# Multiple-Access Relay Wiretap Channel

Bin Dai and Zheng Ma

#### Abstract

In this paper, we investigate the effects of an additional trusted relay node on the secrecy of multiple-access wiretap channel (MAC-WT) by considering the model of multiple-access relay wiretap channel (MARC-WT). More specifically, first, we investigate the discrete memoryless MARC-WT. Three inner bounds (with respect to decode-forward (DF), noise-forward (NF) and compress-forward (CF) strategies) on the secrecy capacity region are provided. Second, we investigate the degraded discrete memoryless MARC-WT, and present an outer bound on the secrecy capacity region of this degraded model. Finally, we investigate the Gaussian MARC-WT, and find that the NF and CF strategies help to enhance Tekin-Yener's achievable secrecy rate region of Gaussian MAC-WT. Moreover, we find that if the channel from the transmitters to the relay is less noisy than the channels from the transmitters to the legitimate receiver and the wiretapper, the achievable secrecy rate region of the DF strategy is even larger than the corresponding regions of the NF and CF strategies.

## Index Terms

Multiple-access wiretap channel, relay channel, secrecy capacity region.

# I. INTRODUCTION

Equivocation was first introduced into channel coding by Wyner in his study of wiretap channel [1]. It is a kind of discrete memoryless degraded broadcast channels. The objective is to transmit messages to the legitimate receiver, while keeping the wiretapper as ignorant of the messages as possible. Based on Wyners work, Leung-Yan-Cheong and Hellman studied the Gaussian wiretap channel (GWC) [2], and showed that its secrecy capacity was the difference between the main channel capacity and the overall wiretap channel capacity (the cascade of main channel and wiretap channel).

After the publication of Wyner's work, Csiszár and Körner [3] investigated a more general situation: the broadcast channels with confidential messages (BCC). In this model, a common message and a confidential message were sent through a general broadcast channel. The common message was assumed to be decoded correctly by the legitimate receiver and the wiretapper, while the confidential message was only allowed to be obtained by the legitimate receiver. This model is also a generalization of [4], where no confidentiality condition is imposed. The capacity-equivocation region and the secrecy capacity region of BCC [3] were totally determined, and the results

B. Dai and Z. Ma are with the School of Information Science and Technology, Southwest JiaoTong University, Chengdu 610031, China e-mail: daibin@home.swjtu.edu.cn, zma@home.swjtu.edu.cn.

were also a generalization of those in [1]. Furthermore, the capacity-equivocation region of Gaussian BCC was determined in [22].

By using the approach of [1] and [3], the information-theoretic security for other multi-user communication systems has been widely studied, see the followings.

- For the broadcast channel, Liu et al. [5] studied the broadcast channel with two confidential messages (no common message), and provided an inner bound on the secrecy capacity region. Furthermore, Xu et al. [6] studied the broadcast channel with two confidential messages and one common message, and provided inner and outer bounds on the capacity-equivocation region.
- For the multiple-access channel (MAC), the security problems are split into two directions.
  - The first is that two users wish to transmit their corresponding messages to a destination, and meanwhile, they also receive the channel output. Each user treats the other user as a wiretapper, and wishes to keep its confidential message as secret as possible from the wiretapper. This model is usually called the MAC with confidential messages, and it was studied by Liang and Poor [7]. An inner bound on the capacity-equivocation region is provided for the model with two confidential messages, and the capacity-equivocation region is still not known. Furthermore, for the model of MAC with one confidential message [7], both inner and outer bounds on capacity-equivocation region are derived. Moreover, for the degraded MAC with one confidential message, the capacity-equivocation region is totally determined.
  - The second is that an additional wiretapper has access to the MAC output via a wiretap channel, and therefore, how to keep the confidential messages of the two users as secret as possible from the additional wiretapper is the main concern of the system designer. This model is usually called the multiple-access wiretap channel (MAC-WT). The Gaussian MAC-WT was investigated in [8], [9]. An inner bound on the capacity-equivocation region is provided for the Gaussian MAC-WT. Other related works on MAC-WT can be found in [10], [11], [12], [13], [14], [15], [16].
- For the interference channel, Liu et al. [5] studied the interference channel with two confidential messages, and provided inner and outer bounds on the secrecy capacity region. In addition, Liang et al. [17] studied the cognitive interference channel with one common message and one confidential message, and the capacity-equivocation region was totally determined for this model.
- For the relay channel, Lai and Gamal [18] studied the relay-eavesdropper channel, where a source wishes to send messages to a destination while leveraging the help of a trusted relay node to hide those messages from the eavesdropper. Three inner bounds (with respect to decode-forward, noise-forward and compress-forward strategies) and one outer bound on the capacity-equivocation region were provided in [18]. Furthermore, Tang et. al. [27] introduced the noise-forward strategy of [18] into the wireless communication networks, and found that with the help of an independent interferer, the security of the wireless communication networks is enhanced. In addition, Oohama [19] studied the relay channel with confidential messages, where a relay helps the transmission of messages from one sender to one receiver. The relay is considered not only as a

sender that helps the message transmission but also as a wiretapper who can obtain some knowledge about the transmitted messages. Measuring the uncertainty of the relay by equivocation, the inner and outer bounds on the capacity-equivocation region were provided in [19].

