,z) = P_{ZY|XV}(z,y|x,v)P_{X|KV}(x|k,v)P_{KV}(k,v),$$

which implies the Markov chain $K \to (X,V) \to (Y,Z)$.

Proof. See Sections A and B. □

Remark 1.

- The range of the random variable K satisfies $\|\mathcal{K}\| \leq \|\mathcal{X}\|\|\mathcal{V}\| + 1$. The proof is standard and easily obtained by using the support lemma (see [15]), and thus, we omit the proof here.
- **Corollary 1.** The secrecy capacity $C_s^{(nf)}$ satisfies:

$$C_s^{(nf)} = \max_{P_{X|KV}P_{KV}} \min\{I(K;Y) - I(K;V), H(Y|Z)\}. \tag{12}$$

Proof. Substituting $R_e = R$ into the region $\mathcal{R}^{(nf)}$ in Theorem 1, we have:

$$\begin{aligned} R &\leq I(K;Y) - I(K;V), &(13)\\ R &\leq H(Y|Z), &(14)\end{aligned}$$

By using (10), (13) and (14), Formula (12) is achieved; thus, the proof is completed. □

- Here, note that if Z^N is a degraded version of Y^N (which implies the existence of the Markov chain $K \to (X,V) \to Y \to Z$), the capacity-equivocation region $\mathcal{R}^{(nf)}$ still holds. The proof of this degraded case is along the lines of the proof of Theorem 1, and thus, we omit the proof here. In [10,12], an achievable rate-equivocation region $\mathcal{R}_i^n$ is provided for the wiretap channel with noncausal CSI, and it is given by:

$$\begin{aligned} \mathcal{R}_i^n = \{&(R,R_e) : R_e \leq R,\\ &R \leq I(K;Y) - I(K;V),\ \ R_e \leq I(K;Y) - I(K;Z)\}, \end{aligned}$$

where the joint probability distribution $P_{KVXYZ}(k,v,x,y,z)$ of $\mathcal{R}_i^n$ satisfies:

$$P_{KVXYZ}(k,v,x,y,z) = P_{Z|Y}(z|y)P_{Y|XV}(y|x,v)P_{X|KV}(x|k,v)P_{KV}(k,v).$$

Here, note that:

$$\begin{aligned} I(K;Y) - I(K;Z) &= H(K|Z) - H(K|Y)\\ &\overset{(a)}{=} H(K|Z) - H(K|Y,Z) = I(K;Y|Z)\\ &\leq H(Y|Z), \end{aligned} \tag{15}$$

where (a) is from $K \to Y \to Z$. Therefore, it is easy to see that the achievable rate-equivocation region $\mathcal{R}_i^n$ of [10] and [12] is enhanced by using this noiseless feedback.

The following Theorem 2 characterizes the capacity-equivocation region $\mathcal{R}^{(cf)}$ of the model of Figure 1 with causal CSI at the transmitter; see the following.

Theorem 2. *A single-letter characterization of the region* $\mathcal{R}^{(cf)}$ *is as follows,*

$$\begin{aligned} \mathcal{R}^{(cf)} = \{&(R,R_e) : 0 \leq R_e \leq R,\\ &0 \leq R \leq I(K;Y),\\ &R_e \leq H(Y|Z)\}, \end{aligned}$$

for some distribution:

$$P_{KVXYZ}(k,v,x,y,z) = P_{YZ|XV}(y,z|x,v)P_{X|KV}(x|k,v)P_K(k)P_V(v).$$

which implies the Markov chain $K \to (X,V) \to (Y,Z)$ *and the fact that* V *is independent of* K.

Proof.
- Proof of the converse: Using the fact that V_i is independent of Y^{i-1} and Z^{i-1}, the converse proof of Theorem 2 is along the lines of that of Theorem 1 (see Section A), and thus, we omit the proof here.
- Proof of the achievability: The achievability proof of Theorem 2 is along the lines of the achievability proof of Theorem 1 (see Section B), and the only difference is that for the causal case, there is no need to use the binning technique. Thus, we omit the proof here.

The proof of Theorem 2 is completed. □

Remark 2.

- The range of the auxiliary random variable K satisfies $\|\mathcal{K}\| \leq \|\mathcal{X}\|\|\mathcal{V}\|$. The proof is standard and easily obtained by using the support lemma (see p. 310 of [16]), and thus, we omit the proof here.
- **Corollary 2.** The secrecy capacity $C_s^{(cf)}$ satisfies:

$$C_s^{(cf)} = \max_{P_{X|KV}P_K} \min\{I(K;Y), H(Y|Z)\}. \tag{16}$$

Proof. Proof of (16): Substituting $R_e = R$ into the region $\mathcal{R}^{(cf)}$, we have:

$$R \leq I(K;Y), \tag{17}$$
$$R \leq H(Y|Z),. \tag{18}$$

By using (11), (17) and (18), Formula (16) is achieved; thus, the proof is completed. □

- Here, note that if Z^N is a degraded version of Y^N, the capacity-equivocation region $\mathcal{R}^{(cf)}$ still holds. The proof of this degraded case is along the lines of the proof of Theorem 2, and thus, we omit the proof here. In [12], an achievable rate-equivocation region $\mathcal{R}_i^c$ is provided for the wiretap channel with causal CSI, and it is given by:

$$\mathcal{R}_i^c = \{(R,R_e) : R_e \leq R, \\ R \leq I(K;Y),\ \ R_e \leq I(K;Y) - I(K;Z)\}, \tag{19}$$

where the joint probability distribution $P_{KVXYZ}(k,v,x,y,z)$ of $\mathcal{R}_i^c$ satisfies:

$$P_{KVXYZ}(k,v,x,y,z) = P_{Z|Y}(z|y)P_{Y|XV}(y|x,v)P_{X|KV}(x|k,v)P_K(k)P_V(v).$$

By using (15), it is easy to see that the achievable rate-equivocation region $\mathcal{R}_i^c$ is enhanced by using this noiseless feedback.

3. Examples of the Model of Figure 1

3.1. Gaussian Case of the Model of Figure 1 with Noncausal CSI at the Transmitter

For the Gaussian case of the model of Figure 1 with noncausal CSI at the transmitter, the i-th time ($1 \leq i \leq N$) channel inputs and outputs are given by:

$$Y_i = X_i + V_i + Z_{1,i},\ \ Z_i = X_i + V_i + Z_{2,i}, \tag{20}$$

where $V_i \sim \mathcal{N}(0, Q)$, $Z_{1,i} \sim \mathcal{N}(0, N_1)$ and $Z_{2,i} \sim \mathcal{N}(0, N_2)$. Here, note that V_i, $Z_{1,i}$ and $Z_{2,i}$ are independent random variables, X_i is independent of $Z_{1,i}$ and $Z_{2,i}$ and $\frac{1}{N}\sum_{i=1}^{N} E(X_i^2) \leq P$. The noise V_i is non-causally known by the transmitter. The following Theorem 3 shows the secrecy capacity of the Gaussian case of the model of Figure 1 with noncausal CSI at the transmitter.

Theorem 3. *For the Gaussian case of the model of Figure 1 with noncausal CSI at the transmitter, the secrecy capacity C_s^{gf} is characterized in the following two cases.*

Case 1: If $N_1 \leq N_2$, the secrecy capacity C_s^{gf} is given by:

$$\begin{aligned} C_s^{gf} &= \max_{\alpha} \min \left\{ \begin{array}{l} \frac{1}{2}\ln \frac{(P+Q+N_1)(P+\alpha^2 Q)}{PQ(1-\alpha)^2+N(P+\alpha^2 Q)} - \frac{1}{2}\ln \frac{P+\alpha^2 Q}{P}, \\ \frac{1}{2}\ln \frac{2\pi e(P+Q+N_1)(N_2-N_1)}{P+Q+N_2} \end{array} \right\} \\ &= \min\{\frac{1}{2}\ln(1+\frac{P}{N_1}), \frac{1}{2}\ln\frac{2\pi e(P+Q+N_1)(N_2-N_1)}{P+Q+N_2}\}, \end{aligned} \tag{21}$$

where the maximum is achieved when $\alpha = \frac{P}{P+N_1}$.

Case 2: If $N_1 > N_2$, the secrecy capacity C_s^{gf} is given by:

$$C_s^{gf} = \min\{\frac{1}{2}\ln(1+\frac{P}{N_1}), \frac{1}{2}\ln 2\pi e(N_1-N_2)\}. \tag{22}$$

Remark 3.

If $N_1 \leq N_2$, the relationship of the channel inputs and outputs defined in (20) can be equivalently characterized by:

$$Y_i = X_i + V_i + Z_{1,i}, \;\; Z_i = X_i + V_i + Z_{1,i} + Z_{2,i}^*, \tag{23}$$

where $Z_{2,i}^* \sim \mathcal{N}(0, N_2 - N_1)$, and it is independent of $Z_{1,i}$. Similar to the determination of the capacity region of the Gaussian broadcast channel (pp. 117–118 of [17]), the relationship (23) implies that there exists a Markov chain $(X_i, V_i) \to Y_i \to Z_i$, *i.e.*, the Gaussian case of the model of Figure 1 reduces to a degraded model of Figure 1.

Analogously, if $N_1 > N_2$, the relationship of the channel inputs and outputs defined in (20) can be equivalently characterized by:

$$Y_i = X_i + V_i + Z_{1,i}^* + Z_{2,i}, \;\; Z_i = X_i + V_i + Z_{2,i}, \tag{24}$$

where $Z_{1,i}^* \sim \mathcal{N}(0, N_1 - N_2)$, and it is independent of $Z_{2,i}$, X_i and V_i. The relationship (24) implies that there exists a Markov chain $(X_i, V_i) \to Z_i \to Y_i$ in the Gaussian case of the model of Figure 1.

Proof. For the direct part of Theorem 3, like [18] and [10], the achievability of C_s^{gf} is proven by substituting $K = X + \alpha V$, $X \sim \mathcal{N}(0, P)$, $V \sim \mathcal{N}(0, Q)$ and the fact that X is independent of V in Theorem 1; the details of the proof are omitted in this paper. Here, note that the calculation of $I(K;Y) - I(K;V)$ is exactly the same as that of the dirty paper channel (page 440 of [18]), and it is easy to see that the maximum of $I(K;Y) - I(K;V)$ is achieved when $\alpha = \frac{P}{P+N_1}$.

For the converse part of Theorem 3, note that the transmitter-receiver channel is Costa's dirty paper channel [18]; thus, the secrecy capacity is upper bounded by the capacity of the dirty paper channel, *i.e.*, $C_s^{gf} \leq \frac{1}{2}\ln(1+\frac{P}{N_1})$. Now, it remains to show $C_s^{gf} \leq \frac{1}{2}\ln\frac{2\pi e(P+Q+N_1)(N_2-N_1)}{P+Q+N_2}$ for $N_1 \leq N_2$ and $C_s^{gf} \leq \frac{1}{2}\ln 2\pi e(N_1-N_2)$ for $N_1 > N_2$; see the following.

