The Relay–Eavesdropper Channel: Cooperation for Secrecy

Lifeng Lai, *Member, IEEE*, and Hesham El Gamal, *Senior Member, IEEE*

***Abstract*—This paper establishes the utility of user cooperation in facilitating secure wireless communications. In particular, the four-terminal relay–eavesdropper channel is introduced and an outer-bound on the optimal rate-equivocation region is derived. Several cooperation strategies are then devised and the corresponding achievable rate-equivocation region are characterized. Of particular interest is the novel noise-forwarding (NF) strategy, where the relay node sends codewords independent of the source message to confuse the eavesdropper. This strategy is used to illustrate the deaf helper phenomenon, where the relay is able to facilitate secure communications while being totally ignorant of the transmitted messages. Furthermore, NF is shown to increase the secrecy capacity in the reversely degraded scenario, where the relay node fails to offer performance gains in the classical setting. The gain offered by the proposed cooperation strategies is then proved theoretically and validated numerically in the additive white Gaussian noise (AWGN) channel.**

***Index Terms*—Cooperation, eavesdropper, noise-forwarding (NF), relay, security.**

I. INTRODUCTION

SHANNON introduced the notion of information theoretic secrecy in [1]. The model in [1] assumed that the transmission is noiseless, and used a key K to protect the confidential message W. Taking the transmission uncertainty into consideration, Wyner introduced the wiretap channel in [2]. In the three-terminal wiretap channel, a source wishes to transmit confidential messages to a destination while keeping the messages as secret as possible from a wiretapper. The wiretapper is assumed to have an unlimited computation ability and to know the coding/decoding scheme used in the main (source-destination) channel. Under the assumption that the source-wiretapper channel is a degraded version of the main channel, Wyner characterized the trade-off between the throughput of the main channel and the level of ignorance of the message at the wiretapper using the rate-equivocation region concept. Loosely speaking, the equivocation rate measures the residual ambiguity about the transmitted message at the wiretapper. If the equivocation rate at the wiretapper is arbitrarily close to the information rate, the transmission is called perfectly secure. Csiszár and Körner extended this work to the broadcast channel with confidential messages, where the source sends common information to both the destination and the wiretapper, and confidential messages are sent only to the destination [3].

Our work here is motivated by the fact that if the wiretapper channel is less noisy than the main channel,[1] the perfect secrecy capacity of the channel is zero [3]. In this case, it is **infeasible** to establish a secure link under Wyner's wiretap channel model. Our main idea is to exploit user cooperation in facilitating the transmission of confidential messages from the source to the destination. More specially, we consider a four-terminal relay–eavesdropper channel, where a source wishes to send messages to a destination while leveraging the help of a relay node to hide those messages from the eavesdropper. The eavesdropper in our model can be viewed as the wireless counterpart of Wyner's wiretapper. This model generalizes the relay channel [4] and the wiretap channel [2].

The relay channel without security constraints was studied under various scenarios [4]–[12]. In most of these works, cooperation strategies were constructed to increase the transmission rate and/or reliability function. In this paper, we identify a **novel** role of the relay node in establishing a secure link from the source to the destination. Toward this end, several cooperation strategies for the relay–eavesdropper channel are constructed and the corresponding achieved rate-equivocation regions are characterized. An outer bound on the optimal rate-equivocation region is also derived. The proposed schemes are shown to achieve a positive perfect secrecy rate in several scenarios where the secrecy capacity in the absence of the relay node is zero. Quite interestingly, we establish the deaf-helper phenomenon where the relay can help while being totally ignorant of the transmitted message from the source. Furthermore, we show that the relay node can aid in the transmission of confidential messages in some settings where classical cooperation fails to offer performance gains, e.g., the reversely degraded relay channel. Finally, we observe that the proposed noise-forwarding (NF) is intimately related with the multiple access channel with security constraints, as evident in the sequel.

At this point, we wish to differentiate our investigation from earlier relevant works. The relay channel with confidential messages was studied in [13], [14], where the relay node acts both as an eavesdropper and a helper. In the model of [14], the source sends common messages to the destination using the help of the

Manuscript received December 6, 2006; revised April 28, 2008. Published August 27, 2008 (projected). The material in this paper was presented in part at the IEEE Symposium on Information Theory, Nice, France, June 2007.

L. Lai is with the Department of Electrical Engineering at Princeton University, Princeton, NJ 08544 USA (e-mail: llai@princeton.edu).

H. El Gamal is with the Department of Electrical and Computer Engineering at the Ohio State University, Columbus, OH 43210 USA. He also serves as the Director for the Wireless Intelligent Networks Center (WINC), Nile University, Cairo, Egypt (e-mail: helgamal@ece.osu.edu).

Communicated by P. Viswanath, Associate Editor for Communications.

Color versions of Figures 3 and 5 of this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2008.928272

[1]The source-wiretapper channel is said to be less noisy than the source-receiver channel, if for every $V \rightarrow X \rightarrow YZ$, $I(V;Z) \geq I(V;Y)$, where X is the signal transmitted by the source, Y, Z are the received signal of the receiver and the wiretapper respectively.

relay node, but also sends private messages to the destination while keeping them secret from the relay. In contrast with [14], the relay node in our work acts as a trusted "third-party" whose sole goal is to facilitate secure communications (imposing an additional security constraint on the relay node is also considered in Section IV). The idea of using a "third-party" to facilitate secure communications also appeared in [15]. Contrary to our work, which considers **noisy** channels, [15] focused on the generation of common random secret keys at two nodes under the assist of a third-party using a **noiseless** public discussion channel. The users then use the secret key to establish a secure link between the source-destination pair. Other recent works on secure communications investigated the multiple-access channel (MAC) with confidential messages [16], [17], the MAC with a degraded wiretapper [18], and multiple-input–multiple-output (MIMO) secure communications [19]. In summary, it appears that our relay–eavesdropper model is fundamentally different from the models considered in all previous works.

Throughout the paper, upper case letter X denotes a random variable, lower case letter x denotes a realization of the random variable, calligraphic letter $\mathcal{X}$ denotes a finite alphabet set. Boldface letter $\mathbf{x}$ denotes a vector, $\{\cdot\}^T$ denotes transpose and $\{\cdot\}^H$ denotes conjugate transpose. We also let $[x]^+ = \max\{0, x\}$.

The rest of the paper is organized as follows. In Section II, we introduce the system model and our notation. Section III describes the proposed cooperation strategies and characterizes the corresponding achievable performance. The rate-equivocation outer-bound is also developed in this section. In Section IV, we discuss several examples that illustrate interesting aspects of the relay–eavesdropper channel. Finally, Section V offers some concluding remarks and briefly outlines possible venues for future research.

II. The Relay-Eavesdropper Channel

We consider a four-terminal discrete channel consisting of finite sets $\mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}, \mathcal{Y}_1, \mathcal{Y}_2$ and a transition probability distribution $p(y, y_1, y_2 \mid x_1, x_2)$, as shown in Fig. 1. Here, $\mathcal{X}_1, \mathcal{X}_2$ are the channel inputs from the source and the relay respectively, while $\mathcal{Y}, \mathcal{Y}_1, \mathcal{Y}_2$ are the channel outputs at the destination, relay and eavesdropper respectively. We impose the memoryless assumption, i.e., the channel outputs $(y_i, y_{1,i}, y_{2,i})$ at time i only depend on the channel inputs $(x_{1,i}, x_{2,i})$ at time i. The source wishes to send the message $W_1 \in \mathcal{W}_1 = \{1, \ldots, M\}$ to the destination using the (M, n) code consisting: 1) a stochastic encoder f_n at the source that maps the message w_1 to a codeword $\mathbf{x}_1 \in \mathcal{X}_1^n$, 2) a relay encoder that maps the signals $(y_{1,1}, y_{1,2}, \ldots, y_{1,i-1})$ received before time i to the channel input $x_{2,i}$, using the mapping φ_i: $(Y_{1,1}, Y_{1,2}, \cdots, Y_{1,i-1}) \to X_{2,i}$, 3) a decoding function ϕ: $\mathcal{Y}^n \to \mathcal{W}_1$. The average error probability of a (M, n) code is defined as

$$P_e^n = \sum_{w_1 \in \mathcal{W}_1} \frac{1}{M} \Pr\{\phi(\mathbf{y}) \neq w_1 | w_1 \text{ was sent}\}. \tag{1}$$

The equivocation rate at the eavesdropper is defined as

$$R_e = \frac{1}{n} H(W_1 | \mathbf{Y}_2). \tag{2}$$

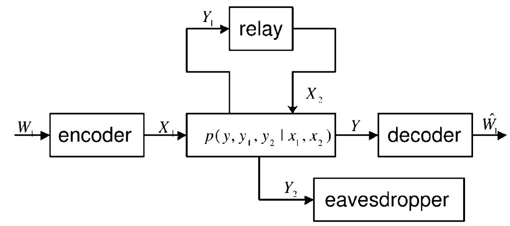

Fig. 1. The relay eavesdropper channel.

The rate-equivocation pair (R_1, R_e) is said to be achievable if for any $\epsilon > 0$, there exists a sequence of codes (M, n) such that for any $n \geq n(\epsilon)$, we have

$$R_1 = \frac{1}{n} \log_2 M \tag{3}$$

$$P_e^n \leq \epsilon \tag{4}$$

$$\frac{1}{n} H(W_1 | \mathbf{Y}_2) \geq R_e - \epsilon. \tag{5}$$

We further say that the perfect secrecy rate R_1 is achievable if the rate-equivocation pair (R_1, R_1) is achievable. Notice that if $\mathcal{Y}_2 = \Phi$, our model reduces to the classical relay channel without security constraints.

III. Main Results

Our first result establishes an outer bound on the optimal rate-equivocation region of the relay–eavesdropper channel.

Theorem 1: In the relay–eavesdropper channel, for any rate-equivocation pair $\{R_1, R_e\}$ with $P_e^n \to 0$ and the equivocation rate at the eavesdropper larger than $R_e - \epsilon$, there exist some random variables $U \to (V_1, V_2) \to (X_1, X_2) \to (Y, Y_1, Y_2)$, such that (R_1, R_e) satisfies the following conditions:

$$\begin{aligned} &R_1 \leq \min\{I(V_1, V_2; Y), I(V_1; Y, Y_1 | V_2)\} \\ &R_e \leq R_1 \\ &R_e \leq \min\{I(V_1, V_2; Y | U), I(V_1; Y, Y_1 | V_2, U)\} \\ &\quad - I(V_1, V_2; Y_2 | U). \end{aligned} \tag{6}$$

Proof: Please refer to Appendix A. □

We now turn our attention to constructing cooperation strategies for the relay–eavesdropper channel. Our first step is to characterize the achievable rate-equivocation region of Cover–El Gamal Decode and Forward (DF) Strategy [4]. In DF cooperation strategy, the relay node will first decode codewords and then re-encode the message to cooperate with the source. Here, we use the regular coding and backward decoding scheme developed in the classical relay setting [20], [7], with the important difference that each message will be associated with many codewords in order to confuse the eavesdropper.

Theorem 2: The rate pairs in the closure of the convex hull of all (R_1, R_e) satisfying

$$\begin{aligned} &R_1 < \min\{I(V_1, V_2; Y | U), I(V_1; Y_1 | V_2, U)\} \\ &R_e < [\min\{I(V_1, V_2; Y | U), I(V_1; Y_1 | V_2, U)\} \\ &\quad - I(V_1, V_2; Y_2 | U)]^+ \end{aligned} \tag{7}$$

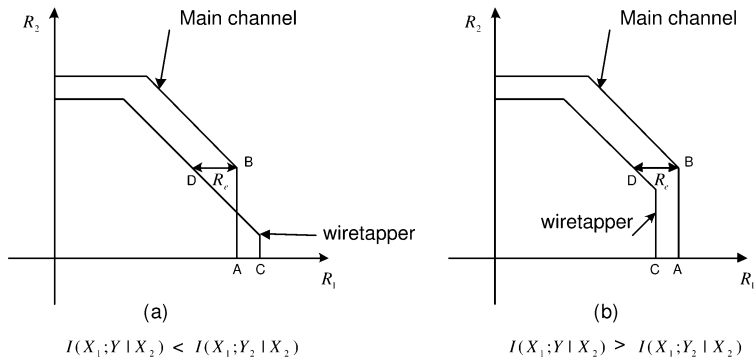

Fig. 2. The rate region of the compound MACs of the relay eavesdropper channel for a fixed input distribution $p(x_1)p(x_2)$.

for some distribution $p(u, v_1, v_2, x_1, x_2, y_1, y_2, y) = p(u)p(v_1, v_2|u)p(x_1, x_2|v_1, v_2)p(y_1, y_2, y|x_1, x_2)$, are achievable using the DF strategy.

Hence, for the DF scheme, the following perfect secrecy rate is achievable

$$R_s^{(DF)} = \sup_{p(u)p(v_1,v_2|u)p(x_1,x_2|v_1,v_2)} [\min\{I(V_1, V_2; Y|U), I(V_1; Y_1|V_2, U)\} - I(V_1, V_2; Y_2|U)]^+. \quad (8)$$

Proof: Please refer to Appendix B.

The channel between the source and the relay becomes a **bottleneck** for the DF strategy when it is noisier than the source–destination channel. This motivates our noise-forwarding (NF) scheme, where the relay node does not attempt to decode the message but sends codewords that are *independent* of the source's message. The enabling observation behind this scheme is that, in the wiretap channel, in addition to its own information, the source should send extra codewords to confuse the wiretapper. In our setting, this task can be accomplished by the relay by allowing it to send independent codewords, which aid in confusing the eavesdropper.