Recently, Ekrem and Ulukus [20] investigated the effects of user cooperation on the secrecy of broadcast channels by considering a cooperative relay broadcast channel. They showed that user cooperation can increase the achievable secrecy rate region of [5].

In this paper, we study the multiple-access relay wiretap channel (MARC-WT), see Figure 1. This model generalizes the MAC-WT by considering an additional trusted relay node. The motivation of this work is to investigate the effects of the trusted relay node on the secrecy of MAC-WT, and whether the achievable secrecy rate region of [9] can be enhanced by using an additional relay node.



Fig. 1: The multiple-access relay wiretap channel

First, we provide three inner bounds on the secrecy capacity region (achievable secrecy rate regions) of the discrete memoryless model of Figure 1. The decode-forward (DF), noise-forward (NF) and compress-forward (CF) relay strategies are used in the construction of the inner bounds. Second, we investigate the degraded discrete memoryless MARC-WT, and present an outer bound on the secrecy capacity region of this degraded case. Finally, the Gaussian model of Figure 1 is investigated, and we find that with the help of this additional trusted relay node, Tekin-Yeners achievable secrecy rate region of the Gaussian MAC-WT [9] is enhanced.

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 Nth 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 2.

The organization of this paper is as follows. Section II provides the achievable secrecy rate regions of the discrete memoryless model of Figure 1. The Gaussian model of Figure 1 is investigated in Section III. Final conclusions are provided in Section IV.

#### II. DISCRETE MEMORYLESS MULTIPLE-ACCESS RELAY WIRETAP CHANNEL

#### A. Inner bounds on the secrecy capacity region of the discrete memoryless MARC-WT

The discrete memoryless model of Figure 1 is a five-terminal discrete channel consisting of finite sets  $\mathcal{X}_1$ ,  $\mathcal{X}_2$ ,  $\mathcal{X}_r$ ,  $\mathcal{Y}$ ,  $\mathcal{Y}_r$ ,  $\mathcal{Z}$  and a transition probability distribution  $P_{Y,Y_r,Z|X_1,X_2,X_r}(y, y_r, z|x_1, x_2, x_r)$ .  $X_1^N$ ,  $X_2^N$  and  $X_r^N$  are the channel inputs from the transmitters and the relay respectively, while  $Y^N$ ,  $Y_r^N$ ,  $Z^N$  are the channel outputs at the legitimate receiver, the relay and the wiretapper, respectively. The channel is discrete memoryless, i.e., the channel outputs  $(y_i, y_{r,i}, z_i)$  at time *i* only depend on the channel inputs  $(x_{1,i}, x_{2,i}, x_{r,i})$  at time *i*.

Definition 1: (Channel encoders) The confidential messages  $W_1$  and  $W_2$  take values in  $W_1$ ,  $W_2$ , respectively.  $W_1$  and  $W_2$  are independent and uniformly distributed over their ranges. The channel encoders  $f_{E1}$  and  $f_{E2}$  are stochastic encoders that map the messages  $w_1$  and  $w_2$  into the codewords  $x_1^N \in \mathcal{X}_1^N$  and  $x_2^N \in \mathcal{X}_2^N$ , respectively. The transmission rates of the confidential messages  $W_1$  and  $W_2$  are  $\frac{\log ||W_1||}{N}$  and  $\frac{\log ||W_2||}{N}$ , respectively.

Definition 2: (**Relay encoder**) The relay encoder  $\varphi_i$  is also a stochastic encoder that maps the signals  $(y_{r,1}, y_{r,2}, ..., y_{r,i-1})$  received before time *i* to the channel input  $x_{r,i}$ .

Definition 3: (**Decoder**) The decoder for the legitimate receiver is a mapping  $f_D : \mathcal{Y}^N \to \mathcal{W}_1 \times \mathcal{W}_2$ , with input  $Y^N$  and outputs  $\hat{W}_1$ ,  $\hat{W}_2$ . Let  $P_e$  be the error probability of the legitimate receiver, and it is defined as  $Pr\{(W_1, W_2) \neq (\hat{W}_1, \hat{W}_2)\}.$ 

The equivocation rate at the wiretapper is defined as

$$\Delta = \frac{1}{N} H(W_1, W_2 | Z^N).$$
(2.1)

A rate pair  $(R_1, R_2)$  (where  $R_1, R_2 \ge 0$ ) is called achievable with perfect secrecy if, for any  $\epsilon > 0$  (where  $\epsilon$  is an arbitrary small positive real number), there exists a sequence of codes  $(2^{NR_1}, 2^{NR_2}, N)$  such that

$$\frac{\log \parallel \mathcal{W}_1 \parallel}{N} = R_1, \frac{\log \parallel \mathcal{W}_2 \parallel}{N} = R_2,$$
  
$$\Delta \ge R_1 + R_2 - \epsilon, \quad P_e \le \epsilon.$$
(2.2)

Note that the above secrecy requirement on the full message set also ensures the secrecy of individual message, i.e.,  $\frac{1}{N}H(W_1, W_2|Z^N) \ge R_1 + R_2 - \epsilon$  implies that  $\frac{1}{N}H(W_t|Z^N) \ge R_t - \epsilon$  for t = 1, 2, and the proof is as follows.