Proof of $C_s^{gf} \leq \frac{1}{2}\ln\frac{2\pi e(P+Q+N_1)(N_2-N_1)}{P+Q+N_2}$ for $N_1 \leq N_2$:

First, note that:

$$\begin{aligned}\frac{1}{N}H(W|Z^N) &\overset{(a)}{\leq} \frac{1}{N}(I(W;Y^N|Z^N)+\delta(P_e))\\ &\leq \frac{1}{N}\sum_{i=1}^{N}h(Y_i|Z_i)+\frac{\delta(P_e)}{N},\end{aligned} \tag{25}$$

where (a) is from Fano's inequality. The conditional differential entropy $h(Y_i|Z_i)$ in (25) is bounded by:

$$\begin{aligned}h(Y_i|Z_i) &= h(Y_i,Z_i)-h(Z_i)\\ &= h(Z_i|Y_i)+h(Y_i)-h(Z_i)\\ &\overset{(1)}{=} h(X_i+V_i+Z_{1,i}+Z^*_{2,i}|X_i+V_i+Z_{1,i})+h(X_i+V_i+Z_{1,i})-h(X_i+V_i+Z_{1,i}+Z^*_{2,i})\\ &\overset{(2)}{=} h(Z^*_{2,i})+h(X_i+V_i+Z_{1,i})-h(X_i+V_i+Z_{1,i}+Z^*_{2,i})\\ &\overset{(3)}{\leq} h(Z^*_{2,i})+h(X_i+V_i+Z_{1,i})-\frac{1}{2}\ln(e^{2h(X_i+V_i+Z_{1,i})}+e^{2h(Z^*_{2,i})})\\ &= h(Z^*_{2,i})+\frac{1}{2}\ln(e^{2h(X_i+V_i+Z_{1,i})})-\frac{1}{2}\ln(e^{2h(X_i+V_i+Z_{1,i})}+e^{2h(Z^*_{2,i})})\\ &\overset{(4)}{=} \frac{1}{2}\ln(2\pi e(N_2-N_1))+\frac{1}{2}\ln\frac{e^{2h(X_i+V_i+Z_{1,i})}}{e^{2h(X_i+V_i+Z_{1,i})}+2\pi e(N_2-N_1)}\\ &\overset{(5)}{\leq} \frac{1}{2}\ln(2\pi e(N_2-N_1))+\frac{1}{2}\ln\frac{2\pi e(P+Q+N_1)}{2\pi e(P+Q+N_1)+2\pi e(N_2-N_1)}\\ &= \frac{1}{2}\ln\frac{2\pi e(N_2-N_1)(P+Q+N_1)}{P+Q+N_2},\end{aligned} \tag{26}$$

where (1) is from Definition (23), (2) is from the fact that $Z^*_{2,i}$ is independent of X_i, V_i and $Z_{1,i}$, (3) is from the entropy power inequality $e^{2h(X_i+V_i+Z_{1,i}+Z^*_{2,i})} \geq e^{2h(X_i+V_i+Z_{1,i})}+e^{2h(Z^*_{2,i})}$ (see [19]), (4) is from the fact that the differential entropy of a Gaussian distributed random variable X is $h(X)=\frac{1}{2}\ln(2\pi eD(X))$ (here, $D(X)$ is the variance of the Gaussian random variable X) and (5) is from $\frac{1}{2}\ln\frac{e^{2h(X_i+V_i+Z_{1,i})}}{e^{2h(X_i+V_i+Z_{1,i})}+2\pi e(N_2-N_1)}$ increasing while $h(X_i+V_i+Z_{1,i})$ is increasing and the fact that $h(X_i+V_i+Z_{1,i}) \leq \frac{1}{2}\ln(2\pi e(P+Q+N_1))$ (here, note that "=" is achieved if $X_i \sim \mathcal{N}(0,P)$).

Substituting (26) into (25), we have:

$$\begin{aligned}\frac{1}{N}H(W|Z^N) &\leq \frac{1}{N}\sum_{i=1}^{N}\frac{1}{2}\ln\frac{2\pi e(N_2-N_1)(P+Q+N_1)}{P+Q+N_2}+\frac{\delta(P_e)}{N}\\ &= \frac{1}{2}\ln\frac{2\pi e(N_2-N_1)(P+Q+N_1)}{P+Q+N_2}+\frac{\delta(P_e)}{N}.\end{aligned} \tag{27}$$

Substituting $P_e \leq \epsilon$ into (27) and letting $N \to \infty$, it is easy to see that $C_s^{gf} \leq \frac{1}{2}\ln\frac{2\pi e(P+Q+N_1)(N_2-N_1)}{P+Q+N_2}$ for $N_1 \leq N_2$.

Proof of $C_s^{gf} \leq \frac{1}{2}\ln 2\pi e(N_1-N_2)$ for $N_1 > N_2$:

For the case $N_1 > N_2$, the conditional differential entropy $h(Y_i|Z_i)$ in (25) can be bounded by:

$$\begin{aligned}h(Y_i|Z_i) &\overset{(a)}{=} h(X_i+V_i+Z^*_{1,i}+Z_{2,i}|X_i+V_i+Z_{2,i})\\ &\overset{(b)}{=} h(Z^*_{1,i})\\ &\overset{(c)}{=} \frac{1}{2}\ln 2\pi e(N_1-N_2),\end{aligned} \tag{28}$$

where (a) is from (24), (b) is from the fact that $Z_{1,i}^*$ is independent of $Z_{2,i}$, X_i and V_i and (c) is from the fact that the differential entropy of a Gaussian distributed random variable X is $h(X) = \frac{1}{2}\ln(2\pi e D(X))$ (here, $D(X)$ is the variance of the Gaussian random variable X). Substituting (28) and $P_e \leq \epsilon$ into (25) and letting $N \to \infty$, it is easy to see that $C_s^{gf} \leq \frac{1}{2}\ln 2\pi e(N_1 - N_2)$ for $N_1 > N_2$. Thus, the converse part of Theorem 3 is proven. The proof of Theorem 3 is completed. □

In [13] (p.2841, Theorem 3), Chia and El Gamal showed that if Y is less noisy than Z ($I(X;Y|V) \geq I(X;Z|V)$ for every $P_{X|V}(x|v)$), the secrecy capacity of the wiretap channel with CSI non-causally known by both the transmitter and the legitimate receiver is given by:

$$C_{s-both} = \max_{p(x|v)} \min\{I(X;Y|V), I(X;Y|V) - I(X;Z|V) + H(V|Z)\}.$$

Here, the $I(X;Z|V) - H(V|Z)$ in the above C_{s-both} can be rewritten as follows.

$$\begin{aligned} I(X;Z|V) - H(V|Z) &= H(Z|V) - H(Z|X,V) - H(V|Z) \\ &= H(V,Z) - H(V) - H(Z|X,V) - H(V,Z) + H(Z) \\ &= H(Z) - H(V) - H(Z|X,V). \end{aligned} \tag{29}$$

Substituting (29) into C_{s-both}, we have:

$$C_{s-both} = \max_{p(x|v)} \min\{I(X;Y|V), I(X;Y|V) - H(Z) + H(V) + H(Z|X,V)\}. \tag{30}$$

On the other hand, for Z less noisy than Y ($I(X;Z|V) \geq I(X;Y|V)$ for every $P_{X|V}(x|v)$), Chia and El Gamal provided an achievable secrecy rate (lower bound on the secrecy capacity) for the wiretap channel with CSI non-causally known by both the transmitter and the legitimate receiver, and it is given by:

$$C_{s-both}^{i} = \max_{p(x|v)} \min\{I(X;Y|V), H(V|Z,X)\}. \tag{31}$$

The following Theorem 4 shows the results on the secrecy capacity of the Gaussian case of the wiretap channel with CSI non-causally known by both the transmitter and the legitimate receiver.

Theorem 4. *For the Gaussian wiretap channel with part of the Gaussian noise non-causally known by both the transmitter and the legitimate receiver, the secrecy capacity C_{s-both}^{g} is characterized by the following two cases.*

Case 1: If $N_1 \leq N_2$, the secrecy capacity C_{s-both}^{g} is given by:

$$C_{s-both}^{g} = \min\left\{ \begin{array}{l} \frac{1}{2}\ln(1+\frac{P}{N_1}), \\ \frac{1}{2}\ln(1+\frac{P}{N_1}) + \frac{1}{2}\ln(2\pi eQ) - \frac{1}{2}\ln(\frac{P+Q+N_2}{N_2}) \end{array} \right\}. \tag{32}$$

Case 2: If $N_1 > N_2$, a lower bound C_{s-both}^{gi} on the secrecy capacity C_{s-both}^{g} is given by:

$$C_{s-both}^{g} \geq C_{s-both}^{gi} = \min\{\frac{1}{2}\ln\frac{2\pi eQN_2}{Q+N_2}, \frac{1}{2}\ln(1+\frac{P}{N_1})\}. \tag{33}$$

Remark 4.

For the Gaussian case, the conditional mutual information $I(X;Y|V)$ is calculated by using the fact that when the CSI is known by both the legitimate receiver and the transmitter, it can be simply subtracted off, which in effect reduces the channel to a Gaussian channel with no CSI, *i.e.*, $I(X;Y|V) = \frac{1}{2}\ln(1+\frac{P}{N_1})$. Analogously, we have $I(X;Z|V) = \frac{1}{2}\ln(1+\frac{P}{N_2})$. Then, it is easy to see that Y is less noisy than Z ($I(X;Z|V) \geq I(X;Y|V)$ for every $P_{X|V}(x|v)$), which can be further expressed

by $N_1 \leq N_2$, and Z is less noisy than Y ($I(X;Z|V) \geq I(X;Y|V)$ for every $P_{X|V}(x|v)$), which can be further expressed by $N_1 \geq N_2$.

Proof. The achievability proof of (32) and (33) is easily obtained by substituting $X \sim \mathcal{N}(0,P)$, $V \sim \mathcal{N}(0,Q)$ and (20) into (30) and (31), respectively. Now, it remains to prove the converse of (32); see the following.

The converse part of (32) is based on the converse proof of (30), (see p.2846, Proof of Theorem 2 of [13] and the left bottom and right top of page 2841[13]). However, the converse proof of (30) is for the discrete memoryless case, and it needs to be further processed for the Gaussian case. Based on the converse proof of (30) [13] and the fact that $N_1 \leq N_2$, we have the following (34) and (35),

$$
\begin{aligned}
C^{g}_{s-both} &\leq \frac{1}{N}\sum_{i=1}^{N}(I(X_i;Y_i|V_i) - I(X_i;Z_i|V_i) + h(V_i|Z_i)) \\
&\overset{(1)}{=} \frac{1}{N}\sum_{i=1}^{N}(I(X_i;Y_i|V_i) - h(Z_i) + h(V_i) + h(Z_i|X_i,V_i)) \\
&\overset{(2)}{=} \frac{1}{N}\sum_{i=1}^{N}(h(X_i+Z_{1,i}|V_i) - h(Z_{1,i}) - h(Z_i) + h(V_i) + h(Z_{1,i}+Z^*_{2,i})) \\
&\leq \frac{1}{N}\sum_{i=1}^{N}(h(X_i+Z_{1,i}) - h(Z_{1,i}) - h(Z_i) + h(V_i) + h(Z_{1,i}+Z^*_{2,i})) \\
&\overset{(3)}{=} \frac{1}{N}\sum_{i=1}^{N}(h(X_i+Z_{1,i}) - \frac{1}{2}\ln(2\pi e N_1) - h(X_i+V_i+Z_{1,i}+Z^*_{2,i}) + \frac{1}{2}\ln(2\pi e Q) \\
&\quad + \frac{1}{2}\ln(2\pi e N_2)) \\
&\overset{(4)}{\leq} \frac{1}{N}\sum_{i=1}^{N}(\frac{1}{2}\ln(e^{2h(X_i+Z_{1,i})}) - \frac{1}{2}\ln(2\pi e N_1) - \frac{1}{2}\ln(e^{2h(X_i+Z_{1,i})} + e^{2h(V_i+Z^*_{2,i})}) \\
&\quad + \frac{1}{2}\ln(2\pi e Q) + \frac{1}{2}\ln(2\pi e N_2)) \\
&\overset{(5)}{=} \frac{1}{N}\sum_{i=1}^{N}(\frac{1}{2}\ln(e^{2h(X_i+Z_{1,i})}) - \frac{1}{2}\ln(2\pi e N_1) - \frac{1}{2}\ln(e^{2h(X_i+Z_{1,i})} + 2\pi e(Q+N_2-N_1)) \\
&\quad + \frac{1}{2}\ln(2\pi e Q) + \frac{1}{2}\ln(2\pi e N_2)) \\
&= \frac{1}{N}\sum_{i=1}^{N}(\frac{1}{2}\ln\frac{e^{2h(X_i+Z_{1,i})}}{e^{2h(X_i+Z_{1,i})} + 2\pi e(Q+N_2-N_1)} - \frac{1}{2}\ln(2\pi e N_1) + \frac{1}{2}\ln(2\pi e Q) + \frac{1}{2}\ln(2\pi e N_2)) \\
&\overset{(6)}{\leq} \frac{1}{N}\sum_{i=1}^{N}(\frac{1}{2}\ln\frac{2\pi e(P+N_1)}{2\pi e(P+N_1) + 2\pi e(Q+N_2-N_1)} - \frac{1}{2}\ln(2\pi e N_1) \\
&\quad + \frac{1}{2}\ln(2\pi e Q) + \frac{1}{2}\ln(2\pi e N_2)) \\
&= \frac{1}{2}\ln(1+\frac{P}{N_1}) + \frac{1}{2}\ln(2\pi e Q) - \frac{1}{2}\ln(\frac{P+Q+N_2}{N_2}),
\end{aligned}
\tag{34}
$$

and:

$$\begin{aligned} C_{s-both}^{g} &\leq \frac{1}{N}\sum_{i=1}^{N} I(X_i;Y_i|V_i) \\ &= \frac{1}{N}\sum_{i=1}^{N}(h(Y_i|V_i)-h(Y_i|V_i,X_i)) \\ &\overset{(7)}{=} \frac{1}{N}\sum_{i=1}^{N}(h(X_i+Z_{1,i}|V_i)-h(Z_{1,i})) \\ &\leq \frac{1}{N}\sum_{i=1}^{N}(h(X_i+Z_{1,i})-h(Z_{1,i})) \\ &\overset{(8)}{\leq} \frac{1}{N}\sum_{i=1}^{N}(\frac{1}{2}\ln(2\pi e(P+N_1))-\frac{1}{2}\ln(2\pi eN_1)) \\ &= \frac{1}{2}\ln(1+\frac{P}{N_1}), \end{aligned} \tag{35}$$

where (1) is from (29), (2) is from Definition (23) and $Z_{2,i}^{*} \sim \mathcal{N}(0, N_2 - N_1)$, (3) is from the fact that the differential entropy of a Gaussian distributed random variable X is $h(X) = \frac{1}{2}\ln(2\pi eD(X))$ (here, $D(X)$ is the variance of the Gaussian random variable X), (4) is from the entropy power inequality $e^{2h(X_i+V_i+Z_{1,i}+Z_{2,i})} \geq e^{2h(X_i+Z_{1,i})} + e^{2h(V_i+Z_{2,i})}$ (see [19]), (5) is from $h(V_i+Z_{2,i}) = \frac{1}{2}\ln(2\pi e(Q+N_2))$, (6) is from $\frac{1}{2}\ln\frac{e^{2h(X_i+Z_{1,i})}}{e^{2h(X_i+Z_{1,i})}+2\pi e(Q+N_2)}$ increasing while $h(X_i+Z_{1,i})$ is increasing and the fact that $h(X_i+Z_{1,i}) \leq \frac{1}{2}\ln(2\pi e(P+N_1))$ (here, note that "=" is achieved if $X_i \sim \mathcal{N}(0,P)$), (7) is from Definition (23) and (8) is from $h(X_i+Z_{1,i}) \leq \frac{1}{2}\ln(2\pi e(P+N_1))$. Thus, the converse part of (32) is proven. The proof of Theorem 4 is completed. □

Recall that for the degraded Gaussian wiretap channel with noncausal CSI at the transmitter $((X,V) \to Y \to Z)$, an achievable secrecy rate (a lower bound on the secrecy capacity) is provided [10]; see the following Theorem 5.

Theorem 5. *For the Gaussian non-feedback model of Figure 1 with the condition that $N_1 \leq N_2$, an achievable secrecy rate C_s^{gi} is denoted by:*

$$C_s^{gi} = \max_{0\leq\alpha\leq 1} \min \left\{ \begin{array}{l} \frac{1}{2}\ln\frac{(P+N_1)(P+\alpha^2Q)}{\alpha^2Q(P+N_1)+N_1P} - \frac{1}{2}\ln\frac{P+\alpha^2Q}{P}, \\ \frac{1}{2}\ln\frac{(P+N_1)(P+\alpha^2Q)}{\alpha^2Q(P+N_1)+N_1P} - \frac{1}{2}\ln\frac{(P+N_2)(P+\alpha^2Q)}{\alpha^2QP+N_2(P+\alpha^2Q)} \end{array} \right\}.$$

Proof. The result is directly obtained from [10], and therefore, the proof is omitted here. □

Remark 5.

- For the case $N_1 \leq N_2$, the relationship (20) of the channel inputs and outputs can be equivalently characterized by (23), which implies the Markov chain $(X,V) \to Y \to Z$.
- To the best of the authors' knowledge, for the case $N_1 > N_2$, the bounds on the secrecy capacity of the Gaussian wiretap channel with noncausal CSI at the transmitter are still unknown.

Finally, note that if the CSI is not available at the legitimate receiver, the wiretapper and the transmitter and there is no feedback link from the legitimate receiver to the transmitter, the Gaussian case of the model of Figure 1 (see (20)) reduces to the model of the Gaussian wiretap channel, where V_i and $Z_{1,i}$ of (20) are the legitimate receiver's channel noises and V_i and $Z_{2,i}$ are the wiretapper's channel noises. From [20], it is easy to see that the secrecy capacity C_s^{*} of the Gaussian wiretap channel is given by:

$$C_s^{*} = \frac{1}{2}\ln\frac{P+Q+N_1}{Q+N_1} - \frac{1}{2}\ln\frac{P+Q+N_2}{Q+N_2}. \tag{36}$$

Comparing Theorem 3 to Theorem 4, we can conclude that if $N_1 \leq N_2$, for given P, N_1 and N_2, C_s^{gf} is larger than C_{s-both}^{g} if and only if:

$$Q \leq \frac{N_1(N_2 - N_1)(P + N_1)}{N_1^2 + N_2 P}. \tag{37}$$

For the case that $N_1 > N_2$, we find that if $\frac{N_1}{2} < N_2 < N_1$, for given P, N_1 and N_2, C_s^{gf} is larger than C_{s-both}^{gi} if and only if:

$$Q \leq \frac{N_2(N_1 - N_2)}{2N_2 - N_1}. \tag{38}$$

If $N_2 = \frac{N_1}{2}$, C_s^{gf} is always larger than C_{s-both}^{gi}.
If $N_2 < \frac{N_1}{2}$, for given P, N_1 and N_2, C_s^{gf} is larger than C_{s-both}^{gi} if and only if:

$$Q \geq \frac{N_2(N_1 - N_2)}{2N_2 - N_1}. \tag{39}$$

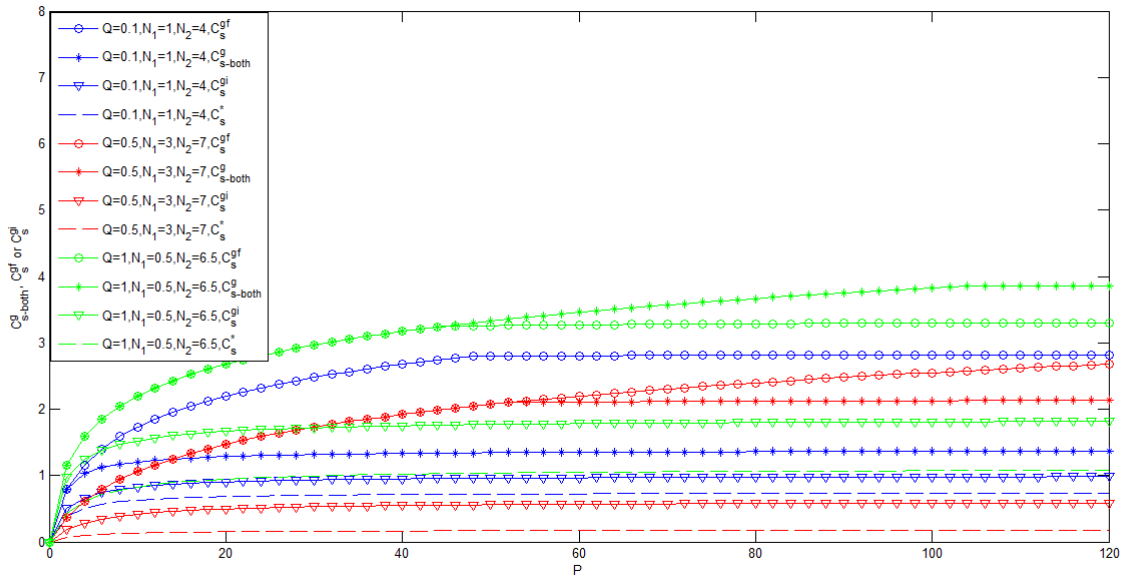

Figure 2. For $N_1 \leq N_2$, the relationships of $P - C_s^{gf}$, $P - C_{s-both}^{g}$, $P - C_s^{gi}$ and $P - C_s^*$ for several values of N_1, N_2 and Q.

For the case $N_1 \leq N_2$, Figure 2 plots the relationships of $P - C_s^*$, $P - C_s^{gi}$, $P - C_s^{gf}$ and $P - C_{s-both}^{g}$ for several values of N_1, N_2 and Q. It is easy to see that the noiseless feedback (C_s^{gf}), the CSI sharing scheme (C_{s-both}^{g}) and the CSI only available at the transmitter (C_s^{gi}) help to enhance the secrecy capacity C_s^* of the Gaussian wiretap channel. Furthermore, we can see that both the noiseless feedback and the CSI sharing scheme perform better than the CSI only available at the transmitter. Moreover, when Q is small ($Q = 0.1, 0.5$), the noiseless feedback performs better than the CSI sharing scheme, and while Q is increasing ($Q = 1$), the CSI sharing scheme is beginning to take advantage of the noiseless feedback.

For the case $N_1 > N_2$, the following Figure 3 plots the relationships of $P - C_s^{gf}$ and $P - C_{s-both}^{g}$ for several values of N_1, N_2 and Q. Since $C_s^* = 0$ for the case that $N_1 > N_2$, both the noiseless feedback (C_s^{gf}) and the CSI sharing scheme (C_{s-both}^{gi}) enhance the secrecy capacity C_s^* of the Gaussian wiretap channel. Moreover, we can see that for fixed Q, if the gap between the legitimate receiver's channel

noise variance N_1 and the wiretapper's channel noise variance N_2 is large, the noiseless feedback performs better than the CSI sharing scheme, and *vice versa*.

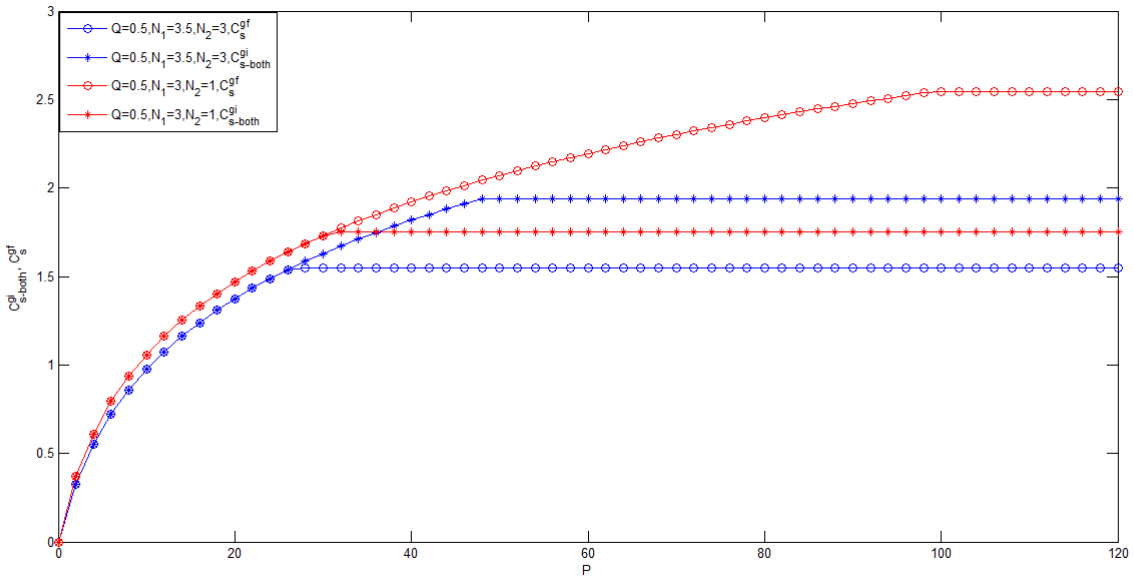

Figure 3. For $N_1 > N_2$, the relationships of $P - C_s^{gf}$ and $P - C_{s-both}^{gi}$ for several values of N_1, N_2 and Q.

3.2. Binary Case of the Model of Figure 1

In this subsection, we calculate the secrecy capacity of a degraded binary case of the model of Figure 1 with causal CSI at the transmitter, where "degraded" means that there exists a Markov chain $(X, V) \to Y \to Z$.