Our NF scheme transforms the relay–eavesdropper channel into a compound MAC, where the source/relay to the receiver is the first MAC and source/relay to the eavesdropper is the second one. Fig. 2 shows the rate region of these two MACs for a fixed input distribution $p(x_1)p(x_2)$. In the figure, R_1 is the codeword rate of the source, and R_2 is the codeword rate of the relay. We can observe from Fig. 2(a) that if the relay node does not transmit, the perfect secrecy rate is zero for this input distribution since $R_1(A) < R_1(C)$. On the other hand, if the relay and the source coordinate their transmissions and operate at point B, we can achieve the equivocation rate R_e, which is strictly larger than zero. On the other hand, in Fig. 2(b), we can still get a positive perfect secrecy rate by operating at point A in the absence of the relay. But by moving the operating point to B, we can get a larger secrecy rate. This illustrates the main idea of our NF scheme. The next result establishes the achievable rate-equivocation region for the NF scheme.

Theorem 3: The rate pairs in the closure of the convex hull of all (R_1, R_e) satisfying

$$\begin{aligned} R_1 &< I(V_1; Y|V_2) \\ R_e &< R_1 \\ R_e &< [I(V_1; Y|V_2) + \min\{I(V_2; Y), I(V_2; Y_2|V_1)\} \\ &\quad - \min\{I(V_2; Y), I(V_2; Y_2)\} \\ &\quad - I(V_1; Y_2|V_2)]^+ \end{aligned} \quad (9)$$

for some distribution

$$p(v_1, v_2, x_1, x_2, y_1, y_2, y) = p(v_1)p(v_2)p(x_1|v_1)p(x_2|v_2)p(y_1, y_2, y|x_1, x_2)$$

are achievable using the NF scheme.

Hence, for the NF scheme, the following perfect secrecy rate is achievable:

$$R_s^{(\mathrm{NF})} = \sup_{p(v_1)p(v_2)p(x_1|v_1)p(x_2|v_2)} [I(V_1; Y|V_2) + \min\{I(V_2; Y), I(V_2; Y_2|V_1)\} - \min\{I(V_2; Y), I(V_2; Y_2)\} - I(V_1; Y_2|V_2)]^+ . \quad (10)$$

□

Proof: Please refer to Appendix C.

The following comments are now in order.

1) The NF scheme is customized to the relay channel with security constraints which make the transmission of codewords that are independent of the source message reasonable. Also, in the NF scheme, the relay node does not need to listen to the source, and hence, this scheme works for relay nodes limited by the half-duplex constraint [11], [9], [21].
2) In NF cooperation, each user sends independent messages to the destination, which resembles the MAC. Hence, NF cooperation can be adapted to the multiple access eavesdropper channel where the multiple users in the MAC channel can help each other in communicating securely with the destination without listening to each other (note that the results in [18] were limited only to the case where the eavesdropped channel is a degraded version of the channel seen by the destination). Our related results will be reported elsewhere.

Now, we study another cooperation scheme that does not require decoding at the relay: Compress and Forward (CF). The CF cooperation strategy can be viewed as a generalization of NF where, in addition to the independent codewords, the relay also sends a quantized version of its noisy observations to the destination. This noisy version of the relay's observations helps the destination in decoding the source's message, while the independent codewords help in confusing the eavesdropper. The following result establishes the achievable rate-equivocation pair in the case when $I(X_1;\hat{Y}_1,Y|X_2) \leq I(X_1;\hat{Y}_1,Y_2|X_2)$, i.e., the source–eavesdropper channel is better than the source–receiver channel, a situation of particular interest to us.

Theorem 4: The rate pairs in the closure of the convex hull of all (R_1,R_e) satisfying

$$\begin{aligned} &R_1 < I(X_1;\hat{Y}_1,Y|X_2),\\ &R_e < R_1\\ &R_e < \left[R_0 + I(X_1;\hat{Y}_1,Y|X_2) - I(X_1,X_2;Y_2)\right]^+ \end{aligned} \tag{11}$$

subject to

$$\min\{I(X_2;Y),I(X_2;Y_2|X_1)\} - R_0 \geq I(Y_1;\hat{Y}_1|X_2) \tag{12}$$

for some distribution $p(x_1,x_2,y_1,y_2,y,\hat{y}_1) = p(x_1)p(x_2)p(y_1,y_2,y|x_1,x_2)p(\hat{y}_1|y_1,x_2)$, are achievable using CF strategy.

Proof: Please refer to Appendix D.

Three comments are now in order.

1) In Theorem 4, R_0 is the rate of pure noise generated by the relay to confuse the eavesdropper, while $\min\{I(X_2;Y),I(X_2;Y_2|X_1)\} - R_0$ is the part of the rate allocated to send the compressed signal $\hat{Y}_1$ to help the destination. If we set $R_0 = \min\{I(X_2;Y),I(X_2;Y_2|X_1)\}$, this scheme becomes the NF scheme.
2) In order to enable analytical tractability, the coding/decoding scheme used in the proof is slightly different from that of [4]. In [4], the destination uses sliding-window decoding, while our proof uses backward decoding. Hence, the bound for R_e provided here is a lower bound for the R_e achieved by the CF scheme. One may be able to achieve a larger R_e using exactly the CF scheme proposed in [4]. But, unfortunately, we are not yet able to bound R_e when sliding-window decoding is used.
3) Compared with CF decoding, the proposed NF strategy enjoys the advantage of simplicity. Also, if one only focuses on the perfect secrecy rate, it is easy to see that these two schemes achieve identical performance. Again, this observation is limited to our lower bound on R_e in Theorem 4.

IV. Examples

This section discusses several examples that illustrate some unique features of the relay–eavesdropper channel. For simplicity, we only focus on the perfect secrecy rate of various schemes.

A. Physically Degraded Relay Channel

In the following, we show that the decode-forward strategy achieves the perfect secrecy capacity for a class of physically degraded relay–eavesdropper channel. In this class of physically degraded relay–eavesdropper channel, the output at the destination is a degraded version of the output at the relay node.

Definition 1: The relay–eavesdropper channel is called physically degraded, if

$$p(y,y_1,y_2|x_1,x_2) = p(y_1|x_1,x_2)p(y|y_1)p(y_2|y_1,y,x_1,x_2).$$

Notice that in this definition, we do not put any degradedness assumption at the output of the eavesdropper.

Theorem 5: The perfect secrecy capacity of the physically degraded relay–eavesdropper channel is

$$\begin{aligned} C_s = \sup_{p(u)p(v_1,v_2|u)p(x_1,x_2|v_1,v_2)} &[\min\{I(V_1,V_2;Y|U),\\ &I(V_1;Y_1|V_2,U)\} - I(V_1,V_2;Y_2|U)]^+. \end{aligned}$$

Proof: Please refer to Appendix E. □

B. The Deaf Helper Phenomenon

The security constraints imposed on the network bring about a new phenomenon which we call the *deaf helper phenomenon*, where the relay node can still help even it is totally ignorant of the message transmitted from the source. In this setup, we impose an additional security constraint on the relay node, and say a rate R_s is achievable for a deaf helper if for any $\epsilon > 0$, there exists a sequence of codes (M,n) such that for any $n \geq n(\epsilon)$, we have

$$\begin{aligned} &R_s = \frac{1}{n}\log_2 M\\ &P_e^n \leq \epsilon\\ &\frac{1}{n}H(W_1|\mathbf{Y}_2) \geq R_s - \epsilon\\ &\frac{1}{n}H(W_1|\mathbf{Y}_1,\mathbf{X}_2) \geq R_s - \epsilon. \end{aligned} \tag{13}$$

In this case, the signal received by the relay node does not leak any information about the transmitted message W_1. This model describes a more conservative scenario where the source does not trust the relay but still wishes to exploit the benefit brought by cooperation. We assume that the relay node is not malicious and, hence, is willing to cooperate with the source.[2] The following theorem characterizes the achievable perfect secrecy rate of the NF strategy in the deaf-helper setting.

Theorem 6: The perfect secrecy rate of the NF scheme with an additional security constraint on the relay node is $R_s = \max_{p(v_1)p(v_2)p(x_1|v_1)p(v_2|v_2)} \min\{R_{s1},R_{s2}\}$, where

$$\begin{aligned} R_{s1} = &[I(V_1;Y|V_2) + \min\{I(V_2;Y),I(V_2;Y_2|V_1)\}\\ &- \min\{I(V_2;Y),I(V_2;Y_2)\} - I(V_1;Y_2|V_2)]^+\\ R_{s2} = &[I(V_1;Y|V_2) - I(V_1;Y_1|X_2)]^+. \end{aligned}$$

Proof: Please refer to Appendix F. □

C. The Reversely Degraded Relay-Eavesdropper Channel

In the classical relay channel without security constraints, there exist some scenarios where the relay node does not pro-

[2]If the relay node is malicious, it can then send signals that are dependent with signal received and then could even block the transmission of the main channel.

vide any gain, for example, the reversely degraded relay channel shown in [4]. Here, we focus on this scenario and show that the relay node can still offer a gain in the presence of the eavesdropper.

Definition 2: [[4]] The relay channel is called reversely degraded, if $p(y,y_1|x_1,x_2)=p(y|x_1,x_2)p(y_1|y,x_2)$.

The following result, borrowed from [4], states the capacity of the classical reversely degraded relay channel.

Theorem 7 [4, Th. 2]: The capacity of the reversely degraded relay channel is

$$C_0=\max_{x_2}\max_{p(x_1)} I(X_1;Y|x_2). \tag{14}$$

This result implies that the relay node should send a constant, and hence, does not contribute new information to the destination. In most channel models, the constant sent by the relay does not result in any capacity gain. The question now is whether the same conclusion holds in the presence of an eavesdropper. We first observe that the degradedness of the relay channel implies that DF and CF cooperation will not provide the destination with additional useful information. The relay node, however, can still send codewords independent of the received signal to confuse the eavesdropper, as proposed in the NF scheme. Since we do not require decoding at the relay node in the Proof of Theorem 3, the degradedness imposed here does not affect the performance. Hence, we get the following achievable perfect secrecy rate for the reversely degraded relay–eavesdropper channel.

Corollary 1: The following rate is achievable for the reversely degraded relay–eavesdropper channel:

$$\begin{aligned}R_s=\max_{p(v_1)p(v_2)p(x_1|v_1)p(x_2|v_2)}&\big[I(V_1;Y|V_2)\\&+\min\{I(V_2;Y),I(V_2;Y_2|V_1)\}\\&-\min\{I(V_2;Y),I(V_2;Y_2)\}-I(V_1;Y_2|V_2)\big]^+.\end{aligned} \tag{15}$$

D. The AWGN Channel

Now we consider the Gaussian relay–eavesdropper channel, where the signal received at each node is

$$y_j[n]=\sum_{i\neq j}h_{ij}x_i[n]+z_j[n]$$

here h_{ij} is the channel coefficient between node $i\in\{s,r\}$ and node $j\in\{r,w,d\}$, and z_j is the i.i.d Gaussian noise with unit variance at node j. The source and the relay have average power constraint P_1,P_2, respectively.

In [22], it was shown that the secrecy capacity of the degraded Gaussian wiretap channel is $[C_M-C_{MW}]^+$, where C_M,C_{MW} are the capacity of the main channel and wiretap channel, respectively. This result is also shown to be valid for stochastically degraded channel [17]. In our case, if the relay does not transmit, the relay eavesdropper channel becomes a Gaussian eavesdropper channel, which can always be converted into a stochastically degraded channel as done in the Gaussian broadcast channel [23]. Applying this result to our case, the secrecy capacity of the Gaussian eavesdropper channel without the relay node is given by $[\frac{1}{2}\log_2(1+|h_{sd}|^2P_1)-\frac{1}{2}\log_2(1+|h_{sw}|^2P_1)]^+$. Hence if $|h_{sw}|^2\geq|h_{sd}|^2$ and the relay does not transmit, the secrecy capacity is zero, no matter how large P_1 is. On the other hand, as shown later, the relay can facilitate the source-destination pair to achieve a positive perfect secrecy rate under some conditions even when $|h_{sw}|^2\geq|h_{sd}|^2$. In the following, we focus on such scenarios.

1) DF and NF: At this point, we do not know the optimal input distribution that maximizes $R_s^{(DF)}$, $R_s^{(NF)}$. Here, we let $V_1=X_1, V_2=X_2$, U to be a constant, and use a Gaussian input distribution to obtain an achievable lower bound.

For DF cooperation scheme, we let $X_2\sim\mathcal{N}(0,P_2)$, $X_{10}\sim\mathcal{N}(0,P)$, where $\mathcal{N}(0,P)$ is the Gaussian distribution with zero mean and variance P. Also, we let

$$X_1=c_1X_2+X_{10}$$

where c_1 is a constant to be specified later. In this relationship, the novel information is modeled by X_{10}, whereas X_2 represents the part of the signal which the source and the relay cooperate in beamforming toward the destination. To satisfy the average power constraint at the source, we require $|c_1|^2P_2+P\leq P_1$.