Proof: Since

$$0 \ge R_1 + R_2 - \epsilon - \frac{1}{N} H(W_1, W_2 | Z^N) = \frac{1}{N} H(W_1) + \frac{1}{N} H(W_2) - \frac{1}{N} H(W_1, W_2 | Z^N) - \epsilon$$
  

$$= \frac{1}{N} H(W_1) + \frac{1}{N} H(W_2) - \frac{1}{N} H(W_1 | Z^N) - \frac{1}{N} H(W_2 | W_1, Z^N) - \epsilon$$
  

$$\ge \frac{1}{N} H(W_1) + \frac{1}{N} H(W_2) - \frac{1}{N} H(W_1 | Z^N) - \frac{1}{N} H(W_2 | Z^N) - \epsilon$$
  

$$= \frac{1}{N} I(W_1; Z^N) + \frac{1}{N} I(W_2; Z^N) - \epsilon,$$
(2.3)

and  $\frac{1}{N}I(W_1; Z^N) \ge 0$ ,  $\frac{1}{N}I(W_2; Z^N) \ge 0$ , it is easy to see that  $\frac{1}{N}I(W_1; Z^N) \le \epsilon$ ,  $\frac{1}{N}I(W_2; Z^N) \le \epsilon$ , which implies that  $\frac{1}{N}H(W_t|Z^N) \ge R_t - \epsilon$  for t = 1, 2. The proof is completed.

The secrecy capacity region  $\mathcal{R}^d$  is a set composed of all achievable secrecy rate pairs  $(R_1, R_2)$ . Three inner bounds (with respect to DF, NF and CF strategies) on  $\mathcal{R}^d$  are provided in the following Theorem 1, 2, 3.

Our first step is to characterize the inner bound on the secrecy capacity region  $\mathcal{R}^d$  by using Cover-El Gamal's Decode and Forward (DF) Strategy [23]. In the DF Strategy, the relay node will first decode the confidential messages, and then re-encode them to cooperate with the transmitters. The superposition coding and random binning techniques will be combined with the classical DF strategy [23] to characterize the DF inner bound of the discrete memoryless MARC-WT. The following Theorem 1 shows the DF inner bound on  $\mathcal{R}^d$ .

*Theorem 1:* (Inner bound 1: DF strategy) A single-letter characterization of the region  $\mathcal{R}^{d_1}$  ( $\mathcal{R}^{d_1} \subseteq \mathcal{R}^d$ ) is as follows,

$$\begin{aligned} \mathcal{R}^{d1} &= \{ (R_1, R_2) : R_1, R_2 \ge 0, \\ R_1 \le \min\{I(X_1; Y_r | X_r, X_2, V_1, V_2), I(X_1, X_r; Y | X_2, V_2)\} - I(X_1; Z), \\ R_2 \le \min\{I(X_2; Y_r | X_r, X_1, V_1, V_2), I(X_2, X_r; Y | X_1, V_1)\} - I(X_2; Z), \\ R_1 + R_2 \le \min\{I(X_1, X_2; Y_r | X_r, V_1, V_2), I(X_1, X_2, X_r; Y)\} - I(X_1, X_2; Z) \}. \end{aligned}$$

for some distribution

$$P_{Y,Z,Y_r,X_r,X_1,X_2,V_1,V_2}(y,z,y_r,x_r,x_1,x_2,v_1,v_2) = P_{Y,Z,Y_r|X_r,X_1,X_2}(y,z,y_r|x_r,x_1,x_2)P_{X_r|V_1,V_2}(x_r|v_1,v_2)P_{X_1|V_1}(x_1|v_1)P_{X_2|V_2}(x_2|v_2)P_{V_1}(v_1)P_{V_2}(v_2)$$

Proof:

The achievable coding scheme is a combination of [26], [21] and [9], and the details about the proof are provided in Appendix A.

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

• If we let Z = const (which implies that there is no wiretapper), the region  $\mathcal{R}^{d1}$  reduces to the following achievable region  $\mathcal{R}^{marc}$ , where

$$\mathcal{R}^{marc} = \{ (R_1, R_2) : R_1, R_2 \ge 0, \\ R_1 \le \min\{I(X_1; Y_r | X_r, X_2, V_1, V_2), I(X_1, X_r; Y | X_2, V_2)\}, \\ R_2 \le \min\{I(X_2; Y_r | X_r, X_1, V_1, V_2), I(X_2, X_r; Y | X_1, V_1)\}, \\ R_1 + R_2 \le \min\{I(X_1, X_2; Y_r | X_r, V_1, V_2), I(X_1, X_2, X_r; Y)\}\}.$$

$$(2.4)$$

Here note that the achievable region  $\mathcal{R}^{marc}$  is exactly the same as the achievable DF region (DF inner bound on the capacity region) of the discrete memoryless multiple-access relay channel [26], [21].