Suppose that the random variable V is uniformly distributed over $\{0, 1\}$, *i.e.*, $p_V(0) = p_V(1) = \frac{1}{2}$. Meanwhile, the random variables X, Y and Z take values in $\{0, 1\}$, and the wiretap channel is a BSC (binary symmetric channel) with crossover probability q. The transition probability of the main channel is defined as follows:

When $v = 0$,

$$p_{Y|X,V}(y|x, v = 0) = \begin{cases} 1 - p, & \text{if } y = x, \\ p, & \text{otherwise.} \end{cases} \tag{40}$$

When $v = 1$,

$$p_{Y|X,V}(y|x, v = 1) = \begin{cases} p, & \text{if } y = x, \\ 1 - p, & \text{otherwise.} \end{cases} \tag{41}$$

From Remark 2, we know that the secrecy capacity for the model of Figure 1 with causal CSI at the transmitter is given by:

$$C_s^{(cf)} = \max_{P_K(k)P_{X|K,V}(x|k,v)} \min\{I(K;Y), H(Y|Z)\} \tag{42}$$

and the maximum achievable secrecy rate $C_s^{(ci)}$ of the wiretap channel with causal CSI [12] is given by:

$$C_s^{(ci)} = \max_{P_K(k)P_{X|K,V}(x|k,v)} (I(K;Y) - I(K;Z)), \tag{43}$$

where (43) is from (19).

In addition, from ([13], Theorem 3), we know that the secrecy capacity C_{s-both} of the wiretap channel with CSI causally or non-causally at both the transmitter and the legitimate receiver is given by:

$$C_{s-both} = \max_{P_{X|V}(x|v)} \min\{I(X;Y|V) - I(X;Z|V) + H(V|Z), I(X;Y|V)\}. \tag{44}$$

It remains to calculate $C_s^{(cf)}$, $C_s^{(ci)}$ and C_{s-both}; see the following.

The calculation of $C_s^{(cf)}$ and $C_s^{(ci)}$:

Let K take values in $\{0,1\}$. The probability of K is defined as follows. $p_K(0) = \alpha$, and $p_K(1) = 1-\alpha$. Define the conditional probability mass function $p_{X|K,V}$ as follows.

$p_{X|K,V}(0|0,0) = \beta_1$, $p_{X|K,V}(1|0,0) = 1-\beta_1$, $p_{X|K,V}(0|0,1) = \beta_2$, $p_{X|K,V}(1|0,1) = 1-\beta_2$,
$p_{X|K,V}(0|1,0) = \beta_3$, $p_{X|K,V}(1|1,0) = 1-\beta_3$, $p_{X|K,V}(0|1,1) = \beta_4$, $p_{X|K,V}(1|1,1) = 1-\beta_4$.

The joint probability mass functions p_{KY} is calculated by:

$$\begin{aligned} p_{KY}(k,y) &= \sum_{x,v} p_{KYXV}(k,y,x,v) \\ &= \sum_{x,v} p_{Y|XV}(y|x,v) p_{X|K,V}(x|k,v) p_K(k) p_V(v). \end{aligned} \tag{45}$$

Then, we have:

$$p_{KY}(0,0) = \frac{\alpha}{2}[1-(\beta_1-\beta_2)(1-2p)], \tag{46}$$

$$p_{KY}(0,1) = \frac{\alpha}{2}[1+(\beta_1-\beta_2)(1-2p)], \tag{47}$$

$$p_{KY}(1,0) = \frac{\alpha}{2}[1-(\beta_3-\beta_4)(1-2p)], \tag{48}$$

$$p_{KY}(1,1) = \frac{\alpha}{2}[1+(\beta_3-\beta_4)(1-2p)]. \tag{49}$$

By calculating, we have:

$$C_s^{(cf)} = \min\{1-h(p), h(q)\}, \tag{50}$$

and:

$$C_s^{(ci)} = h(p+q-2pq) - h(p), \tag{51}$$

where $h(x) = -x\log x - (1-x)\log(1-x)$ and $0 \le x \le 1$.

The calculation of C_{s-both}:

Define $p_{X|V}(0|0) = \alpha$, $p_{X|V}(1|0) = 1-\alpha$, $p_{X|V}(0|1) = \beta$, $p_{X|V}(1|1) = 1-\beta$.

By calculating, C_{s-both} is given by:

$$C_{s-both} = \min\{1-h(p), 1-h(p)+h(p+q-2pq)\} = 1-h(p). \tag{52}$$

The following Figures 4–6 show $C_s^{(cf)}$, $C_s^{(ci)}$ and C_{s-both} for several values of q. Here, note that the noise of the wiretap channel is increasing while q is increasing. It is easy to see that when $q < 0.5$, C_{s-both} and $C_s^{(cf)}$ are always larger than $C_s^{(ci)}$, *i.e.*, both the noiseless feedback (the model of this paper) and the shared CSI [13] help to enhance the security of the wiretap channel with causal CSI at the transmitter. When $q = 0.5$, there is no wiretapper in the channel; thus, $C_s^{(cf)} = C_s^{(ci)} = C_{s-both} = 1-h(p)$.

Moreover, from Figures 4–6, we see that the noiseless feedback performs no better than the shared CSI. However, when q is large enough (satisfying $h(q) \ge 1-h(p)$), the two ways perform the same.

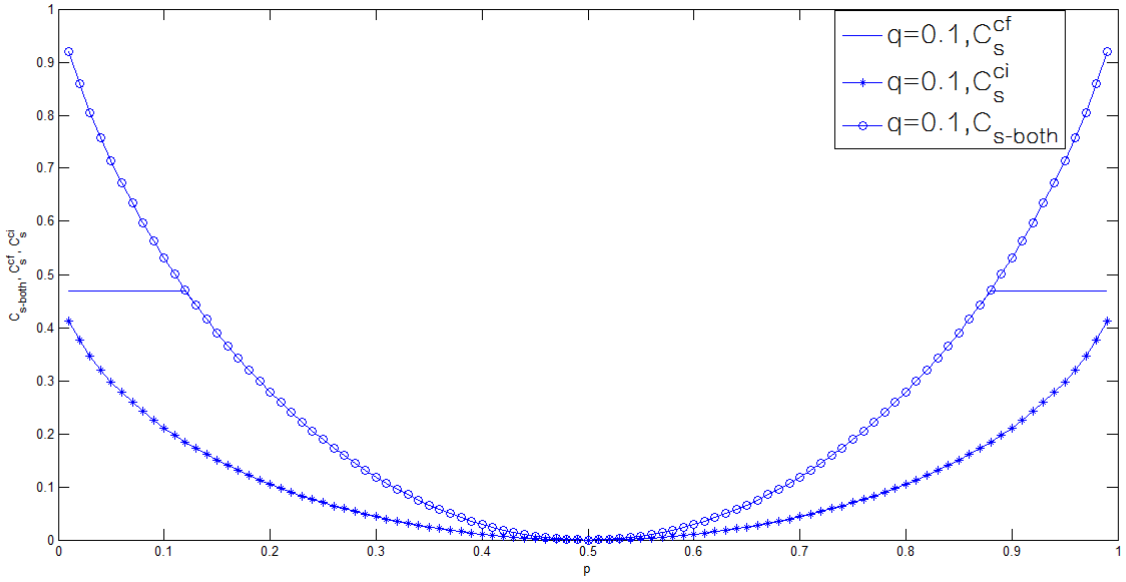

Figure 4. The $C_s^{(cf)}$, $C_s^{(ci)}$ and C_{s-both} for $q = 0.1$.

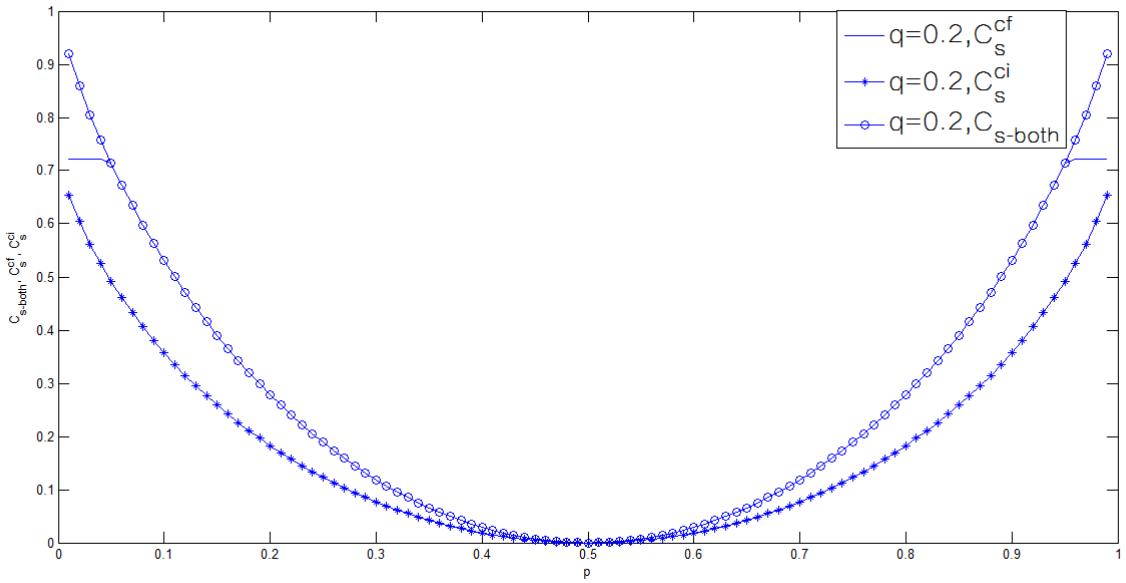

Figure 5. The $C_s^{(cf)}$, $C_s^{(ci)}$ and C_{s-both} for $q = 0.2$.

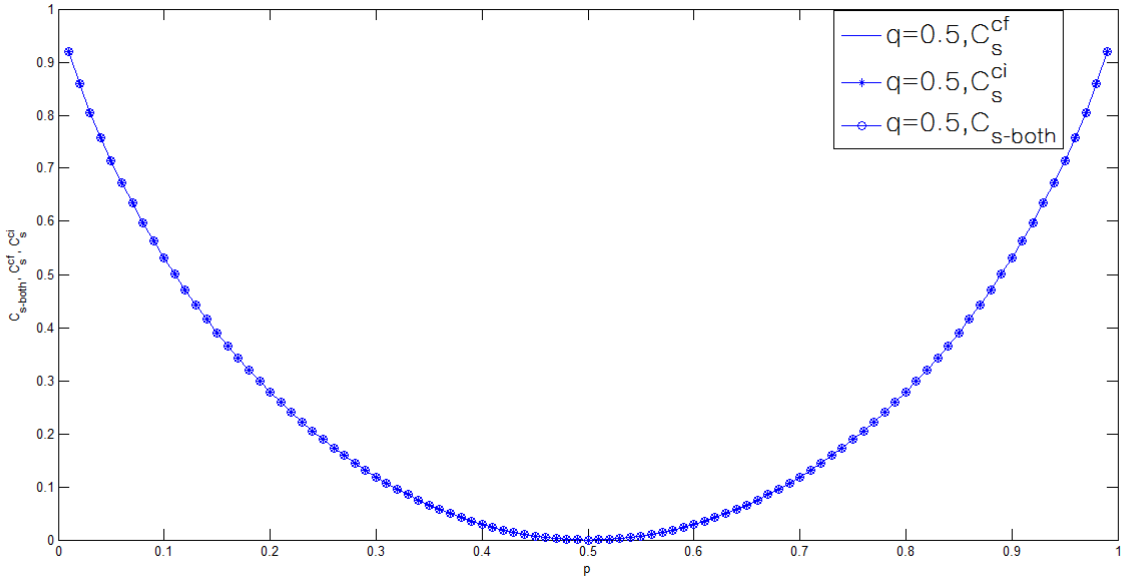

Figure 6. The $C_s^{(cf)}$, $C_s^{(ci)}$ and C_{s-both} for $q = 0.5$.

4. Conclusions

In this paper, we study the general wiretap channel with CSI and noiseless feedback, where the CSI is available at the transmitter in a noncausal or causal manner. Both the capacity-equivocation region and the secrecy capacity are determined for the noncausal and causal cases, and the results are further explained via Gaussian and binary examples. For the Gaussian example, we show that both the noiseless feedback and the CSI sharing scheme [13] help to enhance the security of the Gaussian wiretap channel. Moreover, we show that in some particular cases, the noiseless feedback performs even better than the CSI sharing scheme [13]. For the degraded binary example, we also find that the noiseless feedback enhances the security of the wiretap channel with causal CSI. Unlike the Gaussian example, we find that the noiseless feedback always performs no better than the CSI sharing scheme [13].