Straightforward calculations result in

$$\begin{aligned}I(X_1;Y_1|X_2)&=\frac{1}{2}\log_2(1+|h_{sr}|^2P),\\I(X_1,X_2;Y)&=\frac{1}{2}\log_2(1+|h_{sd}c_1\\&\quad+h_{rd}|^2P_2+|h_{sd}|^2P),\\I(X_1,X_2;Y_2)&=\frac{1}{2}\log_2(1+|h_{sw}c_1\\&\quad+h_{rw}|^2P_2+|h_{sw}|^2P).\end{aligned}$$

Hence, we have (16) shown at the bottom of the page. For NF, we let $X_1\sim\mathcal{N}(0,P_1)$, $X_2\sim\mathcal{N}(0,P_2)$. Here X_1,X_2 are independent, resulting in

$$\begin{aligned}&I(X_1;Y|X_2)\\&\quad=\frac{1}{2}\log_2\left(1+|h_{sd}|^2P_1\right)\\&I(X_1,X_2;Y)-I(X_1,X_2;Y_2)\\&\quad=\frac{1}{2}\log_2\left(\frac{1+|h_{sd}|^2P_1+|h_{rd}|^2P_2}{1+|h_{sw}|^2P_1+|h_{rw}|^2P_2}\right),\\&I(X_2;Y_2|X_1)+I(X_1;Y|X_2)-I(X_1,X_2;Y_2)\\&\quad=\frac{1}{2}\log_2\left(\frac{(1+|h_{rw}|^2P_2)(1+|h_{sd}|^2P_1)}{1+|h_{sw}|^2P_1+|h_{rw}|^2P_2}\right).\end{aligned}$$

$$R_s^{(\mathrm{DF})}=\max_{c_1,P}\left[\min\left\{\frac{1}{2}\log_2\left(\frac{1+|h_{sr}|^2P}{1+|h_{sw}c_1+h_{rw}|^2P_2+|h_{sw}|^2P}\right),\frac{1}{2}\log_2\left(\frac{1+|h_{sd}c_1+h_{rd}|^2P_2+|h_{sd}|^2P}{1+|h_{sw}c_1+h_{rw}|^2P_2+|h_{sw}|^2P}\right)\right\}\right]^+. \tag{16}$$

Hence, we have

$$\begin{aligned} R_s^{(NF)} &= \left[\min\left\{\frac{1}{2}\log_2\left(1+|h_{sd}|^2P_1\right),\right.\right.\\ &\quad \frac{1}{2}\log_2\left(\frac{1+|h_{sd}|^2P_1+|h_{rd}|^2P_2}{1+|h_{sw}|^2P_1+|h_{rw}|^2P_2}\right),\\ &\quad \left.\left.\frac{1}{2}\log_2\left(\frac{(1+|h_{rw}|^2P_2)(1+|h_{sd}|^2P_1)}{1+|h_{sw}|^2P_1+|h_{rw}|^2P_2}\right)\right\}\right]^+ \end{aligned} \tag{17}$$

2) Amplify and Forward: In this subsection, we quantify the achievable secrecy rate of Amplify and Forward (AF) cooperation.[3] In AF, the source encodes its messages into codewords with length ML each, and divides each codeword into L subblocks each with M symbols, where L is chosen to be sufficiently large. At each subblock, the relay sends a linear combination of the received noisy signal of this subblock so far. For simplicity, we limit our discussion to $M=2$. In this case, the source sends $X_1(1)$ at the first symbol interval of each subblock, the relay receives $Y_1(1)=h_{sr}X_1(1)+Z_1(1)$; At the second symbol interval, the source sends $\alpha X_1(1)+\beta X_1(2)$, while the relay sends $\gamma Y_1(1)$. Here α,β,γ are chosen to satisfy the average power constraints of the source and the relay. Thus, this scheme allows beam-forming between the source and relay without requiring the relay to fully decode.

Writing the signal received at the destination and the eavesdropper in matrix form, we have

$$\mathbf{Y}=\mathbf{H}_1\mathbf{X}_1+\mathbf{Z},\quad \mathbf{Y}_2=\mathbf{H}_2\mathbf{X}_1+\mathbf{Z}_2 \tag{18}$$

where

$$\begin{aligned} \mathbf{H}_1 &= \begin{bmatrix} h_{sd} & 0\\ \beta h_{sd}+\gamma h_{sr}h_{rd} & \alpha h_{sd}\end{bmatrix}\\ \mathbf{H}_2 &= \begin{bmatrix} h_{sw} & 0\\ \beta h_{sw}+\gamma h_{sr}h_{rw} & \alpha h_{sw}\end{bmatrix}\\ \mathbf{X}_1 &= [X_1(1),X_1(2)]^T\\ \mathbf{Z} &= [Z(1),\gamma h_{rd}Z_1(1)+Z(2)]^T\\ \mathbf{Z}_2 &= [Z_2(1),\gamma h_{rw}Z_1(1)+Z_2(2)]^T\\ \mathbf{Y} &= [Y(1),Y(2)]^T\\ \mathbf{Y}_2 &= [Y_2(1),Y_2(2)]^T. \end{aligned} \tag{19}$$

The channel under consideration can be viewed as an equivalent standard memoryless eavesdropper channel with input $\mathbf{X}_1$ and outputs $\mathbf{Y},\mathbf{Y}_2$ at the destination and the eavesdropper respectively. Then, based on the result of [3], an achievable perfect secrecy rate is $[I(\mathbf{X}_1;\mathbf{Y})-I(\mathbf{X}_1;\mathbf{Y}_2)]^+$. Choosing a Gaussian input with covariance matrix $\mathbb{E}\{\mathbf{X}\mathbf{X}^H\}=P\mathbf{I}$, where $\mathbf{I}$ is the identity matrix, we get the following perfect secrecy rate:

$$\begin{aligned} R_s^{(AF)} &= \max_{\alpha,\beta,\gamma,P}\left[\frac{1}{4}\log_2\frac{|\det\{P\mathbf{H}_1\mathbf{H}_1^H+\mathbb{E}\{\mathbf{Z}\mathbf{Z}^H\}\}|}{|\det\{\mathbb{E}\{\mathbf{Z}\mathbf{Z}^H\}\}|}\right.\\ &\quad \left.-\frac{1}{4}\log_2\frac{|\det\{P\mathbf{H}_2\mathbf{H}_2^H+\mathbb{E}\{\mathbf{Z}_2\mathbf{Z}_2^H\}\}|}{|\det\{\mathbb{E}\{\mathbf{Z}_2\mathbf{Z}_2^H\}\}|}\right]^+\\ &= \max_{\alpha,\beta,\gamma,P}\left[\frac{1}{4}\log_2\frac{|\det\{P\mathbf{H}_1\mathbf{H}_1^H+\mathbf{A}\}\det\mathbf{B}|}{|\det\{P\mathbf{H}_2\mathbf{H}_2^H+\mathbf{B}\}\det\mathbf{A}|}\right]^+ \end{aligned} \tag{20}$$

where

$$\mathbf{A}=\begin{bmatrix}1 & 0\\ 0 & 1+|\gamma h_{rd}|^2\end{bmatrix}$$
$$\mathbf{B}=\begin{bmatrix}1 & 0\\ 0 & 1+|\gamma h_{rw}|^2\end{bmatrix}$$

and the maximization is over the set of power constraints

$$\begin{aligned} (1+|\alpha|^2+|\beta|^2)P &\le 2P_1\\ |\gamma|^2(|h_{sr}|^2P+1) &\le 2P_2. \end{aligned} \tag{21}$$

3) Numerical Results: In this section, we give numerical results under two channel models. The first is the real channel where $h_{ij}=d_{ij}^{-\gamma}$, with d_{ij} being the distance between node i and j and $\gamma>1$ is the channel attenuation coefficient. In the second model, we assume that each channel experiences an independent phase fading, that is $h_{ij}=d_{ij}^{-\gamma}e^{j\theta_{ij}}$, where θ_{ij} is uniformly distributed over $[0,2\pi)$. We believe that the second model is more practically relevant than the real channel scenario.

Fig. 3 shows the achievable perfect secrecy rate of the proposed schemes for the first channel model. In generating this figure, we use the network topology shown in Fig. 4, where we put the source at $(0,0)$, the destination at $(1,0)$, the eavesdropper at $(0,1)$, and the relay node at $(x,0)$. We let $P_1=1, P_2=8$. Since $d_{sd}=d_{sw}$, the perfect secrecy capacity of the eavesdropper channel without the relay node is zero. But, as shown in the figure, we can achieve a positive secrecy rate by introducing a relay node. In computing the upper bound, we set $V_1\sim\mathcal{N}(0,P_1), V_2\sim\mathcal{N}(0,P_2)$ with a correlation coefficient ρ, and maximize over $\rho\in[-1,1]$. We also set U to be a constant. Notice that the Gaussian input is not necessarily optimal for the upper bound. We can see that, when the relay is near the source, the DF scheme touches the Gaussian upper bound. Also, when $x>1$, it is clear that DF cooperation does not offer any gain, while NF and AF still offer positive rates. Notice that when $x>1$, both d_{sr} is larger than d_{sw}. The interesting observation here is that though both the destination and relay are not in advantageous positions compared with the eavesdropper, they can cooperate with each other and gain some advantage over the eavesdropper. If the relay is at 0, our model is equivalent to the case where the source has two antennas. Notice that the upper-bound of the perfect secrecy capacity is zero under this scenario. Hence, increasing the number of transmitting antenna at the source does not increase the secrecy capacity under the real channel model and this particular network topology. On the other hand, if there is a relay node at an appropriate position, we can exploit this relay node to establish a secure source–destination link.

In the second scenario, we assume that before transmission, the source knows the phases $\theta_{sr},\theta_{sd},\theta_{rd}$, but does not know θ_{sw},θ_{rw}. The random phase will not affect the achievable perfect secrecy rate of NF since it does not depend on beamforming between the source and relay. But, the rates of DF

[3] We did not consider this scheme in the discrete case since, in general, it does not lend itself to a single letter characterization.

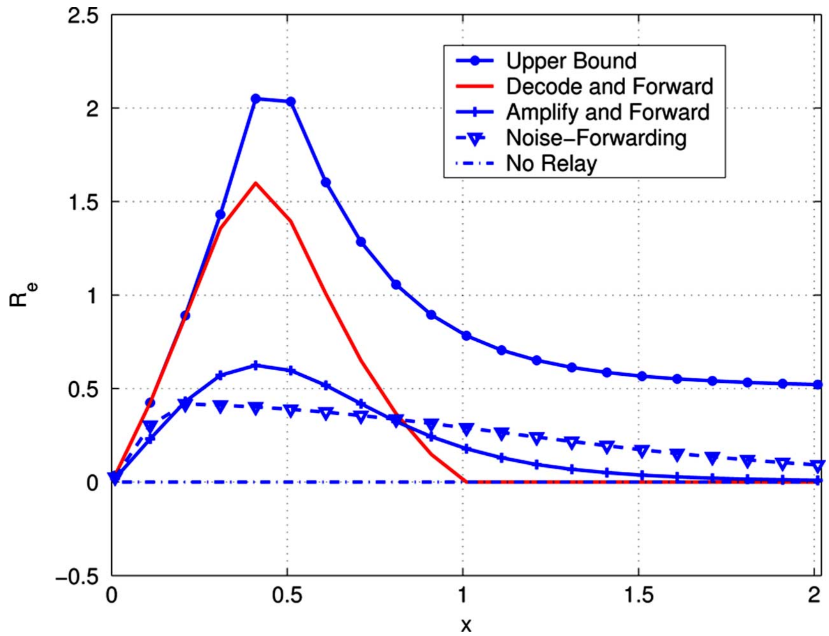

Fig. 3. The achievable perfect secrecy rate of the proposed schemes in the Gaussian relay eavesdropper channel.

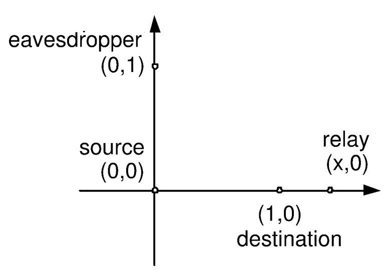

Fig. 4. The network topology.

and AF are different here. In both cases, the source can adjust its phase according to the knowledge of the phase information about $\theta_{sr}, \theta_{sd}, \theta_{rd}$. In this way, the signals of the source and the relay will add up coherently at the destination, but not at the eavesdropper since θ_{sw}, θ_{rw} are independent of $\theta_{sd}, \theta_{rd}, \theta_{sr}$. The secrecy rate of DF and AF could then be obtained by averaging (16), (20) over the random phases. Fig. 5 shows the achievable perfect secrecy rates of the proposed strategies for the same setup as the first scenario. Due to the random phases, the achievable perfect secrecy capacity when the relay is at the same position as the source is not zero anymore. In this case, it will be beneficial to have multiple transmitting antennas at the source. Similar to the first scenario, when $x > 1$, DF cooperation does not offer any benefit. But both NF and AF still enjoy nonzero secrecy rates.

V. Conclusion

In this paper, the relay–eavesdropper channel was studied. In particular, several cooperation strategies were proposed and the corresponding achievable performance bounds were obtained. Furthermore, an outer bound on the optimal rate-equivocation region for this channel was developed. Of particular interest is the proposed NF strategy that was used to illustrate the deaf-helper phenomenon, and to demonstrate the utility of the relay node in the reversely degraded relay–eavesdropper channel. Overall, our results establish the critical role of user cooperation in facilitating secure wireless communications and shed light on the unique feature of the relay–eavesdropper channel.

Among the many open problems posed by our work, how to close the gap between the achievable performance and the outer bound is arguably the most important one. This problem is expected to be challenging since the capacity of the classical relay channel remains unknown. The investigation of the role of feedback in the relay–eavesdropper channel is another interesting problem. In the relay channel without security constraints, noiseless/noisy feedback was shown to be beneficial. On the other hand, in the presence of an eavesdropper, the role and optimal mechanism of feedback is not yet known, since the eavesdropper could also benefit from the feedback signal. Finally, extending our work to a large scale network is expected to be of practical significance.