• If we let  $Y_r = Y$  and  $V_1 = V_2 = X_r = const$  (which implies that there is no relay), the region  $\mathcal{R}^{d1}$  reduces

`

١

to the region  $\mathcal{R}^{mac-wt}$ , where

$$\mathcal{R}^{mac-wt} = \{ (R_1, R_2) : R_1, R_2 \ge 0, \\ R_1 \le I(X_1; Y | X_2) - I(X_1; Z), \\ R_2 \le I(X_2; Y | X_1) - I(X_2; Z), \\ R_1 + R_2 \le I(X_1, X_2; Y) - I(X_1, X_2; Z) \}.$$
(2.5)

Also note that the region  $\mathcal{R}^{mac-wt}$  is exactly the same as the achievable secrecy rate region of discrete memoryless multiple-access wiretap channel [9].

The second step is to characterize the inner bound on the secrecy capacity region  $\mathcal{R}^d$  by using the noise and forward (NF) strategy. In the NF Strategy, the relay node does not attempt to decode the messages but sends sequences that are independent of the transmitters' messages, and these sequences aid in confusing the wiretapper.

More specifically, for a given input distribution of the relay, if the corresponding mutual information with the legitimate receiver's output is not less than that with the wiretapper's output, we allow the legitimate receiver to decode the sequence of the relay, and the wiretapper can not decode it. Therefore, in this case, the sequence of the relay can be viewed as a noise signal to confuse the wiretapper.

On the other hand, if the corresponding mutual information with the legitimate receiver's output is not more than that with the wiretapper's output, we allow both the receivers to decode the sequence of the relay. In this case, the sequence of the relay does not make any contribution to the security of the discrete memoryless MARC-WT.

The following Theorem 2 shows the NF inner bound on  $\mathcal{R}^d$ .

*Theorem 2:* (Inner bound 2: NF strategy) A single-letter characterization of the region  $\mathcal{R}^{d_2}$  ( $\mathcal{R}^{d_2} \subseteq \mathcal{R}^d$ ) is as follows,

$$\mathcal{R}^{d2} =$$
convex closure of  $(\mathcal{L}^1 \bigcup \mathcal{L}^2),$ 

where  $\mathcal{L}^1$  is given by

$$\mathcal{L}^{1} = \bigcup_{\substack{P_{Y,Z,Y_{r},X_{r},X_{1},X_{2}:\\I(X_{r};Y) \ge I(X_{r};Z)}}} \left\{ \begin{array}{l} (R_{1},R_{2}):R_{1},R_{2} \ge 0,\\ R_{1} \le I(X_{1};Y|X_{2},X_{r}) - I(X_{1},X_{r};Z) + R_{r},\\ R_{2} \le I(X_{2};Y|X_{1},X_{r}) - I(X_{2},X_{r};Z) + R_{r},\\ R_{1} + R_{2} \le I(X_{1},X_{2};Y|X_{r}) - I(X_{1},X_{2},X_{r};Z) + R_{r}. \end{array} \right\},$$

 $R_r$  denotes

 $R_r = \min\{I(X_r; Y), I(X_r; Z|X_1), I(X_r; Z|X_2)\},\$ 

and  $\mathcal{L}^2$  is given by

$$\mathcal{L}^{2} = \bigcup_{\substack{P_{Y,Z,Y_{r},X_{r},X_{1},X_{2}:\\I(X_{r};Z) \geq I(X_{r};Y)}}} \left\{ \begin{array}{l} (R_{1},R_{2}):R_{1},R_{2} \geq 0,\\ R_{1} \leq I(X_{1};Y|X_{2},X_{r}) - I(X_{1};Z|X_{r}),\\ R_{2} \leq I(X_{2};Y|X_{1},X_{r}) - I(X_{2};Z|X_{r}),\\ R_{1} + R_{2} \leq I(X_{1},X_{2};Y|X_{r}) - I(X_{1},X_{2};Z|X_{r}). \end{array} \right\}$$

here the joint probability  $P_{Y,Z,Y_r,X_r,X_1,X_2}(y,z,y_r,x_r,x_1,x_2,u)$  satisfies

$$P_{Y,Z,Y_r,X_r,X_1,X_2}(y,z,y_r,x_r,x_1,x_2) = P_{Y,Z,Y_r|X_r,X_1,X_2}(y,z,y_r|x_r,x_1,x_2)P_{X_r}(x_r)P_{X_1}(x_1)P_{X_2}(x_2)$$

Proof:

The achievable coding scheme is a combination of [18, Theorem 3] and [9], and the details about the proof are provided in Appendix B.