Acknowledgment: The authors would like to thank the anonymous reviewers for their valuable suggestions to improve this paper. This work was supported by a sub-project in the 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 61301121, the Fundamental Research Fund for the Central Universities under Grant 2682014CX099, the Key Grant Project of Chinese Ministry of Education (No. 311031 100), the Young Innovative Research Team of Sichuan Province (2011JTD0007) and the Open Research Fund of National Mobile Communications Research Laboratory, Southeast University (No. 2014D01).

Author Contributions: Bin Dai designed research; Bin Dai and Zheng Ma performed research; Linman Yu analyzed the data; Bin Dai wrote the paper. All authors have read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix

A. Converse Proof of Theorem 1

Given an achievable (R, R_e) pair, we need to show that there exists a joint distribution of the form $P_{Z|Y}(z|y)P_{Y|XV}(y|x,v)P_{X|KV}(x|k,v)P_{KV}(k,v)$, such that,

$$0 \leq R_e \leq R, \tag{A.1}$$

$$0 \leq R \leq I(K;Y) - I(K;V), \tag{A.2}$$

$$R_e \leq H(Y|Z). \tag{A.3}$$

A.1. Proof of (A.1)

$$R_e \leq \lim_{N\to\infty} \Delta \leq \lim_{N\to\infty} \frac{1}{N} H(W) = \lim_{N\to\infty} \frac{\log \| \mathcal{W} \|}{N} = R.$$

A.2. Proof of (A.2)

$$\begin{aligned}
\frac{1}{N}H(W) &= \frac{1}{N}(I(W;Y^N) + H(W|Y^N)) \overset{(a)}{\leq} \frac{1}{N}(I(W;Y^N) + \delta(P_e)) \\
&\overset{(b)}{=} \frac{1}{N}(I(W;Y^N) - I(W;V^N) + \delta(P_e)) \\
&\overset{(c)}{=} \frac{1}{N}\sum_{i=1}^{N}(I(Y_i;W,V_{i+1}^N|Y^{i-1}) - I(V_i;W,Y^{i-1}|V_{i+1}^N) + \delta(P_e)) \\
&\overset{(d)}{\leq} \frac{1}{N}\sum_{i=1}^{N}(H(Y_i) - H(Y_i|Y^{i-1},W,V_{i+1}^N) - H(V_i) + H(V_i|V_{i+1}^N,W,Y^{i-1}) + \delta(P_e)) \\
&= \frac{1}{N}\sum_{i=1}^{N}(I(Y_i;W,V_{i+1}^N,Y^{i-1}) - I(V_i;W,Y^{i-1},V_{i+1}^N) + \delta(P_e)) \\
&\overset{(e)}{=} \frac{1}{N}\sum_{i=1}^{N}(I(Y_i;W,V_{i+1}^N,Y^{i-1}|J=i) - I(V_i;W,Y^{i-1},V_{i+1}^N|J=i) + \delta(P_e)) \\
&\overset{(f)}{=} I(Y_J;W,V_{J+1}^N,Y^{J-1}|J) - I(V_J;W,Y^{J-1},V_{J+1}^N|J) + \frac{\delta(P_e)}{N} \\
&\overset{(g)}{\leq} I(Y_J;W,V_{J+1}^N,Y^{J-1},J) - I(V_J;W,Y^{J-1},V_{J+1}^N,J) + \frac{\delta(P_e)}{N} \\
&\overset{(h)}{=} I(K;Y) - I(K;V) + \frac{\delta(P_e)}{N},
\end{aligned} \tag{A.4}$$

where (a) is from Fano's inequality, (b) is from W is independent of V^N, (c) is from Csiszár's equality:

$$\sum_{i=1}^{N} I(Y_i;V_{i+1}^N|Y^{i-1},W) = \sum_{i=1}^{N} I(V_i;Y^{i-1}|V_{i+1}^N,W), \tag{A.5}$$

(d) is from V_i being independent of V_{i+1}^N, (e) and (f) are from J being a random variable (uniformly distributed over $[1,N]$) and being independent of W, V^N and Y^N, (g) is from V_J being independent of J and (h) is from the definitions that $Y \triangleq Y_J$, $V \triangleq V_J$ and $K \triangleq (W, Y^{J-1}, V_{J+1}^N, J)$.

By using $P_e \leq \epsilon$, $\epsilon \to 0$ as $N \to \infty$, $\lim_{N\to\infty} \frac{H(W)}{N} = R$ and (A.4), it is easy to see that $R \leq I(K;Y) - I(K;V)$.

A.3. Proof of (A.3)

$$\begin{aligned}\frac{1}{N}H(W|Z^N) &\overset{(1)}{\leq} \frac{1}{N}(I(W;Y^N|Z^N)+\delta(P_e))\\ &\leq \frac{1}{N}\sum_{i=1}^{N}H(Y_i|Z_i)+\frac{\delta(P_e)}{N}\\ &\overset{(2)}{=} \frac{1}{N}\sum_{i=1}^{N}H(Y_i|Z_i,J=i)+\frac{\delta(P_e)}{N}\\ &\overset{(3)}{\leq} H(Y_J|Z_J,J)+\frac{\delta(P_e)}{N}\\ &\overset{(4)}{\leq} H(Y|Z)+\frac{\delta(P_e)}{N},\end{aligned} \tag{A.6}$$

where (1) is from Fano's inequality, (2) is from J being a random variable (uniformly distributed over $\{1,2,...,N\}$) and being independent of Y^N and Z^N, (3) is from J being uniformly distributed over $\{1,2,...,N\}$ and (4) is from the definitions that $Y \triangleq Y_J$, and $Z \triangleq Z_J$.

By using $P_e \leq \epsilon$, $\epsilon \to 0$ as $N \to \infty$, $\lim_{N\to\infty}\frac{H(W|Z^N)}{N} \geq R_e$ and (A.6), it is easy to see that $R_e \leq H(Y|Z)$.

The converse proof of Theorem 1 is completed.

B. Direct Proof of Theorem 1

The direct part (achievability) of Theorem 1 is proven by considering the following two cases.

- Case 1: If $I(K;Y)-I(K;V) \geq H(Y|Z)$, we need to show that $(R = I(K;Y)-I(K;V)-\epsilon$, $R_e = H(Y|Z))$ is achievable, where $\epsilon \to 0^+$.
- Case 2: If $I(K;Y)-I(K;V) \leq H(Y|Z)$, we need to show that $(R = I(K;Y)-I(K;V)-\epsilon$, $R_e = R = I(K;Y)-I(K;V)-\epsilon)$ is achievable.

The direct proof of Theorem 1 is organized as follows. The balanced coloring lemma introduced by Ahlswede and Cai is provided in Subsection B.1, and it will be used in the remainder of this section. The code-book generation is shown in Subsection B.2, and the equivocation analysis is given in Subsection B.3.

B.1. The Balanced Coloring Lemma

The balanced coloring lemma was first introduced by Ahlswede and Cai; see the following.

Lemma 1. *Balanced coloring lemma: For all $\epsilon_1, \epsilon_2, \epsilon_3, \delta > 0$, sufficiently large N and all N-type $P_Y(y)$, there exists a γ-coloring $c : T_Y^N(\epsilon_1) \to \{1,2,..,\gamma\}$ of $T_Y^N(\epsilon_1)$ such that for all joint N-type $P_{YZ}(y,z)$ with marginal distribution $P_Z(z)$ and $\frac{|T_{Y|Z}^N(z^N)|}{\gamma} > 2^{N\epsilon_2}$, $z^N \in T_Z^N(\epsilon_3)$,*

$$|c^{-1}(k)| \leq \frac{|T_{Y|Z}^N(z^N)|(1+\delta)}{\gamma}, \tag{B.1}$$

for $k = 1,2,...,\gamma$, where c^{-1} is the inverse image of c.

Proof. Letting $U = const$, Lemma 1 is directly from p. 259 of [1], and thus, we omit it here. □

Lemma 1 shows that if y^N and z^N are joint typical, for given z^N, the number of $y^N \in T^N_{Y|Z}(z^N)$ for a certain color k ($k = 1, 2, ..., \gamma$), which is denoted as $|c^{-1}(k)|$, is upper bounded by $\frac{|T^N_{Y|Z}(z^N)|(1+\delta)}{\gamma}$. By using Lemma 1, it is easy to see that the typical set $T^N_{Y|Z}(z^N)$ maps into at least:

$$\frac{|T^N_{Y|Z}(z^N)|}{\frac{|T^N_{Y|Z}(z^N)|(1+\delta)}{\gamma}} = \frac{\gamma}{1+\delta} \tag{B.2}$$

colors. On the other hand, the typical set $T^N_{Y|Z}(z^N)$ maps into at most γ colors.

B.2. Code-Book Generation

Fix the joint probability mass function $P_{Z,Y|X,V}(z, y|x, v)P_{X|K,V}(x|k, v)P_{KV}(k, v)$. The message set $\mathcal{W}$ satisfies:

$$\lim_{N\to\infty} \frac{\log \|\mathcal{W}\|}{N} = R = I(K;Y) - I(K;V) - \epsilon. \tag{B.3}$$

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

The block Markov encoding scheme is used in the direct proof of Theorem 1. The random vectors K^N, V^N, X^N, Y^N and Z^N consist of n blocks of length N. Let $\tilde{K}_i$, $\tilde{V}_i$, $\tilde{Y}_i$ and $\tilde{Z}_i$ ($1 \le i \le n$) be the random vectors for block i. Define $\tilde{k}^n = (\tilde{k}_1, \tilde{k}_2, ..., \tilde{k}_n)$, $\tilde{v}^n = (\tilde{v}_1, \tilde{v}_2, ..., \tilde{v}_n)$, $\tilde{y}^n = (\tilde{y}_1, \tilde{y}_2, ..., \tilde{y}_n)$ and $\tilde{z}^n = (\tilde{z}_1, \tilde{z}_2, ..., \tilde{z}_n)$ to be the specific vectors for all blocks. The message W^n for all n blocks is denoted by $W^n = (W_1, W_2, ..., W_n)$, where W_i ($2 \le i \le n$) is uniformly distributed over the alphabet $\mathcal{W}$, and W_i is independent of W_j ($2 \le j \le n$ and $j \ne i$). Note that w_1 does not exist.

Construction of K^N:

Gel'fand and Pinsker's binning and block Markov coding scheme are used in the construction of K^N.

- Construction of K^N for Case 1:

  For each block, generate $2^{N(I(K;Y)-\epsilon_{2,N})}$ ($\epsilon_{2,N} \to 0$) i.i.d. sequences of k^N, according to $p_K(k)$. Partition these sequences at random into $2^{NR} = 2^{N(I(K;Y)-I(K;V)-\gamma_1)}$ bins, such that each bin has $2^{N(I(K;V)+\gamma_1-\epsilon_{2,N})}$ sequences. Index each bin by $l \in \{1, 2, ..., 2^{NR}\}$.

  Denote the message w_i ($2 \le i \le n$) by $w_i = (w_{i1}, w_{i2})$, where $w_{i1} \in \mathcal{W}_{i1} = \{1, 2, ..., 2^{NH(Y|Z)}\}$ and $w_{i2} \in \mathcal{W}_{i2} = \{1, 2, ..., 2^{N(R-H(Y|Z))}\}$. Here, note that W_{i1} is independent of W_{i2}.

  In the first block, for a given side information $\tilde{v}_1$, try to find a $\tilde{k}_1$, such that $(\tilde{k}_1, \tilde{v}_1) \in T^N_{KV}(\epsilon)$. If multiple sequences exist, randomly choose one for transmission. If there is no such sequence, declare an encoding error.

  For the i-th block ($2 \le i \le n$), the transmitter receives the output $\tilde{y}_{i-1}$ of the $i-1$-th block; he or she gives up if $\tilde{y}_{i-1} \notin T^N_Y(\epsilon_2)$ ($\epsilon_2 \to 0$ as $N \to \infty$). It is easy to see that the probability for giving up at the $i-1$-th block tends to zero as $N \to \infty$. In the case $\tilde{y}_{i-1} \in T^N_Y(\epsilon_2)$, generate a mapping $g_f : T^N_Y(\epsilon_2) \to \{1, 2, ..., 2^{NH(Y|Z)}\}$. Define a random variable K^*_i by $K^*_i = g_f(\tilde{Y}_{i-1})$ ($2 \le i \le n$), and it is uniformly distributed over the set $\mathcal{W}_{i1} = \{1, 2, ..., 2^{NH(Y|Z)}\}$. K^*_i is independent of W_i. Reveal the mapping g_f to the legitimate receiver, the wiretapper and the transmitter. Then, since the transmitter gets $\tilde{y}_{i-1}$, he computes $k^*_i = g_f(\tilde{y}_{i-1}) \in \{1, 2, ..., 2^{NH(Y|Z)}\}$. For a given $w_i = (w_{i1}, w_{i2})$ ($2 \le i \le n$), the transmitter selects a sequence $\tilde{k}_i$ in the bin $(w_{i1} \oplus k^*_i, w_{i2})$ (where $\oplus$ is the modulo addition over $\mathcal{W}_{i1}$), such that $(\tilde{k}_i, \tilde{v}_i) \in T^N_{KV}(\epsilon)$. If multiple sequences in bin $(w_{i1} \oplus k^*_i, w_{i2})$ exist, choose the sequence with the smallest index in the bin. If there is no such sequence, declare an encoding error. Here, note that since K^*_i is independent of $W_i = (W_{i1}, W_{i2})$, $W_{i1} \oplus K^*_i$ is independent of W_i and K^*_i. The proof is given as follows.