Appendix A
Proof of Theorem 1

The proof follows that of [3].

$$
\begin{aligned}
nR_e &= H(W_1|Y_2^n) && (22)\\
&= H(W_1) - I(W_1;Y_2^n) && (23)\\
&= I(W_1;Y^n) - I(W_1;Y_2^n)\\
&\quad + H(W_1|Y^n) && (24)\\
&\le \sum_{i=1}^{n}[I(W_1;Y_i|Y^{i-1})\\
&\quad - I(W_1;Y_{2,i}|Y_{2,i+1}^n)] + n\delta_n && (25)
\end{aligned}
$$

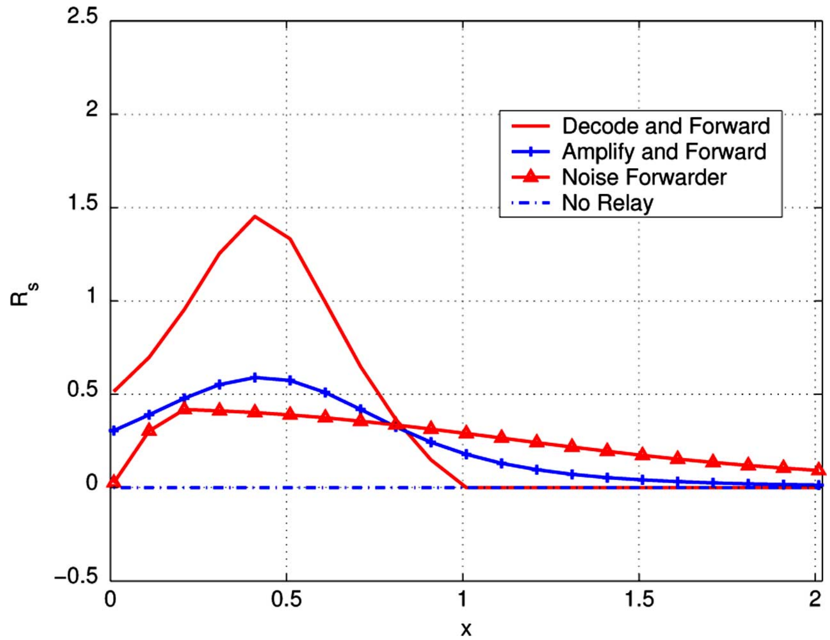

Fig. 5. The achievable perfect secrecy capacity for various schemes in the Gaussian relay eavesdropper channel with phase fading.

where $Y^{i-1} = Y(1,\ldots,i-1), Y_{2,i+1}^n = Y_2(i+1,\ldots,n)$, and $\delta_n \to 0$ as $n \to \infty$. We get this by using the chain rule to expand $I(W_1;Y^n)$ from $i=1$ and expand $I(W_1;Y_2^n)$ from $i=n$, also we use the Fano's inequality to bound $H(W_1|Y^n)$.

We continue

$$\begin{aligned} nR_e \le& \sum_{i=1}^n [I(W_1;Y_i|Y^{i-1}) \\ &- I(W_1;Y_{2,i}|Y_{2,i+1}^n)] + n\delta_n \qquad (26)\\ =& \sum_{i=1}^n [I(W_1,Y_{2,i+1}^n;Y_i|Y^{i-1}) \\ &- I(Y_{2,i+1}^n;Y_i|Y^{i-1},W_1) \qquad (27)\\ &- I(W_1,Y^{i-1};Y_{2,i}|Y_{2,i+1}^n) \\ &+ I(Y^{i-1};Y_{2,i}|Y_{2,i+1}^n,W_1)] + n\delta_n \qquad (28)\\ =& \sum_{i=1}^n [I(W_1,Y_{2,i+1}^n;Y_i|Y^{i-1}) \\ &- I(W_1,Y^{i-1};Y_{2,i}|Y_{2,i+1}^n)] + n\delta_n \qquad (29)\end{aligned}$$

since $\sum_{i=1}^n I(Y_{2,i+1}^n;Y_i|Y^{i-1},W_1) = \sum_{i=1}^n I(Y^{i-1};Y_{2,i}|Y_{2,i+1}^n,W_1)$, which is proved in the Lemma 7 of [3]. Now

$$\begin{aligned} nR_e \le& \sum_{i=1}^n [I(W_1,Y_{2,i+1}^n;Y_i|Y^{i-1}) \\ &- I(W_1,Y^{i-1};Y_{2,i}|Y_{2,i+1}^n)] + n\delta_n \qquad (30)\\ =& \sum_{i=1}^n [I(Y_{2,i+1}^n;Y_i|Y^{i-1}) \\ &+ I(W_1;Y_i|Y^{i-1},Y_{2,i+1}^n) \\ &- I(Y^{i-1};Y_{2,i}|Y_{2,i+1}^n) \\ &- I(W_1;Y_{2,i}|Y^{i-1},Y_{2,i+1}^n)] + n\delta_n \qquad (31)\\ =& \sum_{i=1}^n [I(W_1;Y_i|Y^{i-1},Y_{2,i+1}^n) \\ &- I(W_1;Y_{2,i}|Y^{i-1},Y_{2,i+1}^n)] + n\delta_n \qquad (32)\end{aligned}$$

since $\sum_{i=1}^n I(Y_{2,i+1}^n;Y_i|Y^{i-1}) = \sum_{i=1}^n I(Y^{i-1};Y_{2,i}|Y_{2,i+1}^n)$, which is also proved in [3].

Now, let J be a random variable uniformly distributed over $\{1,\ldots,n\}$, set $U = JY^{i-1}Y_{2,i+1}^n$, $V_1 = JY_{2,i+1}^n W_1, V_2 = JY^{i-1}, Y_1 = Y_{1,J}, Y_2 = Y_{2,J}, Y = Y_J, X_1 = X_{1,J}, X_2 = X_{2,J}$, we have

$$\begin{aligned} R_e \le& \frac{1}{n}\sum_{i=1}^n [I(W_1;Y_i|Y^{i-1},Y_{2,i+1}^n) \\ &- I(W_1;Y_{2,i}|Y^{i-1},Y_{2,i+1}^n)] + \delta_n \\ =& \frac{1}{n}\sum_{i=1}^n [I(W_1,Y^{i-1},Y_{2,i+1}^n;Y_i|Y^{i-1},Y_{2,i+1}^n) \\ &- I(W_1,Y^{i-1},Y_{2,i+1}^n;Y_{2,i}|Y^{i-1},Y_{2,i+1}^n)] + \delta_n \\ =& I(V_1,V_2;Y|U) - I(V_1,V_2;Y_2|U) + \delta_n. \qquad (33)\end{aligned}$$

Since the channel is memoryless, one can then check that $U \to (V_1,V_2) \to (X_1,X_2) \to (Y,Y_1,Y_2)$ is a Markov chain.

Also, continuing from (32), we have

$$\begin{aligned} R_e \le& \frac{1}{n}\sum_{i=1}^n [I(W_1;Y_i|Y^{i-1},Y_{2,i+1}^n) \\ &- I(W_1;Y_{2,i}|Y^{i-1},Y_{2,i+1}^n)] + \delta_n \\ =& \frac{1}{n}\sum_{i=1}^n [I(W_1,Y_{2,i+1}^n;Y_i|Y^{i-1},Y_{2,i+1}^n) \\ &- I(W_1,Y^{i-1},Y_{2,i+1}^n;Y_{2,i}|Y^{i-1},Y_{2,i+1}^n)] + \delta_n \\ \le& \frac{1}{n}\sum_{i=1}^n [I(W_1,Y_{2,i+1}^n;Y_i,Y_{1i}|Y^{i-1},Y_{2,i+1}^n) \\ &- I(W_1,Y^{i-1},Y_{2,i+1}^n;Y_{2,i}|Y^{i-1},Y_{2,i+1}^n)] + \delta_n \\ =& I(V_1;Y,Y_1|V_2,U) - I(V_1,V_2;Y_2|U) + \delta_n. \qquad (34)\end{aligned}$$

In the following, we bound R_1:

$$\begin{aligned} I(W_1;\mathbf{Y}) &= \sum_{i=1}^{n} I(W_1;Y_i|Y^{i-1}) \\ &= \sum_{i=1}^{n}[H(Y_i|Y^{i-1}) - H(Y_i|W_1,Y^{i-1})] \\ &\le \sum_{i=1}^{n}[H(Y_i) - H(Y_i|W_1,Y^{i-1})] \\ &\le \sum_{i=1}^{n}[H(Y_i) - H(Y_i|W_1,Y^{i-1},Y_{2,i+1}^{n})] \\ &= \sum_{i=1}^{n} I(W_1,Y^{i-1},Y_{2,i+1}^{n};Y_i). \end{aligned} \tag{35}$$

Hence

$$R_1 \le \frac{1}{n} I(W_1;\mathbf{Y}) \le I(V_1,V_2;Y). \tag{36}$$

Also

$$\begin{aligned} I(W_1;\mathbf{Y}) &= \sum_{i=1}^{n} I(W_1;Y_i|Y^{i-1}) \\ &\le \sum_{i=1}^{n} I(W_1;Y_i,Y_{1,i}|Y^{i-1}) \\ &= \sum_{i=1}^{n}[H(Y_i,Y_{1,i}|Y^{i-1}) \\ &\quad - H(Y_i,Y_{1,i}|W_1,Y^{i-1})] \\ &\le \sum_{i=1}^{n}[H(Y_i,Y_{1,i}|Y^{i-1}) \\ &\quad - H(Y_i,Y_{1,i}|W_1,Y^{i-1},Y_{2,i+1}^{n})] \\ &= \sum_{i=1}^{n} I(W_1,Y_{2,i+1}^{n};Y_i,Y_{1,i}|Y^{i-1}). \end{aligned} \tag{37}$$

Hence, we have

$$R_1 \le \frac{1}{n} I(W_1;\mathbf{Y}) = I(V_1;Y,Y_1|V_2). \tag{38}$$

So, we have

$$R_1 \le \min\{I(V_1,V_2;Y), I(V_1;Y,Y_1|V_2)\}. \tag{39}$$

The claim is proved.

Appendix B
Proof of Theorem 2

The proof is a combination of the coding schemes of Csiszár *et al.* [3] and the regular coding and backward decoding scheme in the relay channel [7], [20]. We first replace V_1, V_2 in Theorem 2 with X_1, X_2. After proving Theorem 2 with V_1, V_2 replaced by X_1, X_2, we then prefix a memoryless channel with input V_1, V_2 and transmission probability $p(x_1, x_2|v_1, v_2)$ as reasoned in [3] to finish our proof.

1) Codebook Generation: Randomly generated a typical sequence $\mathbf{u}$ with probability $p(\mathbf{u}) = \prod_{i=1}^{n} p(u_i)$. We assume that all the terminals know $\mathbf{u}$. We first generate at random 2^{nR} independent and identically distributed (i.i.d.) n-sequence at the relay node each drawn according to $p(\mathbf{x}_2|\mathbf{u}) = \prod_{i=1}^{n} p(x_{2,i}|u_i)$, index them as $\mathbf{x}_2(a), a \in [1, 2^{nR}]$, where $R = \min\{I(X_1,X_2;Y|U), I(X_1;Y_1|X_2,U)\} - \epsilon_0$. For each $\mathbf{x}_2(a)$, generate 2^{nR} conditionally independent n-sequence $\mathbf{x}_1(k,a), k \in [1, 2^{nR}]$ drawn randomly according to $p(\mathbf{x}_1|\mathbf{x}_2(a),\mathbf{u}) = \prod_{i=1}^{n} p(x_{1,i}|x_{2,i}(a),u_i)$. Define $\mathcal{W} = \{1,\ldots,2^{n[R-I(X_1,X_2;Y_2|U)]}\}$, $\mathcal{L} = \{1,\ldots,2^{nI(X_1,X_2;Y_2|U)}\}$ and $\mathcal{K} = \mathcal{W} \times \mathcal{L} = \{1,\ldots,2^{nR}\}$.

In the following, we assume that $R - I(X_1,X_2;Y_2|U) > 0$. If this is not the case, we set $\mathcal{W} = \{1\}$ and $\mathcal{L} = \{1,\ldots,2^{nR}\}$. This DF strategy does not achieve any security level. In this case, we achieve $(R, 0)$ which is still inside the region given in this theorem.

2) Encoding: We exploit the block Markov coding scheme, as argued in [4], the loss induced by this scheme is negligible as the number of blocks $B \to \infty$.

For a given rate pair (R_1, R_e) with $R_1 \le R$ and $R_e \le R_1$, we give the following coding strategy. Let the message to be transmitted at block i be $w_1(i) \in \mathcal{W}_1 = \{1,\ldots,M\}$, where $M = 2^{nR_1}$.

The stochastic encoder at the transmitter first forms the following mappings.

- If $R_1 > R - I(X_1,X_2;Y_2|U)$, then we let $\mathcal{W}_1 = \mathcal{W} \times \mathcal{J}$, where

$$\mathcal{J} = \{1,\ldots,2^{n(R_1 - [R - I(X_1,X_2;Y_2|U)])}\}.$$

  We let g_1 be the partition that partitions $\mathcal{L}$ into $|\mathcal{J}|$ equal size subsets. The stochastic encoder at transmitter will choose a mapping for each message $w_1(i) = (w(i), j(i)) \to (w(i), l(i))$, where $l(i)$ is chosen randomly from the set $g_1^{-1}(j(i)) \subset \mathcal{L}$ with uniform distribution.
- If $R_1 < R - I(X_1,X_2;Y_2|U)$, the stochastic encoder will choose a mapping $w_1(i) \to (w_1(i), l(i))$, where $l(i)$ is chosen uniformly from the set $\mathcal{L}$.