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

- Since the two regions  $\mathcal{L}^1$  and  $\mathcal{L}^2$  are not necessarily contained by one another, by using time-sharing arguments, it is easy to find a new achievable region which is the convex-closure of the union of the two regions.
- The region L<sup>1</sup> is characterized under the condition that for a given input distribution of the relay, the corresponding mutual information with the legitimate receiver's output is not less than that with the wire-tapper's output (I(X<sub>r</sub>; Y) ≥ I(X<sub>r</sub>; Z)). Then, in this case, the legitimate receiver is allowed to decode the sequence of the relay, and the wiretapper is not allowed to decode it. The rate of the sequence is defined as R<sub>r</sub> = min{I(X<sub>r</sub>; Y), I(X<sub>r</sub>; Z|X<sub>1</sub>), I(X<sub>r</sub>; Z|X<sub>2</sub>)}, and the sequence is viewed as pure noise for the wiretapper.
- The region L<sup>2</sup> is characterized under the condition that for a given input distribution of the relay, the corresponding mutual information with the legitimate receiver's output is not more than that with the wiretapper's output (I(X<sub>r</sub>; Y) ≤ I(X<sub>r</sub>; Z)). Then, in this case, both the legitimate receiver and the wiretapper are allowed to decode the sequence of the relay. The rate of the sequence is defined as R<sub>r</sub> = I(X<sub>r</sub>; Y), and the sequence does not make any contribution to the security of the discrete memoryless MARC-WT.

The third step is to characterize the inner bound on the secrecy capacity region  $\mathcal{R}^d$  by using a combination of Cover- El Gamals compress and forward (CF) strategy [23] and the NF strategy provided in Theorem 2, i.e., in addition to the independent codewords, the relay also sends a quantized version of its noisy observations to the legitimate receiver. This noisy version of the relay's observations helps the legitimate receiver in decoding the transmitters' messages, while the independent codewords help in confusing the wiretapper. The following Theorem 3 shows the CF inner bound on  $\mathcal{R}^d$ .

*Theorem 3:* (Inner bound 3: CF strategy) A single-letter characterization of the region  $\mathcal{R}^{d3}$  ( $\mathcal{R}^{d3} \subseteq \mathcal{R}^{d}$ ) is as follows,

$$\mathcal{R}^{d3} = \text{convex closure of} \quad (\mathcal{L}^3 \bigcup \mathcal{L}^4),$$

where  $\mathcal{L}^3$  is given by

$$\mathcal{L}^{3} = \bigcup_{\substack{P_{Y,Z,Y_{r},\hat{Y}_{r},X_{r},X_{1},X_{2}: I(X_{r};Y) \geq I(X_{r};Z) \\ R_{r1}^{*} - R^{*} \geq I(Y_{r};\hat{Y}_{r}|X_{r})}} \left\{ \begin{array}{l} (R_{1},R_{2}):R_{1},R_{2} \geq 0, \\ R_{1} \leq I(X_{1};Y,\hat{Y}_{r}|X_{2},X_{r}) - I(X_{1},X_{r};Z) + R^{*}, \\ R_{2} \leq I(X_{2};Y,\hat{Y}_{r}|X_{1},X_{r}) - I(X_{2},X_{r};Z) + R^{*}, \\ R_{1} + R_{2} \leq I(X_{1},X_{2};Y,\hat{Y}_{r}|X_{r}) - I(X_{1},X_{2},X_{r};Z) + R^{*}. \end{array} \right\}$$

 $R_{r1}^* = \min\{I(X_r; Z|X_1), I(X_r; Z|X_2), I(X_r; Y)\}, R^*$  is the rate of pure noise generated by the relay to confuse the wiretapper,  $R_{r1}^* - R^*$  is the part of the rate allocated to send the compressed signal  $\hat{Y}_r$  to help the legitimate receiver, and  $\mathcal{L}^4$  is given by

$$\mathcal{L}^{4} = \bigcup_{\substack{P_{Y,Z,Y_{r},\hat{Y}_{r},X_{T},X_{1},X_{2} : I(X_{r};Z) \geq I(X_{r};Y) \\ I(X_{r};Y) \geq I(Y_{r};\hat{Y}_{r}|X_{r})}} \left\{ \begin{array}{l} (R_{1},R_{2}) : R_{1},R_{2} \geq 0, \\ R_{1} \leq I(X_{1};Y,\hat{Y}_{r}|X_{2},X_{r}) - I(X_{1};Z|X_{r}), \\ R_{2} \leq I(X_{2};Y,\hat{Y}_{r}|X_{1},X_{r}) - I(X_{2};Z|X_{r}), \\ R_{1} + R_{2} \leq I(X_{1},X_{2};Y,\hat{Y}_{r}|X_{r}) - I(X_{1},X_{2};Z|X_{r}). \end{array} \right\}.$$

The joint probability  $P_{Y,Z,Y_r,\hat{Y}_r,X_r,X_1,X_2}(y,z,y_r,\hat{y}_r,x_r,x_1,x_2)$  satisfies

$$\begin{split} P_{Y,Z,Y_r,\hat{Y}_r,X_r,X_1,X_2}(y,z,y_r,\hat{y}_r,x_r,x_1,x_2) = \\ P_{\hat{Y}_r|Y_r,X_r}(\hat{y}_r|y_r,x_r)P_{Y,Z,Y_r|X_r,X_1,X_2}(y,z,y_r|x_r,x_1,x_2)P_{X_r}(x_r)P_{X_1}(x_1)P_{X_2}(x_2) \end{split}$$

Proof:

The achievable coding scheme is a combination of [18, Theorem 4] and [9], and the details about the proof are provided in Appendix C.