Proof. Since:

$$\begin{aligned} Pr\{K_i^* \oplus W_{i1} = a\} &= \sum_{k_i^* \in \mathcal{W}_{i1}} Pr\{K_i^* \oplus W_{i1} = a, K_i^* = k_i^*\} \\ &= \sum_{k_i^* \in \mathcal{W}_{i1}} Pr\{W_{i1} = a \ominus k_i^*, K_i^* = k_i^*\} \\ &= \sum_{k_i^* \in \mathcal{W}_{i1}} Pr\{W_{i1} = a \ominus k_i^*\} Pr\{K_i^* = k_i^*\} \\ &= \sum_{k_i^* \in \mathcal{W}_{i1}} \frac{1}{\|\mathcal{W}_{i1}\|^2} = \frac{1}{\|\mathcal{W}_{i1}\|}, \end{aligned} \tag{B.4}$$

and:

$$\begin{aligned} &Pr\{K_i^* \oplus W_{i1} = a, K_i^* = k_i^*\} \\ &= Pr\{W_{i1} = a \ominus k_i^*, K_i^* = k_i^*\} \\ &= Pr\{W_{i1} = a \ominus k_i^*\} Pr\{K_i^* = k_i^*\} \\ &= \frac{1}{\|\mathcal{W}_{i1}\|^2}, \end{aligned} \tag{B.5}$$

it is easy to see that $Pr\{K_i^* \oplus W_{i1} = a, K_i^* = k_i^*\} = Pr\{K_i^* \oplus W_{i1} = a\} \cdot Pr\{K_i^* = k_i^*\}$, which implies that $K_i^* \oplus W_{i1}$ is independent of K_i^*.

Analogously, we can prove that $Pr\{K_i^* \oplus W_{i1} = a, W_{i1} = w_{i1}, W_{i2} = w_{i2}\} = Pr\{K_i^* \oplus W_{i1} = a\} \cdot Pr\{W_{i1} = w_{i1}\} \cdot Pr\{W_{i2} = w_{i2}\}$, which implies that $K_i^* \oplus W_{i1}$ is independent of $W_i = (W_{i1}, W_{i2})$. Thus, the proof of $W_{i1} \oplus K_i^*$ is independent of W_i, and K_i^* is completed.

□

- Construction of K^N for Case 2: The construction of K^N for Case 2 is similar to that of Case 1, except that there is no need to divide w_i into two parts. The detail is as follows. For the i-th block ($2 \leq i \leq n$), if $\tilde{y}_{i-1} \in T_Y^N(\epsilon_2)$, generate a mapping $g_f : T_Y^N(\epsilon_2) \to \mathcal{W}$ (note that $|T_Y^N(\epsilon_2)| \geq |\mathcal{W}|$). Let $K_i^* = g_f(\tilde{Y}_{i-1})$ ($2 \leq i \leq n$), and it is uniformly distributed over the set $\mathcal{W}$. K_i^* is independent of W_i. Reveal the mapping g_f to the legitimate receiver, the wiretapper and the transmitter. When the transmitter receives the feedback $\tilde{y}_{i-1}$ of the $i-1$-th block, he or she computes $k_i^* = g_f(\tilde{y}_{i-1}) \in \mathcal{W}$. For a given transmitted message w_i ($2 \leq i \leq n$), the transmitter selects a codeword $\tilde{k}_i$ in the bin $w_i \oplus k_i^*$ (where $\oplus$ is the modulo addition over $\mathcal{W}$), such that $(\tilde{k}_i, \tilde{v}_i) \in T_{KV}^N(\epsilon)$. If multiple sequences in bin $w_i \oplus k_i^*$ exist, select the one with the smallest index in the bin. If there is no such sequence, declare an encoding error. Here, note that $W_i \oplus K_i^*$ is independent of W_i and K_i^*, and the proof is similar to that of Case 1. Thus, we omit the proof here.

Construction of X^N:

In each block, the channel input x^N is generated by a pre-fixed discrete memoryless channel with transition probability $P_{X|K,V}(x|k,v)$. The inputs of the channel are k^N and v^N, and the output is x^N.

Here, note that for Case 1, the random vector $\tilde{K}_i$ of block i ($2 \leq i \leq n$) is i.i.d. generated corresponding to the encrypted message $(W_{i1} \oplus K_i^*, W_{i2})$ and $\tilde{V}_i$ (here, $\tilde{V}_i$ is also i.i.d. generated according to the probability mass function $P_V(v)$). Since $\tilde{Y}_i$ and $\tilde{Z}_i$ are generated according to $\tilde{K}_i$, $\tilde{V}_i$ and the discrete memoryless channel, the only connection between $(W_i, \tilde{K}_i, \tilde{V}_i, \tilde{Y}_i, \tilde{Z}_i)$ of the i-th block and $(W_{i-1}, \tilde{K}_{i-1}, \tilde{V}_{i-1}, \tilde{Y}_{i-1}, \tilde{Z}_{i-1})$ of the $i-1$-th block is the secret key K_i^*, which is generated by $\tilde{Y}_{i-1}$. As stated above, both the encrypted message $(W_{i1} \oplus K_i^*, W_{i2})$ and the real message $W_i = (W_{i1}, W_{i2})$ are independent of K_i^*, and thus, $(W_i, \tilde{K}_i, \tilde{V}_i, \tilde{Y}_i, \tilde{Z}_i)$ of the i-th block are independent of $(W_{i-1}, \tilde{K}_{i-1}, \tilde{V}_{i-1}, \tilde{Y}_{i-1}, \tilde{Z}_{i-1})$ of the $i-1$-th block. Since $(W_{i1} \oplus K_i^*, W_{i2})$ and W_i are also independent of W_j and K_j^* ($2 \leq i, j \leq n$ and $j \neq i$), it is easy

to see that $(W_i, \tilde{K}_i, \tilde{V}_i, \tilde{Y}_i, \tilde{Z}_i)$ are independent of $(W_j, \tilde{K}_j, \tilde{V}_j, \tilde{Y}_j, \tilde{Z}_j)$. Finally, note that $(W_{i1} \oplus K_i^*, W_{i2})$ $(2 \leq i \leq n)$ is independent of K_2^* (generated by $\tilde{Y}_1$); thus, $(W_i, \tilde{K}_i, \tilde{V}_i, \tilde{Y}_i, \tilde{Z}_i)$ are independent of $(\tilde{K}_1, \tilde{V}_1, \tilde{Y}_1, \tilde{Z}_1)$.

Analogously, in Case 2, for $2 \leq i, j \leq n$ and $j \neq i$, the fact that $(W_i, \tilde{K}_i, \tilde{V}_i, \tilde{Y}_i, \tilde{Z}_i)$ are independent of $(W_j, \tilde{K}_j, \tilde{V}_j, \tilde{Y}_j, \tilde{Z}_j)$ and $(\tilde{K}_1, \tilde{V}_1, \tilde{Y}_1, \tilde{Z}_1)$ also holds.

Decoding: For block i $(2 \leq i \leq n)$, given a vector $\tilde{y}_i \in \mathcal{Y}^N$, try to find a sequence $\tilde{k}_i(\hat{w}_{i1} \oplus k_i^*, \hat{w}_{i2}, \hat{j})$ (Case 1) or $\tilde{k}_i(\hat{w}_i \oplus k_i^*, \hat{j})$ (Case 2), such that $\tilde{k}_i$ and $\tilde{y}_i$ are joint typical. If there exists a unique sequence, put out the corresponding index of the bin $(\hat{w}_{i1} \oplus k_i^*, \hat{w}_{i2})$ or $\hat{w}_i \oplus k_i^*$. Otherwise, declare a decoding error. Since the legitimate receiver has k_i^*, put out the corresponding $\hat{w}_i$ from $(\hat{w}_{i1} \oplus k_i^*, \hat{w}_{i2})$ or $\hat{w}_i \oplus k_i^*$.

B.3. Proof of Achievability

Here, note that the above encoding-decoding scheme for the achievability proof of Theorem 1 is exactly the same as that in [11], except that the transmitter transmits an "encrypted message" by using the secret key k_i^*. Since the legitimate receiver has k_i^*, the decoding scheme for the achievability proof of Theorem 1 is in fact the same as that in [11]. Hence, we omit the proof of $P_e \leq \epsilon$ here. It remains to prove that $\lim_{N\to\infty} \Delta \geq R_e$; see the following.

- For Case 1, part of the message w_i is encrypted by k_i^*. In the analysis of the equivocation, we drop w_{i2} from w_i. Then, the equivocation about w_i is equivalent to the equivocation about k_i^*. Since $k_i^* = g_f(\tilde{y}_{i-1})$, the wiretapper tries to guess k_i^* from $\tilde{y}_{i-1}$. Note that for a given $\tilde{z}_{i-1}$ and sufficiently large N, $Pr\{\tilde{y}_{i-1} \in T_{Y|Z}^N(\tilde{z}_{i-1})\} \to 1$. Thus, the wiretapper can guess $\tilde{y}_{i-1}$ from the conditional typical set $T_{Y|Z}^N(\tilde{z}_{i-1})$. By using the above Lemma 1 and (B.2), the set $T_{Y|Z}^N(\tilde{z}_{i-1})$ maps into at least $\frac{2^{NH(Y|Z)}}{1+\delta}$ (here, $\gamma = 2^{NH(Y|Z)}$) k_i^* (colors). Thus, in the i-th block, the uncertainty about K_i^* is bounded by:

$$\frac{1}{N} H(K_i^* | \tilde{Z}_{i-1}) \geq H(Y|Z) - \frac{\log(1+\delta)}{N}, \tag{B.6}$$

  Here, note that K_i^* is uniformly distributed.
- For Case 2, the alphabet of the secret key k_i^* equals the alphabet $\mathcal{W}_i = \{1, 2, ..., 2^{NR}\}$, and the encrypted message is denoted by $w_i \oplus k_i^*$. Then, by using the above Lemma 1 and (B.2), the set $T_{Y|Z}^N(\tilde{z}_{i-1})$ maps into at least $\frac{2^{NR}}{1+\delta}$ (here, $\gamma = 2^{NR}$) k_i^* (colors). Thus, in the i-th block, the uncertainty about K_i^* is bounded by:

$$\frac{1}{N} H(K_i^* | \tilde{Z}_{i-1}) \geq R - \frac{\log(1+\delta)}{N}. \tag{B.7}$$

Proof of $\lim_{N\to\infty}\Delta \geq R_e$ for Case 1:

$$\begin{aligned}
\Delta &= \frac{H(W^n|Z^n)}{nN} = \frac{\sum_{i=2}^{n} H(W_i|W^{i-1},Z^n)}{nN} \\
&\overset{(a)}{=} \frac{\sum_{i=2}^{n} H(W_i|\tilde{Z}_i,\tilde{Z}_{i-1})}{nN} \\
&\geq \frac{\sum_{i=2}^{n} H(W_{i1}|\tilde{Z}_i,\tilde{Z}_{i-1})}{nN} \\
&\geq \frac{\sum_{i=2}^{n} H(W_{i1}|\tilde{Z}_i,\tilde{Z}_{i-1},W_{i1}\oplus K_i^*)}{nN} \\
&\geq \frac{\sum_{i=2}^{n} H(W_{i1}|W_{i2},\tilde{Z}_i,\tilde{Z}_{i-1},W_{i1}\oplus K_i^*)}{nN} \\
&\overset{(b)}{=} \frac{\sum_{i=2}^{n} H(W_{i1}|W_{i2},\tilde{Z}_{i-1},W_{i1}\oplus K_i^*)}{nN} \\
&\overset{(c)}{=} \frac{\sum_{i=2}^{n} H(W_{i1}|\tilde{Z}_{i-1},W_{i1}\oplus K_i^*)}{nN} \\
&= \frac{\sum_{i=2}^{n} H(K_i^*|\tilde{Z}_{i-1},W_{i1}\oplus K_i^*)}{nN} \\
&\overset{(d)}{=} \frac{\sum_{i=2}^{n} H(K_i^*|\tilde{Z}_{i-1})}{nN} \\
&\overset{(e)}{\geq} \frac{\sum_{i=2}^{n}(NH(Y|Z)-\log(1+\delta))}{nN} \\
&= \frac{(n-1)(NH(Y|Z)-\log(1+\delta))}{nN},
\end{aligned} \tag{B.8}$$

where (a) is from $W_i \to (\tilde{Z}_i,\tilde{Z}_{i-1}) \to (W^{i-1},\tilde{Z}^{i-2},\tilde{Z}_{i+1}^n)$ (proven in the remainder of this section), (b) is from $W_{i1} \to (W_{i2}, W_{i1}\oplus K_i^*, \tilde{Z}_{i-1}) \to \tilde{Z}_i$ (proven in the remainder of this section), (c) is from W_{i2} being independent of $\tilde{Z}_{i-1}$, $W_{i1}\oplus K_i^*$ and W_{i1}, (d) follows from the fact that $W_{i1}\oplus K_i^*$ is independent of K_i^*, W_{i1} and $\tilde{Z}_{i-1}$ and (e) is from (B.6).