Assume that the message $w_1(i-1)$ transmitted at block $i-1$ is associated with $(w(i-1), l(i-1))$ and the message $w_1(i)$ intended to send at block i is associated with $(w(i), l(i))$ by the stochastic encoder at the transmitter. We let $a(i-1) = (w(i-1), l(i-1))$ and $b(i) = (w(i), l(i))$. The encoder then sends $\mathbf{x}_1(b(i), a(i-1))$. The relay has an estimation $\hat{\hat{a}}(i-1)$ (see the decoding part), and thus sends the corresponding codeword $\mathbf{x}_2(\hat{\hat{a}}(i-1))$.

At block 1, the source sends $\mathbf{x}_1(b(1), 1)$, the relay sends $\mathbf{x}_2(1)$.

At block B, the source sends $\mathbf{x}_1(1, a(B-1))$, and the relay sends $\mathbf{x}_2(\hat{\hat{a}}(B-1))$.

3) Decoding: At the end of block i, the relay already has an estimation of the $\hat{\hat{a}}(i-1)$, which was sent at block $i-1$, and will declare that it receives $\hat{\hat{a}}(i)$, if this is the only pair such that $(\mathbf{x}_1(\hat{\hat{a}}(i), \hat{\hat{a}}(i-1)), \mathbf{x}_2(\hat{\hat{a}}(i-1)), \mathbf{y}_1(i), \mathbf{u})$ are jointly typical. Since $R = \min\{I(X_1;Y_1|X_2,U),\ I(X_1,X_2;Y|U)\} - \epsilon \le I(X_1;Y_1|X_2,U) - \epsilon$, then based on the AEP, one has $\hat{\hat{a}}(i) = a(i)$ with probability goes to 1.

The destination decodes from the last block, i.e., block B. Suppose that at the end of block $B-1$, the relay decodes successfully, then the destination will declare that $\hat{a}(B-1)$ is received, if $(\mathbf{x}_1(1, \hat{a}(B-1)), \mathbf{x}_2(\hat{a}(B-1)), \mathbf{y}, \mathbf{u})$ are jointly typical. It's easy to see that if $R \leq I(X_1, X_2; Y|U)$, we will have $\hat{a}(B-1) = a(B-1)$ with probability goes to 1, as n increases.

After getting $\hat{a}(B-1)$, the receiver can get an estimation of $a(i), i \in [1, B-2]$ in a similar way.

Having $\hat{a}(i-1)$, the destination can get the estimation of the message $w_1(i-1)$ by letting 1) $\hat{w}_1(i-1) = (\hat{w}(i-1), \hat{j}(i-1)) = (\hat{w}(i-1), g_1(\hat{l}(i-1)))$ if $R_1 > R - I(X_1, X_2; Y_2|U)$, 2) $\hat{w}_1(i-1) = \hat{w}(i-1)$ if $R_1 < R - I(X_1, X_2; Y_2|U)$.

The probability that $\hat{w}_1(i-1) = w_1(i-1)$ goes to one for sufficiently large n.

4) Equivocation Computation:

$$\begin{aligned} H(W_1|\mathbf{Y}_2) &\geq H(W_1|\mathbf{Y}_2, \mathbf{U}) \quad (40)\\ &= H(W_1, \mathbf{Y}_2|\mathbf{U}) - H(\mathbf{Y}_2|\mathbf{U})\\ &= H(W_1, \mathbf{Y}_2, \mathbf{X}_1, \mathbf{X}_2|\mathbf{U})\\ &\quad - H(\mathbf{X}_1, \mathbf{X}_2|W_1, \mathbf{Y}_2, \mathbf{U}) - H(\mathbf{Y}_2|\mathbf{U})\\ &= H(\mathbf{X}_1, \mathbf{X}_2|\mathbf{U})\\ &\quad + H(W_1, \mathbf{Y}_2|\mathbf{X}_1, \mathbf{X}_2, \mathbf{U})\\ &\quad - H(\mathbf{X}_1, \mathbf{X}_2|W_1, \mathbf{Y}_2, \mathbf{U}) - H(\mathbf{Y}_2|\mathbf{U})\\ &\geq H(\mathbf{X}_1|\mathbf{U}) + H(\mathbf{Y}_2|\mathbf{X}_1, \mathbf{X}_2, \mathbf{U})\\ &\quad - H(\mathbf{X}_1, \mathbf{X}_2|W_1, \mathbf{Y}_2, \mathbf{U}) - H(\mathbf{Y}_2|\mathbf{U}). \quad (41) \end{aligned}$$

First, let us calculate $H(\mathbf{X}_1, \mathbf{X}_2|W_1, \mathbf{Y}_2, \mathbf{U})$. Given W_1, the eavesdropper can also do backward decoding as the receiver. At the end of block B, given W_1, the eavesdropper knows $w(B-1)$, hence it will decode $l(B-1)$, by letting $l(B-1) = \hat{l}(B-1)$, if $\hat{l}(B-1)$ is the only one such that $(\mathbf{x}_1(1, (w(B-1), \hat{l}(B-1))), \mathbf{x}_2((w(B-1), \hat{l}(B-1))), \mathbf{y}, \mathbf{u})$ are jointly typical. Since $l \in [1, 2^{nI(X_1, X_2; Y_2)}]$, we have

$$\begin{aligned} \Pr\{(\mathbf{X}_1(1, a(i-1)), \mathbf{X}_2(\hat{\hat{a}}(i-1)))\\ \neq (\mathbf{X}_1(1, (w(B-1), \hat{l}(B-1))),\\ \mathbf{X}_2((w(B-1), \hat{l}(B-1))))\} \leq \epsilon_1. \quad (42) \end{aligned}$$

Then based on Fano's inequality, we have

$$\frac{1}{n} H(\mathbf{X}_1, \mathbf{X}_2|W_1 = w_1, \mathbf{Y}_2, \mathbf{U}) \leq \frac{1}{n} + \epsilon_1 I(X_1, X_2; Y_2|U) \quad (43)$$

Hence, we have

$$\begin{aligned} \frac{1}{n} H(\mathbf{X}_1, \mathbf{X}_2|W_1, \mathbf{Y}_2, \mathbf{U})\\ &= \frac{1}{n} \sum_{w_1 \in \mathcal{W}_1} p(W_1\\ &= w_1) H(\mathbf{X}_1, \mathbf{X}_2|W_1 = w_1, \mathbf{Y}_2, \mathbf{U}) \leq \epsilon_2 \quad (44) \end{aligned}$$

when n is sufficiently large.

Since the channel is memoryless, we have $H(\mathbf{Y}_2|\mathbf{U}) - H(\mathbf{Y}_2|\mathbf{X}_1, \mathbf{X}_2, \mathbf{U}) \leq nI(X_1, X_2; Y_2|U) + n\delta_n$, where $\delta_n \to 0$, as $n \to \infty$ [2].

Now, from the code construction, we have $H(\mathbf{X}_1|\mathbf{U}) = nR$ if $R_1 > R - I(X_1, X_2; Y_2|U)$. In this case, we get $nR_e = H(W_1|\mathbf{Y}_2) \geq n(R - I(X_1, X_2; Y_2|U) - \epsilon_3)$. If $R_1 \leq R - I(X_1, X_2; Y_2|U)$, $H(\mathbf{X}_1|\mathbf{U}) = n(R_1 + I(X_1, X_2; Y_2|U))$, in this case, we get the perfect secrecy, since

$$\begin{aligned} nR_e &\geq n(R_1 + I(X_1, X_2; Y_2|U))\\ &\quad - nI(X_1, X_2; Y_2|U) - n\epsilon_3\\ &\geq n(R_1 - \epsilon_3). \end{aligned}$$

The claim is proved.

Appendix C
Proof of Theorem 3

As [3], we first prove the result for the case where V_1, V_2 in Theorem 3 are replaced with X_1, X_2, then prefix a memoryless channel with transition probability $p(x_1|v_1)p(x_2|v_2)$ to finish our proof.

We first consider the case $I(X_1; Y|X_2) < I(X_1; Y_2|X_2)$, i.e., the channel between the source and the eavesdropper is better than the channel between the source and the destination. In this case, we only need to consider $\min\{I(X_2; Y), I(X_2; Y_2)\} = I(X_2; Y_2)$, otherwise, the secrecy rate will be zero. Thus in this case, the last equation in (9) changes to $R_e < [I(X_1; Y|X_2) + \min\{I(X_2; Y), I(X_2; Y_2|X_1)\} - I(X_1, X_2; Y_2)]^+$.

1) Codebook Generation: For a given distribution $p(x_1)p(x_2)$, we generate at random 2^{nR_2} i.i.d. n-sequence at the relay node each drawn according to $p(\mathbf{x}_2) = \prod_{i=1}^{n} p(x_{2,i})$, index them as $\mathbf{x}_2(a), a \in [1, 2^{nR_2}]$, where we set $R_2 = \min\{I(X_2; Y), I(X_2; Y_2|X_1)\} - \epsilon$. We also generate random 2^{nR} i.i.d n-sequence at the source each drawn according to $p(\mathbf{x}_1) = \prod_{i=1}^{n} p(x_{1,i})$, index them as $\mathbf{x}_1(k), k \in [1, 2^{nR}]$, where $R = I(X_1; Y|X_2) - \epsilon$. Let

$$\begin{aligned} R' = \min\{I(X_2; Y), I(X_2; Y_2|X_1)\}\\ + I(X_1; Y|X_2) - I(X_1, X_2; Y_2), \end{aligned}$$

and define $\mathcal{W} = \{1, \ldots, 2^{nR'}\}$, $\mathcal{L} = \{1, \ldots, 2^{n(R-R')}\}$ and $\mathcal{K} = \mathcal{W} \times \mathcal{L} = \{1, \ldots, 2^{nR}\}$.

In the following, we assume that $R - R' > 0$. If this is not the case, we set $\mathcal{W} = \{1\}$ and $\mathcal{L} = \{1, \ldots, 2^{nR}\}$. This NF strategy does not achieve any security level. In this case, we achieve $(R, 0)$ which is still inside the region given in this theorem.

2) Encoding: For a given rate pair (R_1, R_e) with $R_1 \leq R, R_e \leq R_1$, we give the following coding strategy. Let the message to be transmitted at block i be $w_1(i) \in \mathcal{W}_1 = [1, M]$, where $M = 2^{nR_1}$.

The stochastic encoder at the transmitter first forms the following mappings.

- If $R_1 > R'$, then we let $\mathcal{W}_1 = \mathcal{W} \times \mathcal{J}$, where $\mathcal{J} = \{1, 2^{n(R_1 - R')}\}$. We let g_1 be the partition that partitions $\mathcal{L}$ into $|\mathcal{J}|$ equal size subsets. The stochastic encoder at transmitter will choose a mapping for each message $w_1(i) = (w(i), j(i)) \to (w(i), l(i))$, where $l(i)$ is chosen randomly from the set $g_1^{-1}(j(i)) \subset \mathcal{L}$ with uniform distribution.

- If $R_1 < R^{'}$, the stochastic encoder will choose a mapping $w_1(i) \to (w_1(i), l(i))$, where $l(i)$ is chosen uniformly from the set $\mathcal{L}$.

Suppose the message $w_1(i)$ intended to send at block i is associated with $(w(i), l(i))$ by the stochastic encoder at the transmitter. The encoder then sends $\mathbf{x}_1((w(i), l(i)))$. The relay uniformly picks a code $\mathbf{x}_2(a)$ from $a \in [1, \ldots, 2^{nR_2}]$, and sends $\mathbf{x}_2(a)$.

3) Decoding: At the end of block i, the destination declares that $\hat{a}(i)$ is received, if $\hat{a}(i)$ is the only one such that $(\mathbf{x}_2(\hat{a}(i)), \mathbf{y})$ are jointly typical. If there does not exist or there exist more than one such sequences, the destination declares an error. Since $R_2 = \min\{I(X_2;Y), I(X_2;Y_2|X_1)\} - \epsilon \leq I(X_2;Y) - \epsilon$, then based on AEP, we know that the error probability will be less than any given positive number ϵ, when the codeword length n is long enough.

The destination then declares that $\hat{k}$ is received, if $\hat{k}$ is the only one such that $(\mathbf{x}_1(\hat{k}), \mathbf{x}_2(\hat{a}), \mathbf{y})$ are jointly typical, otherwise declares an error. Since $R = I(X_1;Y|X_2) - \epsilon$, then based on AEP, we know that we will have error probability goes to zero, when n is sufficiently large.

Having $\hat{k}(i)$, the destination can get the estimation of the message $w_1(i)$ by letting

1) $\hat{w}_1(i) = (\hat{w}(i), \hat{j}(i)) = (\hat{w}(i), g_1(\hat{l}(i)))$, if $R_1 > R^{'}$;
2) $\hat{w}_1(i) = \hat{w}(i)$, if $R_1 < R^{'}$.

The probability that $\hat{w}_1(i) = w_1(i)$ goes to one for sufficiently large n.