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

- Since the two regions  $\mathcal{L}^3$  and  $\mathcal{L}^4$  are not necessarily contained by one another, by using time-sharing arguments, it is easy to find a new achievable region which is the convex-closure of the union of the two regions.
- The region L<sup>3</sup> is characterized under the condition that for a given input distribution of the relay, the corresponding mutual information with the legitimate receiver's output is not less than that with the wiretapper's output (I(X<sub>r</sub>; Y) ≥ I(X<sub>r</sub>; Z)). Then, in this case, the legitimate receiver is allowed to decode the sequence of the relay, and the wiretapper is not allowed to decode it. Here note that if R<sup>\*</sup> = R<sup>\*</sup><sub>r1</sub>, this scheme is exactly the same as the NF scheme.
- The region L<sup>4</sup> is characterized under the condition that for a given input distribution of the relay, the corresponding mutual information with the legitimate receiver's output is not more than that with the wiretapper's output (I(X<sub>r</sub>; Y) ≤ I(X<sub>r</sub>; Z)). Then, in this case, both the legitimate receiver and the wiretapper are allowed to decode the sequence of the relay. However, the relay can still help to enhance the security of the discrete memoryless MARC-WT by sending the compressed signal Ŷ<sub>r</sub> to the legitimate receiver.

## B. Outer bound on the secrecy capacity region of the degraded discrete memoryless MARC-WT

Compared with the discrete memoryless MARC-WT (see Figure 1), the degraded case implies the existence of a Markov chain  $(X_1, X_2, X_r, Y_r) \rightarrow Y \rightarrow Z$ . The secrecy capacity region  $\mathcal{R}^{dd}$  of the degraded discrete memoryless MARC-WT is a set composed of all achievable secrecy rate pairs  $(R_1, R_2)$ . An outer bound on  $\mathcal{R}^{dd}$  is provided in the following Theorem 4.

*Theorem 4:* (Outer bound) A single-letter characterization of the region  $\mathcal{R}^{ddo}$  ( $\mathcal{R}^{dd} \subseteq \mathcal{R}^{ddo}$ ) is as follows,

$$\mathcal{R}^{ddo} = \{ (R_1, R_2) : R_1, R_2 \ge 0, \\ R_1 \le I(X_1, X_r; Y | X_2, U) - I(X_1; Z | U) \\ R_2 \le I(X_2, X_r; Y | X_1, U) - I(X_2; Z | U) \\ R_1 + R_2 \le I(X_1, X_2, X_r; Y | U) - I(X_1, X_2; Z | U) \}$$

for some distribution

$$P_{Z,Y,Y_r,X_r,X_1,X_2,U}(z,y,y_r,x_r,x_1,x_2,u) =$$

$$P_{Z|Y}(z|y)P_{Y,Y_r|X_1,X_2,X_r}(y,y_r|x_1,x_2,x_r)P_{U,X_1,X_2,X_r}(u,x_1,x_2,x_r)$$

Proof:

The details about the proof are provided in Appendix D.

*Remark 4:* The outer bound on the secrecy capacity region of the degraded discrete memoryless MARC-WT is generally loose, but it is still useful for the analysis of the outer bound on the secrecy capacity region of the Gaussian MARC-WT, and this is because the scalar Gaussian MARC-WT is always degraded. The capacity results on the Gaussian MARC-WT will be given in the next section.

# III. GAUSSIAN MULTIPLE-ACCESS RELAY WIRETAP CHANNEL

In this section, we investigate the Gaussian multiple-access relay wiretap channel (GMARC-WT). The signal received at each node is given by

$$Y_r = X_1 + X_2 + Z_r,$$
  

$$Y = X_1 + X_2 + X_r + Z_1,$$
  

$$Z = X_1 + X_2 + X_r + Z_2,$$
(3.1)

where  $Z_r \sim \mathcal{N}(0, N_r)$ ,  $Z_1 \sim \mathcal{N}(0, N_1)$ ,  $Z_2 \sim \mathcal{N}(0, N_2)$ , and they are independent. The Gaussian noise vectors  $Z_r^N$ ,  $Z_1^N$  and  $Z_2^N$  are composed of i.i.d. components with probability distributions  $Z_r \sim \mathcal{N}(0, N_r)$ ,  $Z_1 \sim \mathcal{N}(0, N_1)$  and  $Z_2 \sim \mathcal{N}(0, N_2)$ , respectively. The average power constraints of  $X_1^N$ ,  $X_2^N$  and  $X_r^N$  are  $\frac{1}{N} \sum_{i=1}^N E[X_{1,i}^2] \leq P_1$ ,  $\frac{1}{N} \sum_{i=1}^N E[X_{2,i}^2] \leq P_2$  and  $\frac{1}{N} \sum_{i=1}^N E[X_{r,i}^2] \leq P_r$ , respectively.