Letting $N\to\infty$ and $n\to\infty$, it is easy to see that:

$$\lim_{N\to\infty}\Delta = \lim_{N\to\infty}\lim_{n\to\infty}\frac{H(W^n|Z^n)}{nN} \geq H(Y|Z) = R_e. \tag{B.9}$$

The proof of $\lim_{N\to\infty}\Delta \geq R_e$ for Case 1 is completed.

Proof of $\lim_{N\to\infty}\Delta \geq R_e$ for Case 2:

$$\begin{aligned}
\Delta &= \frac{H(W^n|Z^n)}{nN} \\
&\overset{(a)}{=} \frac{\sum_{i=2}^{n} H(W_i|\tilde{Z}_i,\tilde{Z}_{i-1})}{nN} \\
&\geq \frac{\sum_{i=2}^{n} H(W_i|\tilde{Z}_i,\tilde{Z}_{i-1},W_i\oplus K_i^*)}{nN} \\
&\overset{(b)}{=} \frac{\sum_{i=2}^{n} H(W_i|\tilde{Z}_{i-1},W_i\oplus K_i^*)}{nN} \\
&= \frac{\sum_{i=2}^{n} H(K_i^*|\tilde{Z}_{i-1},W_i\oplus K_i^*)}{nN} \\
&\overset{(c)}{=} \frac{\sum_{i=2}^{n} H(K_i^*|\tilde{Z}_{i-1})}{nN} \\
&\overset{(d)}{\geq} \frac{\sum_{i=2}^{n}(NR-\log(1+\delta))}{nN} \\
&= \frac{(n-1)(NR-\log(1+\delta))}{nN},
\end{aligned} \tag{B.10}$$

where (a) is from $W_i \to (\tilde{Z}_i, \tilde{Z}_{i-1}) \to (W^{i-1}, \tilde{Z}^{i-2}, \tilde{Z}_{i+1}^n)$ (proven in the remainder of this section), (b) is from $W_i \to (W_i \oplus K_i^*, \tilde{Z}_{i-1}) \to \tilde{Z}_i$ (proven in the remainder of this section), (c) follows from the fact that $W_i \oplus K_i^*$ is independent of K_i^* and $\tilde{Z}_{i-1}$ and (d) is from (B.7).

Letting $N \to \infty$ and $n \to \infty$, it is easy to see that:

$$\lim_{N\to\infty} \Delta = \lim_{N\to\infty} \lim_{n\to\infty} \frac{H(W^n|Z^n)}{nN} \geq R = R_e. \tag{B.11}$$

The proof of $\lim_{N\to\infty} \Delta \geq R_e$ for Case 2 is completed.

It remains to prove the Markov chains $W_i \to (\tilde{Z}_i, \tilde{Z}_{i-1}) \to (W^{i-1}, \tilde{Z}^{i-2}, \tilde{Z}_{i+1}^n)$ and $W_{i1} \to (W_{i2}, W_{i1} \oplus K_i^*, \tilde{Z}_{i-1}) \to \tilde{Z}_i$ of the proof of $\lim_{N\to\infty} \Delta \geq R_e$ for Case 1 and $W_i \to (\tilde{Z}_i, \tilde{Z}_{i-1}) \to (W^{i-1}, \tilde{Z}^{i-2}, \tilde{Z}_{i+1}^n)$, $W_i \to (W_i \oplus K_i^*, \tilde{Z}_{i-1}) \to \tilde{Z}_i$ of the proof of $\lim_{N\to\infty} \Delta \geq R_e$ for Case 2.

Proof. Proof of $W_i \to (\tilde{Z}_i, \tilde{Z}_{i-1}) \to (W^{i-1}, \tilde{Z}^{i-2}, \tilde{Z}_{i+1}^n)$ for Case 1:

For convenience, we denote the probability $Pr\{V = v\}$ by $Pr\{v\}$.

By definition, $W_i \to (\tilde{Z}_i, \tilde{Z}_{i-1}) \to (W^{i-1}, \tilde{Z}^{i-2}, \tilde{Z}_{i+1}^n)$ holds if and only if:

$$Pr\{w_i|\tilde{z}_i, \tilde{z}_{i-1}, w^{i-1}, \tilde{z}^{i-2}, \tilde{z}_{i+1}^n\} = Pr\{w_i|\tilde{z}_i, \tilde{z}_{i-1}\}. \tag{B.12}$$

Equation (B.12) can be further expressed as:

$$\frac{Pr\{w_i, \tilde{z}_i, \tilde{z}_{i-1}, w^{i-1}, \tilde{z}^{i-2}, \tilde{z}_{i+1}^n\}}{Pr\{\tilde{z}_i, \tilde{z}_{i-1}, w^{i-1}, \tilde{z}^{i-2}, \tilde{z}_{i+1}^n\}} = \frac{Pr\{w_i, \tilde{z}_i, \tilde{z}_{i-1}\}}{Pr\{\tilde{z}_i, \tilde{z}_{i-1}\}}. \tag{B.13}$$

It remains to calculate the joint probabilities in (B.13); see the following.

$$\begin{aligned}
& Pr\{w_i, \tilde{z}_i, \tilde{z}_{i-1}, w^{i-1}, \tilde{z}^{i-2}, \tilde{z}_{i+1}^n\} = Pr\{w^i, \tilde{z}^n\} \\
= & \sum_{\tilde{v}^n}\sum_{\tilde{y}^n}\sum_{\tilde{k}^n} Pr\{w^i, \tilde{z}^n, \tilde{v}^n, \tilde{y}^n, \tilde{k}^n\} \\
\overset{(a)}{=} & \sum_{\tilde{v}^n}\sum_{\tilde{y}^n}\sum_{\tilde{k}^n} Pr\{w^i, \tilde{z}^i, \tilde{v}^i, \tilde{y}^i, \tilde{k}^i\} \cdot Pr\{\tilde{z}_{i+1}^n, \tilde{v}_{i+1}^n, \tilde{y}_{i+1}^n, \tilde{k}_{i+1}^n\} \\
= & \sum_{\tilde{v}^i}\sum_{\tilde{y}^i}\sum_{\tilde{k}^i} Pr\{w^i, \tilde{z}^i, \tilde{v}^i, \tilde{y}^i, \tilde{k}^i\} \sum_{\tilde{v}_{i+1}^n}\sum_{\tilde{y}_{i+1}^n}\sum_{\tilde{k}_{i+1}^n} Pr\{\tilde{z}_{i+1}^n, \tilde{v}_{i+1}^n, \tilde{y}_{i+1}^n, \tilde{k}_{i+1}^n\} \\
= & \sum_{\tilde{v}^i}\sum_{\tilde{y}^i}\sum_{\tilde{k}^i} Pr\{w^i, \tilde{z}^i, \tilde{v}^i, \tilde{y}^i, \tilde{k}^i\} Pr\{\tilde{z}_{i+1}^n\} \\
\overset{(b)}{=} & Pr\{\tilde{z}_{i+1}\} \cdots Pr\{\tilde{z}_n\} \Big(\sum_{\tilde{v}^i}\sum_{\tilde{y}^i}\sum_{\tilde{k}^i} Pr\{w^i, \tilde{z}^i, \tilde{v}^i, \tilde{y}^i, \tilde{k}^i\}\Big) \\
= & Pr\{\tilde{z}_{i+1}\} \cdots Pr\{\tilde{z}_n\} \Big(\sum_{\tilde{v}^i}\sum_{\tilde{y}^i}\sum_{\tilde{k}^i} Pr\{\tilde{k}^i|w^i, \tilde{z}^i, \tilde{v}^i, \tilde{y}^i\} Pr\{w^i, \tilde{z}^i, \tilde{v}^i, \tilde{y}^i\}\Big) \\
\overset{(c)}{=} & Pr\{\tilde{z}_{i+1}\} \cdots Pr\{\tilde{z}_n\} \Big(\sum_{\tilde{v}^i}\sum_{\tilde{y}^i} Pr\{w^i, \tilde{z}^i, \tilde{v}^i, \tilde{y}^i\}\Big) \\
\overset{(d)}{=} & Pr\{\tilde{z}_{i+1}\} \cdots Pr\{\tilde{z}_n\} \Big(\sum_{\tilde{v}^i}\sum_{\tilde{y}^i} Pr\{\tilde{z}_1, \tilde{v}_1, \tilde{y}_1\} \prod_{j=2}^{i} Pr\{w_j, \tilde{z}_j, \tilde{v}_j, \tilde{y}_j\}\Big) \\
= & Pr\{\tilde{z}_{i+1}\} \cdots Pr\{\tilde{z}_n\} \Big(\sum_{\tilde{v}_1}\sum_{\tilde{y}_1} Pr\{\tilde{z}_1, \tilde{v}_1, \tilde{y}_1\}\Big)\Big(\sum_{\tilde{v}_2}\sum_{\tilde{y}_2} Pr\{w_2, \tilde{z}_2, \tilde{v}_2, \tilde{y}_2\}\Big) \cdots \Big(\sum_{\tilde{v}_i}\sum_{\tilde{y}_i} Pr\{w_i, \tilde{z}_i, \tilde{v}_i, \tilde{y}_i\}\Big) \\
= & Pr\{\tilde{z}_{i+1}\} \cdots Pr\{\tilde{z}_n\} Pr\{\tilde{z}_1\} Pr\{w_2, \tilde{z}_2\} \cdots Pr\{w_i, \tilde{z}_i\},
\end{aligned} \tag{B.14}$$

where (a) is from the fact that $(\tilde{z}_{i+1}^n, \tilde{v}_{i+1}^n, \tilde{y}_{i+1}^n, \tilde{k}_{i+1}^n)$ are independent of $(w^i, \tilde{z}^i, \tilde{v}^i, \tilde{y}^i, \tilde{k}^i)$, (b) is from the fact that $\tilde{Z}_j$ is independent of $\tilde{Z}_l$ for all of the $i+1 \leq j, l \leq n$ and $j \neq l$, (c) is from the fact that given w^i,

$\tilde{z}^i$, $\tilde{v}^i$ and $\tilde{y}^i$, $\tilde{k}^i$ is uniquely determined, and (d) follows from the fact that $(\tilde{z}_1, \tilde{v}_1, \tilde{y}_1)$, $(w_2, \tilde{z}_2, \tilde{v}_2, \tilde{y}_2)$, ..., $(w_i, \tilde{z}_i, \tilde{v}_i, \tilde{y}_i)$ are independent.

Replacing i by $i-1$, the joint probability $Pr\{\tilde{z}_i, \tilde{z}_{i-1}, w^{i-1}, \tilde{z}^{i-2}, \tilde{z}_{i+1}^n\}$ can be calculated by:

$$\begin{aligned} & Pr\{\tilde{z}_i, \tilde{z}_{i-1}, w^{i-1}, \tilde{z}^{i-2}, \tilde{z}_{i+1}^n\} = Pr\{w^{i-1}, \tilde{z}^n\} \\ \overset{(e)}{=} \; & Pr\{\tilde{z}_i\} \cdot \cdots \cdot Pr\{\tilde{z}_n\} Pr\{\tilde{z}_1\} Pr\{w_2, \tilde{z}_2\} \cdots Pr\{w_{i-1}, \tilde{z}_{i-1}\}, \end{aligned} \tag{B.15}$$

where (e) follows from (B.14) (replacing i by $i-1$).