4) Equivocation Computation:

$$\begin{aligned} H(W_1|\mathbf{Y}_2) &= H(W_1, \mathbf{Y}_2) - H(\mathbf{Y}_2) \\ &= H(W_1, \mathbf{Y}_2, \mathbf{X}_1, \mathbf{X}_2) \\ &\quad - H(\mathbf{X}_1, \mathbf{X}_2|W_1, \mathbf{Y}_2) - H(\mathbf{Y}_2) \\ &= H(\mathbf{X}_1, \mathbf{X}_2) + H(W_1, \mathbf{Y}_2|\mathbf{X}_1, \mathbf{X}_2) \\ &\quad - H(\mathbf{X}_1, \mathbf{X}_2|W_1, \mathbf{Y}_2) - H(\mathbf{Y}_2) \\ &\geq H(\mathbf{X}_1, \mathbf{X}_2) + H(\mathbf{Y}_2|\mathbf{X}_1, \mathbf{X}_2) \\ &\quad - H(\mathbf{X}_1, \mathbf{X}_2|W_1, \mathbf{Y}_2) - H(\mathbf{Y}_2). \end{aligned} \tag{45}$$

Now let's calculate $H(\mathbf{X}_1, \mathbf{X}_2|W_1, \mathbf{Y}_2)$. Given W_1, the eavesdropper can do joint decoding. At any block i, given W_1, the eavesdropper knows $w(i)$, hence it will decode $l(i)$ and $a(i)$ sent by the relay, by letting $l(i) = \hat{l}(i), a(i) = \hat{a}(i)$, if $\hat{l}(i), \hat{a}(i)$ are the only one such that $(\mathbf{x}_1(w(i), \hat{l}(i)), \mathbf{x}_2(\hat{a}(i)), \mathbf{y})$ are jointly typical. Then, since $R_2 = \min\{I(X_2;Y), I(X_2;Y_2|X_1)\} - \epsilon \leq I(X_2;Y_2|X_1) - \epsilon$, we get

$$\begin{aligned} &\frac{1}{2}\log_2(|\mathcal{L}|) + R_2 \\ &\quad = R + I(X_1, X_2; Y_2) \\ &\quad - \min\{I(X_2;Y), I(X_2;Y_2|X_1)\} \\ &\quad - I(X_1;Y|X_2) \\ &\quad + \min\{I(X_2;Y), I(X_2;Y_2|X_1)\} - \epsilon \\ &\quad \leq I(X_1, X_2; Y_2) - \epsilon \end{aligned} \tag{46}$$

Also, we have $\frac{1}{2}\log_2(|\mathcal{L}|) < R \leq I(X_1;Y_2|X_2) - \epsilon$.

So

$$\begin{aligned} &\Pr\{(\mathbf{X}_1(w(i), \hat{l}(i)), \mathbf{X}_2(\hat{a}(i))) \\ &\qquad \neq (\mathbf{X}_1(w(i), l(i)), \mathbf{X}_2(a(i)))\} \leq \epsilon_1. \end{aligned}$$

Then based on Fano's inequality, we have

$$\frac{1}{n}H(\mathbf{X}_1, \mathbf{X}_2|W_1 = w_1, \mathbf{Y}_2) \leq \frac{1}{n} + \epsilon_1 I(X_1, X_2; Y_2) \tag{47}$$

Hence, we have

$$\begin{aligned} \frac{1}{n}H(\mathbf{X}_1, \mathbf{X}_2|W_1, \mathbf{Y}_2) &= \frac{1}{n}\sum_{w_1 \in \mathcal{W}_1} p(W_1 = w_1) \\ &\quad \times H(\mathbf{X}_1, \mathbf{X}_2|W_1 = w_1, \mathbf{Y}_2) \leq \epsilon_2 \end{aligned} \tag{48}$$

when n is sufficiently large.

Now, $H(\mathbf{Y}_2) - H(\mathbf{Y}_2|\mathbf{X}_1, \mathbf{X}_2) \leq nI(X_1, X_2; Y_2) + n\delta_n$, where $\delta_n \to 0$, as $n \to \infty$.

Also we have $H(\mathbf{X}_1, \mathbf{X}_2) = H(\mathbf{X}_1) + H(\mathbf{X}_2)$ since $\mathbf{x}_1$ and $\mathbf{x}_2$ are independent. If $R_1 > R^{'}$, we have $H(\mathbf{X}_1, \mathbf{X}_2) = R + R_2 = I(X_1;Y|X_2) + \min\{I(X_2;Y), I(X_2;Y_2|X_1)\}$. Combining these, we get

$$\begin{aligned} nR_e = H(W_1|\mathbf{Y}_2) \geq n(&\min\{I(X_2;Y), I(X_2;Y_2|X_1)\} \\ &+ I(X_1;Y|X_2)) - nI(X_1, X_2; Y_2) - n\epsilon_4. \end{aligned}$$

On the other hand, if $R_1 < R^{'}$, we have

$$H(\mathbf{X}_1) = R_1 + I(X_1, X_2; Y_2) - \min\{I(X_2;Y), I(X_2;Y_2|X_1)\}$$

hence we have $H(\mathbf{X}_1, \mathbf{X}_2) = R_1 + I(X_1, X_2; Y_2) - \epsilon$. We get perfect secrecy rate, since $nR_e = H(W_1|\mathbf{Y}_2) \geq nR_1 - n\epsilon_4$.

This case is proved.

Now, consider the case $I(X_1;Y|X_2) > I(X_1;Y_2|X_2)$. If $\min\{I(X_2;Y), I(X_2;Y_2)\} = I(X_2;Y)$, then we have $\min\{I(X_2;Y), I(X_2;Y_2|X_1)\} = I(X_2;Y)$, because $I(X_2;Y_2|X_1) > I(X_2;Y_2)$ since X_1, X_2 are independent. Under this case, we only need to prove $R_e \leq I(X_1;Y|X_2) - I(X_1;Y_2|X_2)$ are achievable, which can be achieved by letting the codeword rate be $I(X_1;Y|X_2)$ and $R^{'} = I(X_1;Y|X_2) - I(X_1;Y|X_2)$. Now the equivocation rate of the eavesdropper can be calculated as

$$\begin{aligned} H(W_1|\mathbf{Y}_2) &\geq H(W_1|\mathbf{Y}_2, \mathbf{X}_2) \\ &= H(W_1, \mathbf{Y}_2|\mathbf{X}_2) - H(\mathbf{Y}_2|\mathbf{X}_2) \\ &= H(W_1, \mathbf{Y}_2, \mathbf{X}_1|\mathbf{X}_2) \\ &\quad - H(\mathbf{X}_1|W_1, \mathbf{Y}_2, \mathbf{X}_2) - H(\mathbf{Y}_2|\mathbf{X}_2) \\ &= H(\mathbf{X}_1|\mathbf{X}_2) + H(W_1, \mathbf{Y}_2|\mathbf{X}_1, \mathbf{X}_2) \\ &\quad - H(\mathbf{X}_1|W_1, \mathbf{Y}_2, \mathbf{X}_2) - H(\mathbf{Y}_2|\mathbf{X}_2) \\ &\geq H(\mathbf{X}_1) + H(\mathbf{Y}_2|\mathbf{X}_1, \mathbf{X}_2) \\ &\quad - H(\mathbf{X}_1|W_1, \mathbf{Y}_2, \mathbf{X}_2) - H(\mathbf{Y}_2|\mathbf{X}_2) \end{aligned} \tag{49}$$

since $\mathbf{x}_1, \mathbf{x}_2$ are independent. This can then be shown to be larger than $n(I(X_1;Y|X_2) - I(X_1;Y_2|X_2) - \epsilon)$.

If $\min\{I(X_2;Y), I(X_2;Y_2)\} = I(X_2;Y_2)$, the last line in (9) changes to $R_e < [I(X_1;Y|X_2) + \min\{I(X_2;Y), I(X_2;Y_2|X_1)\} - I(X_1, X_2; Y_2)]^+$, then we can use a coding/decoding scheme similar to the one developed above to show the achievability.

The claim is achieved.

APPENDIX D
PROOF OF THEOREM 4

The proof is a combination of the coding scheme of Csiszár *et al.* [3] and a revised CF scheme in the relay channel [4].

1) Codebook Generation: We first generate at random 2^{nR} i.i.d. n-sequence $\mathbf{x}_1$ at the source node each drawn according to $p(\mathbf{x}_1)=\prod_{j=1}^{n}p(x_{1,j})$, index them as $\mathbf{x}_1(k), k\in[1,2^{nR}]$, with $R=I(X_1;\hat{Y}_1,Y|X_2)-\epsilon$.

Generate at random 2^{nR_2} i.i.d. n-sequence $\mathbf{x}_2$ each with probability $p(\mathbf{x}_2)=\prod_{j=1}^{n}p(x_{2,j})$. Index these as $\mathbf{x}_2(s), s\in[1,2^{nR_2}]$, where

$$R_2=\min\{I(X_2;Y),I(X_2;Y_2|X_1)\}-\epsilon.$$

For each $\mathbf{x}_2(s)$, generate at random $2^{n(R_2-R_0)}$ i.i.d. $\hat{\mathbf{y}}_1$, each with probability $p(\hat{\mathbf{y}}_1|\mathbf{x}_2(s))=\prod_{j=1}^{n}p(\hat{y}_{1,j}|x_{2,j}(s))$. Label these $\hat{\mathbf{y}}_1(z,s), z\in[1,2^{n\hat{R}}], s\in[1,2^{nR_2}]$, where we set $\hat{R}=R_2-R_0$. Equally divide these $2^{nR_2}\mathbf{x}_2$ sequences into $2^{n\hat{R}}$ bins, hence there are $2^{nR_0}\mathbf{x}_2$ sequences at each bin. Let f be this mapping, that is $z=f(s)$.

Let $R^{'}=\min\{I(X_2;Y),I(X_2;Y_2|X_1)\}+I(X_1;\hat{Y}_1,Y|X_2)-I(X_1,X_2;Y_2)$.

Define $\mathcal{W}=\{1,\dots,2^{nR^{'}}\}$, $\mathcal{L}=\{1,\dots,2^{n(R-R^{'})}\}$ and $\mathcal{K}=\mathcal{W}\times\mathcal{L}=\{1,\dots,2^{nR}\}$.

In the following, we assume that $R-R^{'}>0$. If this is not the case, we set $\mathcal{W}=\{1\}$ and $\mathcal{L}=\{1,\dots,2^{nR}\}$. This NF strategy does not achieve any security level. In this case, we achieve $(R,0)$ which is still inside the region given in this theorem.

2) Encoding: We exploit the block Markov coding scheme.

For a given rate pair (R_1,R_e), where $R_1\le R, R_e\le R_1$, we give the following coding strategy. Let the message to be transmitted at block i be $w_1(i)\in\mathcal{W}_1=[1,M]$, where $M=2^{nR_1}$. We require $R_1\le R$.

The stochastic encoder at the transmitter first forms the following mappings.

- If $R_1>R^{'}$, we let $\mathcal{W}_1=\mathcal{W}\times\mathcal{J}$, where $\mathcal{J}=\{1,2^{n(R_1-R^{'})}\}$. We let g_1 be the partition that partitions $\mathcal{L}$ into $|\mathcal{J}|$ equal size subsets. The stochastic encoder at transmitter will choose a mapping for each message $w_1(i)=(w(i),j(i))\to(w(i),l(i))$, where $l(i)$ is chosen randomly from the set $g_1^{-1}(j(i))\subset\mathcal{L}$ with uniform distribution.
- If $R_1<R^{'}$, the stochastic encoder will choose a mapping $w_1(i)\to(w_1(i),l(i))$, where $l(i)$ is chosen uniformly from the set $\mathcal{L}$.

At first consider block i, where $i\neq1,B$, which means it is not the first or the last block. Assume that the message $w_1(i)$ intended to send at block i is associated with $(w(i),l(i))$ by the stochastic encoder at the transmitter. We let $k(i)=(w(i),l(i))$. Then the encoder at the source sends $\mathbf{x}_1(k(i))$ at block i. At the end of block $i-1$, we assume that $(\mathbf{x}_2(s(i-1)),\hat{\mathbf{y}}_1(z(i-1),s(i-1)),\mathbf{y}_1(i-1))$ are jointly typical,[4] then we choose $s(i)$ uniformly from bin $z(i-1)$, and the relay sends $\mathbf{x}_2(s(i))$ at block i.

When $i=1$, the source sends $\mathbf{x}_1(k(1))$, the relay sends $\mathbf{x}_2(1)$. When $i=B$, the source sends $\mathbf{x}_1(1)$, the relay sends $\mathbf{x}_2(s(B))$.

3) Decoding: First consider the relay node. At the end of block i, the relay already has $s(i)$,[5] it then decides $z(i)$ by choosing $z(i)$ such that $(\mathbf{x}_2(s(i)),\hat{\mathbf{y}}_1(z(i),s(i)),\mathbf{y}_1(i))$ are jointly typical. There exists such $z(i)$, if

$$\hat{R}\ge I(Y_1;\hat{Y}_1|X_2) \tag{50}$$

and n is sufficiently large. Choose $s(i+1)$ uniformly from bin $z(i)$.

The destination does backward decoding. The decoding process starts at the last block B, the destination decodes $s(B)$ by choosing unique $\hat{s}(B)$ such that $(\mathbf{x}_2(\hat{s}(B)),\mathbf{y}(B))$ are jointly typical. We will have $\hat{s}(B)=s(B)$, if

$$R_2\le I(X_2;Y) \tag{51}$$

and n is sufficiently large.