The remainder of this section is organized as follows. Subsection III-A shows the achievable secrecy rate regions of GMARC-WT, and the numerical examples and discussions are given in Subsection III-B.

# A. Capacity results on GMARC-WT

Theorem 5: The DF inner bound on the secrecy capacity region of the GMARC-WT is given by

$$\mathcal{R}^{g_{1}} = \bigcup_{0 \leq \gamma \leq 1} \left\{ \begin{array}{l} (R_{1}, R_{2}) : R_{1}, R_{2} \geq 0, \\ R_{1} \leq \min\{\frac{1}{2}\log(1 + \frac{P_{1}}{N_{r}}), \frac{1}{2}\log(1 + \frac{P_{1} + \gamma P_{r}}{N_{1}})\} - \frac{1}{2}\log\frac{P_{1} + P_{2} + P_{r} + N_{2}}{P_{2} + P_{r} + N_{2}}, \\ R_{2} \leq \min\{\frac{1}{2}\log(1 + \frac{P_{2}}{N_{r}}), \frac{1}{2}\log(1 + \frac{P_{2} + (1 - \gamma)P_{r}}{N_{1}})\} - \frac{1}{2}\log\frac{P_{1} + P_{2} + P_{r} + N_{2}}{P_{1} + P_{r} + N_{2}}, \\ R_{1} + R_{2} \leq \min\{\frac{1}{2}\log(1 + \frac{P_{1} + P_{2}}{N_{r}}), \frac{1}{2}\log(1 + \frac{P_{1} + P_{2} + P_{r}}{N_{1}})\} - \frac{1}{2}\log\frac{P_{1} + P_{2} + P_{r} + N_{2}}{P_{r} + N_{2}}. \end{array} \right\}.$$
(3.2)

Proof:

First, let  $X_r = V_1 + V_2$ , where  $V_1 \sim \mathcal{N}(0, \gamma P_r)$  and  $V_2 \sim \mathcal{N}(0, (1 - \gamma)P_r)$ . Let  $X_1 = \sqrt{\frac{(1-\alpha)P_1}{\gamma P_r}}V_1 + X_{10}$ , where  $0 \le \alpha \le 1$  and  $X_{10} \sim \mathcal{N}(0, \alpha P_1)$ . Analogously, let  $X_2 = \sqrt{\frac{(1-\beta)P_2}{(1-\gamma)P_r}}V_2 + X_{20}$ , where  $0 \le \beta \le 1$  and  $X_{20} \sim \mathcal{N}(0, \beta P_2)$ . Here note that  $V_1$ ,  $V_2$ ,  $X_{10}$  and  $X_{20}$  are independent random variables.

The region  $\mathcal{R}^{g_1}$  is obtained by substituting the above definitions into Theorem 1, and maximizing  $\alpha$  and  $\beta$  (the maximum of  $\mathcal{R}^{g_1}$  is achieved when  $\alpha = \beta = 1$ ). Thus, the proof of Theorem 5 is completed.

Theorem 6: The NF inner bound on the secrecy capacity region of the GMARC-WT is given by

$$\mathcal{R}^{g_2} = \text{convex closure of} \quad (\mathcal{G}^1 \bigcup \mathcal{G}^2),$$

where  $\mathcal{G}^1$  is given by

$$\mathcal{G}^{1} = \bigcup_{N_{1} \leq N_{2}} \left\{ \begin{array}{l} (R_{1}, R_{2}) : R_{1}, R_{2} \geq 0, \\ R_{1} \leq \frac{1}{2} \log(1 + \frac{P_{1}}{N_{1}}) - \frac{1}{2} \log(1 + \frac{P_{1}+P_{r}}{P_{2}+N_{2}}) + R_{r}, \\ R_{2} \leq \frac{1}{2} \log(1 + \frac{P_{2}}{N_{1}}) - \frac{1}{2} \log(1 + \frac{P_{2}+P_{r}}{P_{1}+N_{2}}) + R_{r}, \\ R_{1} + R_{2} \leq \frac{1}{2} \log(1 + \frac{P_{1}+P_{2}}{N_{1}}) - \frac{1}{2} \log(1 + \frac{P_{1}+P_{2}+P_{r}}{N_{2}}) + R_{r}. \end{array} \right\},$$

$$R_{r} = \min\{\frac{1}{2} \log(1 + \frac{P_{r}}{P_{1}+P_{2}+N_{1}}), \frac{1}{2} \log(1 + \frac{P_{r}}{P_{2}+N_{2}}), \frac{1}{2} \log(1 + \frac{P_{r}}{P_{1}+N_{2}})\}, \text{ and } \mathcal{G}^{2} \text{ is given by}$$