Substituting (B.14) and (B.15) into the left-hand side of (B.13), we have:

$$\begin{aligned} & \frac{Pr\{w_i, \tilde{z}_i, \tilde{z}_{i-1}, w^{i-1}, \tilde{z}^{i-2}, \tilde{z}_{i+1}^n\}}{Pr\{\tilde{z}_i, \tilde{z}_{i-1}, w^{i-1}, \tilde{z}^{i-2}, \tilde{z}_{i+1}^n\}} \\ = \; & \frac{Pr\{w_i, \tilde{z}_i\}}{Pr\{\tilde{z}_i\}}. \end{aligned} \tag{B.16}$$

Next, we need to calculate the right-hand side of (B.13); see the following.

$$Pr\{w_i, \tilde{z}_i, \tilde{z}_{i-1}\} \overset{(1)}{=} Pr\{w_i, \tilde{z}_i\} Pr\{\tilde{z}_{i-1}\}, \tag{B.17}$$

where (1) is from the fact that W_i and $\tilde{Z}_i$ are independent of $\tilde{Z}_{i-1}$.

The joint probability $Pr\{\tilde{z}_i, \tilde{z}_{i-1}\}$ is calculated by:

$$Pr\{\tilde{z}_i, \tilde{z}_{i-1}\} \overset{(2)}{=} Pr\{\tilde{z}_i\} \cdot Pr\{\tilde{z}_{i-1}\}, \tag{B.18}$$

where (1) is from the fact that $\widetilde{Z}_i$ is independent of $\widetilde{Z}_{i-1}$.

Substituting (B.17) and (B.18) into the right-hand side of (B.13), we have:

$$\frac{Pr\{w_i, \tilde{z}_i, \tilde{z}_{i-1}\}}{Pr\{\tilde{z}_i, \tilde{z}_{i-1}\}} = \frac{Pr\{w_i, \tilde{z}_i\}}{Pr\{\tilde{z}_i\}}. \tag{B.19}$$

By checking (B.16) and (B.19), the Markov chain $W_i \rightarrow (\tilde{Z}_i, \tilde{Z}_{i-1}) \rightarrow (W^{i-1}, \tilde{Z}^{i-2}, \tilde{Z}_{i+1}^n)$ is proven. □

Proof. Proof of $W_{i1} \rightarrow (W_{i2}, W_{i1} \oplus K_i^*, \tilde{Z}_{i-1}) \rightarrow \tilde{Z}_i$ for Case 1:

By definition, $W_{i1} \rightarrow (W_{i2}, W_{i1} \oplus K_i^*, \tilde{Z}_{i-1}) \rightarrow \tilde{Z}_i$ holds if and only if:

$$Pr\{w_{i1} | w_{i2}, w_{i1} \oplus k_i^*, \tilde{z}_{i-1}, \tilde{z}_i\} = Pr\{w_{i1} | w_{i2}, w_{i1} \oplus k_i^*, \tilde{z}_{i-1}\}. \tag{B.20}$$

Equation (B.20) can be further expressed as:

$$\frac{Pr\{w_{i1}, w_{i2}, w_{i1} \oplus k_i^*, \tilde{z}_{i-1}, \tilde{z}_i\}}{Pr\{w_{i2}, w_{i1} \oplus k_i^*, \tilde{z}_{i-1}, \tilde{z}_i\}} = \frac{Pr\{w_{i1}, w_{i2}, w_{i1} \oplus k_i^*, \tilde{z}_{i-1}\}}{Pr\{w_{i2}, w_{i1} \oplus k_i^*, \tilde{z}_{i-1}\}}. \tag{B.21}$$

It remains to calculate the joint probabilities in (B.21); see the following.

$$Pr\{w_{i1}, w_{i2}, w_{i1} \oplus k_i^*, \tilde{z}_{i-1}, \tilde{z}_i\} \overset{(a)}{=} Pr\{w_{i1}\} \cdot Pr\{w_{i2}, w_{i1} \oplus k_i^*, \tilde{z}_i\} \cdot Pr\{\tilde{z}_{i-1}\}, \tag{B.22}$$

where (a) is from the fact that W_{i1} is independent of W_{i2}, $W_{i1} \oplus K_i^*$, $\tilde{Z}_i$ and $\tilde{Z}_{i-1}$ and $\tilde{Z}_{i-1}$ is independent of W_{i2}, $W_{i1} \oplus K_i^*$, $\tilde{Z}_i$.

Similarly, we have:

$$Pr\{w_{i2}, w_{i1} \oplus k_i^*, \tilde{z}_{i-1}, \tilde{z}_i\} \overset{(b)}{=} Pr\{w_{i2}, w_{i1} \oplus k_i^*, \tilde{z}_i\} \cdot Pr\{\tilde{z}_{i-1}\}, \tag{B.23}$$

where (b) is from the fact that $\tilde{Z}_{i-1}$ is independent of W_{i2}, $W_{i1} \oplus K_i^*$ and $\tilde{Z}_i$.

Substituting (B.22) and (B.23) into the left-hand side of (B.21), we have:

$$\frac{Pr\{w_{i1}, w_{i2}, w_{i1} \oplus k_i^*, \tilde{z}_{i-1}, \tilde{z}_i\}}{Pr\{w_{i2}, w_{i1} \oplus k_i^*, \tilde{z}_{i-1}, \tilde{z}_i\}} = Pr\{w_{i1}\}. \tag{B.24}$$

Next, we need to calculate the right-hand side of (B.21); see the following.

$$Pr\{w_{i1}, w_{i2}, w_{i1} \oplus k_i^*, \tilde{z}_{i-1}\} \overset{(c)}{=} Pr\{w_{i1}\} \cdot Pr\{w_{i2}\} \cdot Pr\{w_{i1} \oplus k_i^*\} \cdot Pr\{\tilde{z}_{i-1}\}, \tag{B.25}$$

where (c) is from the fact that W_{i1}, W_{i2}, $W_{i1} \oplus K_i^*$ and $\tilde{Z}_{i-1}$ are independent.

The joint probability $Pr\{w_{i2}, w_{i1} \oplus k_i^*, \tilde{z}_{i-1}\}$ is calculated by:

$$Pr\{w_{i2}, w_{i1} \oplus k_i^*, \tilde{z}_{i-1}\} \overset{(d)}{=} Pr\{w_{i2}\} \cdot Pr\{w_{i1} \oplus k_i^*\} \cdot Pr\{\tilde{z}_{i-1}\}, \tag{B.26}$$

where (d) is from the fact that W_{i2}, $W_{i1} \oplus K_i^*$ and $\tilde{Z}_{i-1}$ are independent.

Substituting (B.25) and (B.26) into the right-hand side of (B.21), we have:

$$\frac{Pr\{w_{i1}, w_{i2}, w_{i1} \oplus k_i^*, \tilde{z}_{i-1}\}}{Pr\{w_{i2}, w_{i1} \oplus k_i^*, \tilde{z}_{i-1}\}} = Pr\{w_{i1}\}. \tag{B.27}$$

By checking (B.24) and (B.27), the Markov chain $W_{i1} \to (W_{i2}, W_{i1} \oplus K_i^*, \tilde{Z}_{i-1}) \to \tilde{Z}_i$ is proven. □

Proof. Proof of $W_i \to (\tilde{Z}_i, \tilde{Z}_{i-1}) \to (W^{i-1}, \tilde{Z}^{i-2}, \tilde{Z}_{i+1}^n)$ for Case 2:

Letting $W_{i2} = \varnothing$ and $W_{i1} = W_i$ for all $2 \leq i \leq n$, the proof of $W_i \to (\tilde{Z}_i, \tilde{Z}_{i-1}) \to (W^{i-1}, \tilde{Z}^{i-2}, \tilde{Z}_{i+1}^n)$ for Case 2 is along the lines of that for Case 1, and therefore, we omit it here. □

Proof. Proof of $W_i \to (W_i \oplus K_i^*, \tilde{Z}_{i-1}) \to \tilde{Z}_i$ for Case 2:

Letting $W_{i2} = \varnothing$ and $W_{i1} = W_i$ for all $2 \leq i \leq n$, the proof of $W_i \to (W_i \oplus K_i^*, \tilde{Z}_{i-1}) \to \tilde{Z}_i$ for Case 2 is along the lines of that for Case 1, and therefore, we omit it here. □

Thus, the direct proof of Theorem 1 is completed.

References

1. Ahlswede, R.; Cai, N. Transmission, Identification and Common Randomness Capacities for Wire-Tap Channels with Secure Feedback from the Decoder. In *General Theory of Information Transfer and Combinatorics*; Springer-Verlag: Berlin/Heidelberg, Germany, 2006; Volume 4123, pp. 258–275.
2. Dai, B.; Vinck, A.J.H.; Luo, Y.; Zhuang, Z. Capacity region of non-degraded wiretap channel with noiseless feedback. In Proceedings of 2012 IEEE International Symposium on Information Theory, Cambridge, MA, USA, 1–6 July 2012.
3. Wyner, A.D. The wire-tap channel. *Bell Syst. Tech. J.* **1975**, *54*, 1355–1387.
4. Ardestanizadeh, E.; Franceschetti, M.; Javidi, T.; Kim, Y. Wiretap channel with secure rate-limited feedback. *IEEE Trans. Inf. Theory* **2009**, *55*, 5353–5361.
5. Lai, L.; El Gamal, H.; Poor, H.V. The wiretap channel with feedback: Encryption over the channel. *IEEE Trans. Inf. Theory* **2008**, *54*, 5059–5067.
6. He, X.; Yener, A. The role of feedback in two-way secure communication. *IEEE Trans. Inf. Theory* **2013**, *59*, 8115–8130.
7. Bassi, G.; Piantanida, P.; Shamai, S. On the capacity of the wiretap channel with generalized feedback. In Processdings of 2015 IEEE International Symposium on Information Theory, Hong Kong, China, 14–19 June 2015.
8. Mitrpant, C.; Vinck, A.J.H.; Luo, Y. An achievable region for the gaussian wiretap channel with side information. *IEEE Trans. Inf. Theory* **2006**, *52*, 2181–2190.

9. El-Halabi, M.; Liu, T.; Georghiades, C.N.; Shamai, S. Secret writing on dirty paper: A deterministic view. *IEEE Trans. Inf. Theory* **2012**, *58*, 3419–3429.
10. Chen, Y.; Vinck, A.J.H. Wiretap channel with side information. *IEEE Trans. Inf. Theory* **2008**, *54*, 395–402.
11. Gel'fand, S.I.; Pinsker, M.S. Coding for channel with random parameters. *Probl. Control Inf. Theory* **1980**, *9*, 19–31.
12. Dai, B.; Luo, Y. Some new results on wiretap channel with side information. *Entropy* **2012**, *14*, 1671–1702.
13. Chia, Y.K.; El Gamal, A. Wiretap channel with causal state information. *IEEE Trans. Inf. Theory* **2012**, *58*, 2838–2849.
14. Liu, T.; Mukherjee, P.; Ulukus, S.; Lin, S.; Hong, Y.W.P. Secure degrees of freedom of MIMO Rayleigh block fading wiretap channels with no CSI anywhere. *IEEE Trans. Wireless Commun.* **2015**, *14*, 2655–2669.
15. Csiszár, I.; Körner, J. *Information Theory: Coding Theorems for Discrete Memoryless Systems*; Academic Press: New York, NY, USA, 1981; pp. 123–124.
16. Csiszár, I.; Körner, J. Broadcast channels with confidential messages. *IEEE Trans. Inf. Theory* **1978**, *24*, 339–348.
17. Gamal, A.E.; Kim, Y.H. *Network Information Theory*; Cambridge University Press: Cambridge, UK, 2011.
18. Costa, M.H.M. Writing on dirty paper. *IEEE Trans. Inf. Theory* **1983**, *29*, 439–441.
19. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423, 623–656.
20. Leung-Yan-Cheong, S.; Hellman, M.E. The Gaussian wire-tap channel. *IEEE Trans. Inf. Theory* **1978**, *24*, 451–456.