Next, the destination moves to the block $B-1$. Now it already has $s(B)$, hence we also have $z(B-1)=f(s(B))$. It first declares that $\hat{s}(B-1)$ is received, if $\hat{s}(B-1)$ is the unique one such that $(\mathbf{x}_2(\hat{s}(B-1)),\mathbf{y}(B-1))$ are jointly typical. If (51) is satisfied, $\hat{s}(B-1)=s(B-1)$ with high probability. After knowing $\hat{s}(B-1)$, the destination gets an estimation of $\hat{k}(B-1)$, by picking the unique $\hat{k}(B-1)$ such that $(\mathbf{x}_1(\hat{k}(B-1)),\hat{\mathbf{y}}_1(z(B-1),\hat{s}(B-1)),\mathbf{y}(B-1),\mathbf{x}_2(\hat{s}(B-1)))$ are jointly typical. We will have $\hat{k}(B-1)=k(B-1)$ with high probability, if

$$R\le I(X_1;\hat{Y}_1,Y|X_2) \tag{52}$$

and n is sufficiently large.

When the destination moves to block i, the destination has $s(i+1)$ and hence $z(i)=f(s(i+1))$. It first declares that $\hat{s}(i)$ is received, by choosing unique $\hat{s}(i)$ such that $(\mathbf{x}_2(\hat{s}(i)),\mathbf{y}(i))$ are jointly typical. If (51) is satisfied, $\hat{s}(i)=s(i)$ with high probability. After knowing $\hat{s}(i)$, the destination declares that $\hat{k}(i)$ is received, if $\hat{k}(i)$ is the unique one such that $(\mathbf{x}_1(\hat{k}(i)),\hat{\mathbf{y}}_1(z(i),\hat{s}(i)),\ \mathbf{y}(i),\mathbf{x}_2(\hat{s}(i)))$ are jointly typical. If (52) is satisfied, $\hat{k}(i)=k(i)$ with high probability when n is sufficiently large.

Having $\hat{k}(i)$, the destination can get the estimation of the message $w_1(i)$ by letting 1) $\hat{w}_1(i)=(\hat{w}(i),\hat{j}(i))=(\hat{w}(i),g_1(\hat{l}(i)))$, if $R_1>R-R^{'}$, 2) $\hat{w}_1(i)=\hat{w}(i)$, if $R_1<R-R^{'}$. The probability that $\hat{w}_1(i)=w_1(i)$ goes to one for sufficiently large n.

4) Equivocation Computation:

$$\begin{aligned}
&H(W_1|\mathbf{Y}_2)\\
&\quad=H(W_1,\mathbf{Y}_2)-H(\mathbf{Y}_2)\\
&\quad=H(W_1,\mathbf{Y}_2,\mathbf{X}_1,\mathbf{X}_2)\\
&\qquad-H(\mathbf{X}_1,\mathbf{X}_2|W_1,\mathbf{Y}_2)-H(\mathbf{Y}_2)\\
&\quad=H(\mathbf{X}_1,\mathbf{X}_2)+H(W_1,\mathbf{Y}_2|\mathbf{X}_1,\mathbf{X}_2)\\
&\qquad-H(\mathbf{X}_1,\mathbf{X}_2|W_1,\mathbf{Y}_2)-H(\mathbf{Y}_2)
\end{aligned}$$

[4]See the decoding part, such $z(i-1)$ exists.

[5]At the end of block 1, relay knows $s(i)=1$, this is the starting point.

$$\geq H(\mathbf{X}_1, \mathbf{X}_2) + H(\mathbf{Y}_2|\mathbf{X}_1, \mathbf{X}_2) - H(\mathbf{X}_1, \mathbf{X}_2|W_1, \mathbf{Y}_2) - H(\mathbf{Y}_2). \quad (53)$$

Following [2], we will have $H(\mathbf{Y}_2) - H(\mathbf{Y}_2|\mathbf{X}_1, \mathbf{X}_2) \leq nI(X_1, X_2; Y_2) + n\delta_n$, where $\delta_n \to 0$ as $n \to \infty$.

Now let's calculate $H(\mathbf{X}_1, \mathbf{X}_2|W_1, \mathbf{Y}_2)$. Given W_1, the eavesdropper can do joint decoding. It does backward decoding. We pick up the story at block i, we suppose it already decodes $s(i+1)$ and hence $z(i) = f(s(i+1))$. Given W_1, the eavesdropper knows $w(i)$, hence it will decode $l(i)$ and $s(i)$ sent by the relay, by letting $l(i) = \hat{l}(i), s(i) = \hat{s}(i)$, if $\hat{l}(i), \hat{s}(i)$ are the only ones such that $(\mathbf{x}_1(w(i), \hat{l}(i)), \mathbf{x}_2(\hat{s}(i)), \hat{\mathbf{y}}_1(z(i), \hat{s}(i)), \mathbf{y}_2(i))$ are jointly typical. Then, if $R_2 \leq I(X_2; Y_2|X_1)$ and (52) is satisfied, we have

$$\begin{aligned} \frac{1}{2}\log_2(|\mathcal{L}|) + R_2 &= R - R' + R_2 \\ &= R - \min\{I(X_2; Y), I(X_2; Y_2|X_1)\} \\ &\quad - I(X_1; \hat{Y}_1, Y|X_2) \\ &\quad + I(X_1, X_2; Y_2) + \min\{I(X_2; Y), I(X_2; Y_2|X_1)\} \\ &\leq I(X_1, X_2; Y_2). \end{aligned} \quad (54)$$

Also, we have

$$\frac{1}{2}\log_2(|\mathcal{L}|) < R \leq I(X_1; \hat{Y}_1, Y_2|X_2).$$

Thus, we have

$$\Pr\{(\mathbf{X}_1(w(i), \hat{l}(i)), \mathbf{X}_2(\hat{s}(i))) \neq (\mathbf{X}_1(w(i), l(i)), \mathbf{X}_2(s(i)))\} \leq \epsilon_1. \quad (55)$$

Then based on Fano's inequality, we have

$$\frac{1}{n}H(\mathbf{X}_1, \mathbf{X}_2, |W_1 = w_1, \mathbf{Y}_2) \leq \frac{1}{n} + \epsilon_1 I(X_1, X_2; Y_2). \quad (56)$$

Hence, we have

$$\begin{aligned} \frac{1}{n}H(\mathbf{X}_1, \mathbf{X}_2|W_1, \mathbf{Y}_2) &= \frac{1}{n}\sum_{w_1 \in \mathcal{W}_1} p(W_1 = w_1) \\ &\quad \times H(\mathbf{X}_1, \mathbf{X}_2|W_1 = w_1, \mathbf{Y}_2) \leq \epsilon_2 \end{aligned} \quad (57)$$

when n is sufficiently large.

We know $H(\mathbf{X}_1, \mathbf{X}_2) = H(\mathbf{X}_1) + H(\mathbf{X}_2|\mathbf{X}_1) \geq n(R + R_0)$.

If $R_1 > R'$, we have $H(\mathbf{X}_1) = nR$, then we get

$$nR_e = H(W_1|\mathbf{Y}_2) \geq n(R_0 + I(X_1; \hat{Y}_1, Y|X_2) - I(X_1, X_2; Y_2) - \epsilon_4).$$

If $R_1 < R'$, we have $H(\mathbf{X}_1) = n(R_1 + R - R')$, hence

$$nR_e = H(W_1|\mathbf{Y}_2) \geq nR_1 + n(R_0 - \min\{I(X_2; Y), I(X_2; Y_2|X_1))\} - \epsilon_4).$$

The claim is proved.

Appendix E
Proof of Theorem 5

The achievable part follows from the achievability Proof of Theorem 2.

In the following, we modified the converse Proof of Theorem 1 to take the degradedness assumption into consideration. The condition

$$R_e \leq I(V_1, V_2; Y|V_2, U) - I(V_1, V_2; Y_2|U) + \delta_n$$

is the same as Theorem 1.

Meanwhile, from (34), we have

$$\begin{aligned} R_e &\leq \frac{1}{n}\sum_{i=1}^{n}[I(W_1, Y_{2,i+1}^n; Y_i, Y_{1i}|Y^{i-1}, Y_{2,i+1}^n) \\ &\quad - I(W_1, Y^{i-1}, Y_{2,i+1}^n; Y_{2,i}|Y^{i-1}, Y_{2,i+1}^n)] + \delta_n \\ &= \frac{1}{n}\sum_{i=1}^{n}[I(W_1, Y_{2,i+1}^n; Y_{1i}|Y^{i-1}, Y_{2,i+1}^n) \\ &\quad + I(W_1, Y_{2,i+1}^n; Y_i|Y_{1i}, Y^{i-1}, Y_{2,i+1}^n) \\ &\quad - I(W_1, Y^{i-1}, Y_{2,i+1}^n; Y_{2,i}|Y^{i-1}, Y_{2,i+1}^n)] + \delta_n \\ &\stackrel{(a)}{=} \frac{1}{n}\sum_{i=1}^{n}[I(W_1, Y_{2,i+1}^n; Y_{1i}|Y^{i-1}, Y_{2,i+1}^n) \\ &\quad - I(W_1, Y^{i-1}, Y_{2,i+1}^n; Y_{2,i}|Y^{i-1}, Y_{2,i+1}^n)] + \delta_n \\ &= I(V_1; Y_1|V_2, U) - I(V_1, V_2; Y_2|U) + \delta_n. \end{aligned} \quad (58)$$

In (a), $I(W_1, Y_{2,i+1}^n; Y_i|Y_{1i}, Y^{i-1}, Y_{2,i+1}^n) = 0$, because given Y_{1i}, Y_i is independent with everything, according to the physically degradedness definition. Thus, the proof is complete.

Appendix F
Proof of Theorem 6

The proof follows closely with that of Theorem 3. We first consider the case $I(X_1; Y|X_2) < I(X_1; Y_2|X_2)$, i.e., the channel between the source and the eavesdropper is better than the channel between the source and the destination. In this case, we only need to consider the case $\min\{I(X_2; Y), I(X_2; Y_2)\} = I(X_2; Y_2)$, otherwise, the perfect secrecy rate will be zero. Thus, in this case, $R_{s1} = [I(X_1; Y|X_2) + \min\{I(X_2; Y), I(X_2; Y_2|X_1)\} - I(X_1, X_2; Y_2)]^+$.

1) Codebook Generation: For a given distribution $p(x_1)p(x_2)$, we generate at random 2^{nR_2} i.i.d. n-sequence at the relay node each drawn according to $p(\mathbf{x}_2) = \prod_{i=1}^n p(x_{2,i})$, index them as $\mathbf{x}_2(a), a \in [1, 2^{nR_2}]$. Here we set $R_2 = \min\{I(X_2; Y), I(X_2; Y_2|X_1)\} - \epsilon$. We also generate random 2^{nR} i.i.d n-sequence at the source each drawn according to $p(\mathbf{x}_1) = \prod_{i=1}^n p(x_{1,i})$, index them as $\mathbf{x}_1(k), k \in [1, 2^{nR}]$ with $R = I(X_1; Y|X_2) - \epsilon$. Let

$$R_{\min} = \min\{R_{s1}, R_{s2}\}, R_{\max} = \max\{R_{s1}, R_{s2}\}$$

where $R_{s1} = \min\{I(X_2; Y), I(X_2; Y_2|X_1)\} + I(X_1; Y|X_2) - I(X_1, X_2; Y_2)$, $R_{s2} = I(X_1; Y|X_2) - I(X_1; Y_1|X_2)$. We now define $\mathcal{W} = \{1,\ldots,2^{nR_{\min}}\}$, $\mathcal{L}_1 = \{1,\ldots,2^{n(R_{\max}-R_{\min})}\}$, $\mathcal{L}_2 = \{1,\ldots,2^{n(R-R_{\max})}\}$ and $\mathcal{L} = \mathcal{L}_1 \times \mathcal{L}_2$, $\mathcal{K} = \mathcal{W} \times \mathcal{L}$.

2) Encoding: Here, we consider perfect secrecy rate. For a given rate $R_1 \le R_{\min}$, we give the following coding strategy to show that for any given $\epsilon \ge 0$, the equivocation rate at the eavesdropper and the relay node can be made to be larger or equal $R_1 - \epsilon$.

Let the message to be transmitted at block i be $w_1(i) \in \mathcal{W}_1 = [1, M]$, where $M = 2^{nR_1}$. The stochastic encoder will choose a mapping $w_1(i) \to (w_1(i), l_1(i), l_2(i))$, where $l_1(i), l_2(i)$ are chosen uniformly from the set $\mathcal{L}_1, \mathcal{L}_2$, respectively. We write $l(i) = (l_1(i), l_2(i))$.

Suppose the message $w_1(i)$ intended to send at block i is associated with $(w(i), l(i))$ by the stochastic encoder at the transmitter. The encoder then sends $\mathbf{x}_1((w(i), l(i)))$. The relay uniformly picks a code $\mathbf{x}_2(a)$ from $a \in [1,\ldots,2^{nR_2}]$, and sends $\mathbf{x}_2(a)$.

3) Decoding: At the end of block i, the destination declares that $\hat{a}(i)$ is received, if $\hat{a}(i)$ is the only one such that $(\mathbf{x}_2(\hat{a}(i)), \mathbf{y})$ are jointly typical. If there does not exist or there exist more than one such sequences, the destination declares an error. Since $R_2 = \min\{I(X_2;Y), I(X_2;Y_2|X_1)\} - \epsilon \le I(X_2;Y) - \epsilon$, then based on AEP, we know that the error probability will be less than any given positive number ϵ, when the codeword length n is long enough.

The destination then declares that $\hat{k}$ is received, if $\hat{k}$ is the only one such that $(\mathbf{x}_1(\hat{k}), \mathbf{x}_2(\hat{a}), \mathbf{y})$ are jointly typical, otherwise, declares an error. Since $R = I(X_1;Y|X_2) - \epsilon$, then based on AEP, we know that we will have error probability goes to zero, when n is sufficiently large.

Having $\hat{k}(i)$, the destination can get the estimation of the message $w_1(i)$ by letting $\hat{w}_1(i) = \hat{w}(i)$. The probability that $\hat{w}_1(i) = w_1(i)$ goes to one for sufficiently large n.

4) Equivocation Computation: We first calculate the equivocation rate of the eavesdropper when $R_{s1} \le R_{s2}$.

$$\begin{aligned} H(W_1|\mathbf{Y}_2) &= H(W_1,\mathbf{Y}_2) - H(\mathbf{Y}_2) \\ &= H(W_1,\mathbf{Y}_2,\mathbf{X}_1,\mathbf{X}_2) \\ &\quad - H(\mathbf{X}_1,\mathbf{X}_2|W_1,\mathbf{Y}_2) - H(\mathbf{Y}_2) \\ &= H(\mathbf{X}_1,\mathbf{X}_2) + H(W_1,\mathbf{Y}_2|\mathbf{X}_1,\mathbf{X}_2) \\ &\quad - H(\mathbf{X}_1,\mathbf{X}_2|W_1,\mathbf{Y}_2) - H(\mathbf{Y}_2) \\ &\ge H(\mathbf{X}_1,\mathbf{X}_2) + H(\mathbf{Y}_2|\mathbf{X}_1,\mathbf{X}_2) \\ &\quad - H(\mathbf{X}_1,\mathbf{X}_2|W_1,\mathbf{Y}_2) - H(\mathbf{Y}_2). \end{aligned} \tag{59}$$

Now let us calculate $H(\mathbf{X}_1,\mathbf{X}_2|W_1,\mathbf{Y}_2)$. Given W_1, the eavesdropper can do joint decoding. At any block i, given W_1, the eavesdropper knows $w(i)$, hence it will decode $l(i) = (l_1(i), l_2(i))$ and $a(i)$ sent by the relay, by letting $l(i) = \hat{l}(i), a(i) = \hat{a}(i)$, if $\hat{l}(i), \hat{a}(i)$ are the only one pair such that $(\mathbf{x}_1(w(i),\hat{l}(i)), \mathbf{x}_2(\hat{a}(i)), \mathbf{y})$ are jointly typical. Since $R_2 = \min\{I(X_2;Y), I(X_2;Y_2|X_1)\} - \epsilon \le I(X_2;Y_2|X_1) - \epsilon$, we have

$$\begin{aligned} &\frac{1}{2}\log_2(|\mathcal{L}|) + R_2 \\ &\quad = R + I(X_1,X_2;Y_2) - \min\{I(X_2;Y), I(X_2;Y_2|X_1)\} \\ &\qquad - I(X_1;Y|X_2) + \min\{I(X_2;Y), I(X_2;Y_2|X_1)\} - \epsilon \\ &\quad \le I(X_1,X_2;Y_2) - \epsilon. \end{aligned} \tag{60}$$

Also, we have $\frac{1}{2}\log_2(|\mathcal{L}|) < R \le I(X_1;Y_2|X_2)$.

So

$$\begin{aligned} \Pr\{&(\mathbf{X}_1(w(i),\hat{l}(i)), \mathbf{X}_2(\hat{a}(i))) \\ &\neq (\mathbf{X}_1(w(i),l(i)), \mathbf{X}_2(a(i)))\} \le \epsilon_1. \end{aligned}$$

Then based on Fano's inequality, we have

$$\frac{1}{n}H(\mathbf{X}_1,\mathbf{X}_2|W_1=w_1,\mathbf{Y}_2) \le \frac{1}{n} + \epsilon_1 I(X_1,X_2;Y_2). \tag{61}$$

Hence, we have

$$\begin{aligned} &\frac{1}{n}H(\mathbf{X}_1,\mathbf{X}_2|W_1,\mathbf{Y}_2) \\ &\quad = \frac{1}{n}\sum_{w_1\in\mathcal{W}_1} p(W_1=w_1)H(\mathbf{X}_1,\mathbf{X}_2|W_1=w_1,\mathbf{Y}_2) \le \epsilon_2 \end{aligned} \tag{62}$$

when n is sufficiently large.

Now, $H(\mathbf{Y}_2) - H(\mathbf{Y}_2|\mathbf{X}_1,\mathbf{X}_2) \le nI(X_1,X_2;Y_2) + n\delta_n$, where $\delta_n \to 0$, as $n \to \infty$. Also, we have $H(\mathbf{X}_1,\mathbf{X}_2) = H(\mathbf{X}_1) + H(\mathbf{X}_2)$ since $\mathbf{x}_1$ and $\mathbf{x}_2$ are independent. Now $H(\mathbf{X}_1) = R_1 + I(X_1,X_2;Y_2) - \min\{I(X_2;Y), I(X_2;Y_2|X_1)\}$, hence $H(\mathbf{X}_1,\mathbf{X}_2) = R_1 + I(X_1,X_2;Y_2) - \epsilon$.

We get $nR_e = H(W_1|\mathbf{Y}_2) \ge nR_1 - n\epsilon_4$.

Now we calculate the equivocation rate at the relay node

$$\begin{aligned} &H(W_1|\mathbf{Y}_1,\mathbf{X}_2) \\ &\quad \ge H(W_1|\mathbf{Y}_1,\mathbf{X}_2,L_1) \\ &\quad = H(W_1,\mathbf{Y}_1,L_1|\mathbf{X}_2) - H(\mathbf{Y}_1,L_1|\mathbf{X}_2) \\ &\quad = H(W_1,L_1,\mathbf{Y}_1,\mathbf{X}_1|\mathbf{X}_2) \\ &\qquad - H(\mathbf{X}_1|W_1,L_1,\mathbf{Y}_1,\mathbf{X}_2) - H(\mathbf{Y}_1,L_1|\mathbf{X}_2) \\ &\quad = H(\mathbf{X}_1|\mathbf{X}_2) + H(W_1,L_1,\mathbf{Y}_1|\mathbf{X}_1,\mathbf{X}_2) \\ &\qquad - H(\mathbf{X}_1|W_1,L_1,\mathbf{Y}_1,\mathbf{X}_2) - H(\mathbf{Y}_1,L_1|\mathbf{X}_2) \\ &\quad \overset{(a)}{\ge} H(\mathbf{X}_1) + H(\mathbf{Y}_1|\mathbf{X}_1,\mathbf{X}_2) \\ &\qquad - H(\mathbf{X}_1|W_1,L_1,\mathbf{Y}_1,\mathbf{X}_2) - H(L_1) - H(\mathbf{Y}_1|\mathbf{X}_2) \end{aligned}$$

where the first term of (a) comes from the fact that $\mathbf{x}_1, \mathbf{x}_2$ are independent, and the fourth term comes from the fact that $l_1, \mathbf{x}_2$ are independent.

Now, $H(L_1) = n(R_{\max} - R_{\min})$, $H(\mathbf{Y}_1|\mathbf{X}_1,\mathbf{X}_2) - H(\mathbf{Y}_1|\mathbf{X}_2) \le nI(X_1;Y_1|X_2) + n\delta_n$. Given $w_1, l_1, \mathbf{x}_2$, the relay can just choose the $\mathbf{x}_1$ in the bin (w_1, l_1) which is jointly typical with $\mathbf{x}_2, \mathbf{y}_1$. Since $\frac{1}{n}\log_2(|\mathcal{L}_2|) \le I(X_1;Y_1|X_2)$, we have $\Pr\{\hat{\mathbf{X}}_1 \neq \mathbf{X}_1\} \le \epsilon_2$.

Then based on Fano's inequality, we have

$$\begin{aligned} &\frac{1}{n}H(\mathbf{X}_1|W_1=w_1,L_1=l_1,\mathbf{Y}_1,\mathbf{X}_2=\mathbf{x}_2) \\ &\qquad \le \frac{1}{n} + \epsilon_1 I(X_1;Y_1|X_2) \end{aligned} \tag{63}$$

Hence, we have

$$\begin{aligned} &\frac{1}{n}H(\mathbf{X}_1|W_1,L_1,\mathbf{Y}_1,\mathbf{X}_2) \\ &\qquad = \frac{1}{n}\sum_{w_1,l_1,\mathbf{x}_2} p(W_1=w_1,L_1=l_1,\mathbf{x}_2) \\ &\qquad \times H(\mathbf{X}_1,\mathbf{X}_2|W_1=w_1,L_1=l_i,\mathbf{x}_2,\mathbf{Y}_1) \le \epsilon_2 \end{aligned} \tag{64}$$

when n is sufficiently large.

Also, based on the encoding part, we have $H(\mathbf{X}_1) = n(R_1 + I(X_1;Y|X_2) - R_{\min})$.

Combining these, we get

$$\begin{aligned} H(W_1|\mathbf{Y}_1,\mathbf{X}_2) &\geq n(R_1 + I(X_1;Y|X_2) \\ &\quad - R_{\min} - (R_{\max} - R_{\min}) \\ &\quad - I(X_1;Y_1|X_2) - \delta_n) \\ &= n(R_1 - \delta_n). \end{aligned} \tag{65}$$

The equivocation rate of the relay and the eavesdropper when $R_{1s} \geq R_{2s}$ can be calculated similarly, with the only difference that we bound the equivocation rate of the eavesdropper by giving it L_1. This case is proved.

Now, consider the case $I(X_1;Y|X_2) > I(X_1;Y_2|X_2)$. If $\min\{I(X_2;Y), I(X_2;Y_2)\} = I(X_2;Y)$, then we have $\min\{I(X_2;Y), I(X_2;Y_2|X_1)\} = I(X_2;Y)$, because $I(X_2;Y_2|X_1) > I(X_2;Y_2)$ since X_1, X_2 are independent. Under this case, we only need to prove the case $R_{s1} = [I(X_1;Y|X_2) - I(X_1;Y_2|X_2)]^+$, which can be achieved by using a scheme similar to the one developed in proving (49). If $\min\{I(X_2;Y),\ I(X_2;Y_2)\} = I(X_2;Y_2)$, and we only need to consider $R_{s1} = [I(X_1;Y|X_2) + \min\{I(X_2;Y), I(X_2;Y_2|X_1)\} - I(X_1,X_2;Y_2)]^+$, then we can use a coding/decoding scheme similar to the one developed above to show the achievability.

The claim is achieved.

Acknowledgment

The authors would like to thank Dr. Yingbin Liang of the University of Hawaii for helpful discussions. The authors would also like to thank Onur Ozan Koyluoglu of Ohio State University for an insight that improved the paper.

References

[1] C. E. Shannon, "Communication theory of secrecy systems," *Bell Syst. Tech. J.*, vol. 28, pp. 656–715, Oct. 1949.

[2] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, 1975.

[3] I. Csiszar and J. Korner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 24, pp. 339–348, May 1978.

[4] T. Cover and A. E. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, pp. 572–584, Sep. 1979.

[5] E. C. van der Meulen, "Three-terminal communication channels," *Adv. Appl. Probab.*, vol. 3, pp. 120–154, 1971.

[6] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity- part I: System description," *IEEE Trans. Commun.*, vol. 51, pp. 1927–1938, Nov. 2003.

[7] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, pp. 3037–3063, Sep. 2005.

[8] L. L. Xie and P. R. Kumar, "An achievable rate for the multiple-level relay channel," *IEEE Trans. Inf. Theory*, vol. 51, pp. 1348–1358, Apr. 2005.

[9] L. Lai, K. Liu, and H. El-Gamal, "The three node wireless network: Achievable rates and cooperation strategies," *IEEE Trans. Inf. Theory*, vol. 52, pp. 805–828, Mar. 2006.

[10] Y. Liang and V. V. Veeravalli, "Cooperative relay broadcast channels," *IEEE Trans. Inf. Theory*, vol. 53, pp. 900–928, Mar. 2007.

[11] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, pp. 3062–3080, Dec. 2004.

[12] K. Azarian, H. El-Gamal, and P. Schniter, "On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels," *IEEE Trans. Inf. Theory*, vol. 51, pp. 4152–4172, Dec. 2005.

[13] Y. Oohama, "Coding for relay channels with confidential messages," in *Proc. IEEE Inf. Theory Workshop*, Cairns, Australia, Sep. 2-7, 2001, pp. 87–89.

[14] Y. Oohama, "Relay channels with confidential messages," *IEEE Trans. Inf. Theory*, Nov. 2006, Submitted.

[15] I. Csiszar and P. Narayan, "Common randomness and secret key generation with a helper," *IEEE Trans. Inf. Theory*, vol. 24, pp. 339–348, May 1978.

[16] R. Liu, I. Maric, R. D. Yates, and P. Spasojevic, "The discrete memoryless multiple access channel with confidential messages," in *Proc. IEEE Int. Symp. Inf. Theory*, Seattle, WA, Jul. 9-14, 2006.

[17] Y. Liang and H. V. Poor, "Multiple access channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 54, pp. 976–1002, Mar. 2008.

[18] E. Tekin and A. Yener, "The gaussian multiple access wire-tap channel," *IEEE Trans. Inf. Theory*, 2006, submitted for publication.

[19] A. O. Hero, "Secure space-time communication," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1–16, Dec. 2003.

[20] C.-M. Zing, F. Kuhlmann, and A. Buzo, "Achivability proof of some multiuser channel coding theorems using backward decoding," *IEEE Trans. Inf. Theory*, vol. 35, no. 6, pp. 1160–1165, 1989.

[21] M. A. Khojastepour, A. Sabharwal, and B. Aazhang, "Cut-set theorems for multi-state networks," *IEEE Trans. Inf. Theory*, submitted for publication.

[22] S. K. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian wiretap channel," *IEEE Trans. Inf. Theory*, vol. 24, pp. 451–456, July 1978.

[23] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.