$$\mathcal{G}^{2} = \bigcup_{N_{1} \ge N_{2}} \left\{ \begin{array}{l} (R_{1}, R_{2}) : R_{1}, R_{2} \ge 0, \\ R_{1} \le \frac{1}{2} \log(1 + \frac{P_{1}}{N_{1}}) - \frac{1}{2} \log(1 + \frac{P_{1}}{P_{2} + N_{2}}), \\ R_{2} \le \frac{1}{2} \log(1 + \frac{P_{2}}{N_{1}}) - \frac{1}{2} \log(1 + \frac{P_{2}}{P_{1} + N_{2}}), \\ R_{1} + R_{2} \le \frac{1}{2} \log(1 + \frac{P_{1} + P_{2}}{N_{1}}) - \frac{1}{2} \log(1 + \frac{P_{1} + P_{2}}{N_{2}}). \end{array} \right\}.$$

Proof:

Here note that  $N_1 \leq N_2$  implies  $I(X_r; Y) \geq I(X_r; Z)$ . The region  $\mathcal{G}^1$  is obtained by substituting  $X_1 \sim \mathcal{N}(0, P_1)$ ,  $X_2 \sim \mathcal{N}(0, P_2)$  and  $X_r \sim \mathcal{N}(0, P_r)$  into the region  $\mathcal{L}^1$  of Theorem 2, and using the fact that  $X_1$ ,  $X_2$  and  $X_r$  are independent random variables.

Analogously,  $N_1 \ge N_2$  implies  $I(X_r; Y) \le I(X_r; Z)$ . The region  $\mathcal{G}^2$  is obtained by substituting  $X_1 \sim \mathcal{N}(0, P_1)$ ,  $X_2 \sim \mathcal{N}(0, P_2)$  and  $X_r \sim \mathcal{N}(0, P_r)$  into the region  $\mathcal{L}^2$  of Theorem 2, and using the fact that  $X_1$ ,  $X_2$  and  $X_r$  are independent random variables. Thus, the proof of Theorem 6 is completed.

Theorem 7: The CF inner bound on the secrecy capacity region of the GMARC-WT is given by

$$\mathcal{R}^{g3} = \text{convex closure of} \quad (\mathcal{G}^3 \bigcup \mathcal{G}^4).$$

where  $\mathcal{G}^3$  is given by

$$\mathcal{G}^{3} = \bigcup_{N_{1} \leq N_{2}} \left\{ \begin{array}{l} (R_{1}, R_{2}) : R_{1}, R_{2} \geq 0, \\ R_{1} \leq \frac{1}{2} \log(1 + \frac{P_{1}(Q+N_{1}+N_{r})}{N_{1}(N_{r}+Q)}) - \frac{1}{2} \log(1 + \frac{P_{1}+P_{r}}{P_{2}+N_{2}}) + R^{*}, \\ R_{2} \leq \frac{1}{2} \log(1 + \frac{P_{2}(Q+N_{1}+N_{r})}{N_{1}(N_{r}+Q)}) - \frac{1}{2} \log(1 + \frac{P_{2}+P_{r}}{P_{1}+N_{2}}) + R^{*}, \\ R_{1} + R_{2} \leq \frac{1}{2} \log(1 + \frac{(P_{1}+P_{2})(Q+N_{1}+N_{r})}{N_{1}(N_{r}+Q)}) - \frac{1}{2} \log(1 + \frac{P_{1}+P_{2}+P_{r}}{N_{2}}) + R^{*}. \end{array} \right\},$$

Q satisfies

$$\frac{1}{2}\log(1+\frac{P_1+P_2+N_r}{Q}) \leq \min\{\frac{1}{2}\log(1+\frac{P_r}{P_1+P_2+N_1}), \frac{1}{2}\log(1+\frac{P_r}{P_2+N_2}), \frac{1}{2}\log(1+\frac{P_r}{P_1+N_2})\}, \frac{1}{2}\log(1+\frac{P_r}{P_1+N_2})\}$$

and  $R^*$  satisfies

$$\begin{split} 0 &\leq R^* &\leq & \min\{\frac{1}{2}\log(1+\frac{P_r}{P_1+P_2+N_1}), \frac{1}{2}\log(1+\frac{P_r}{P_2+N_2}), \frac{1}{2}\log(1+\frac{P_r}{P_1+N_2})\}\\ &\quad -\frac{1}{2}\log(1+\frac{P_1+P_2+N_r}{Q}), \end{split}$$

and  $\mathcal{G}^4$  is given by

$$\mathcal{G}^{4} = \bigcup_{N_{1} \ge N_{2}} \left\{ \begin{array}{l} (R_{1}, R_{2}) : R_{1}, R_{2} \ge 0, \\ R_{1} \le \frac{1}{2} \log(1 + \frac{P_{1}(Q+N_{1}+N_{r})}{N_{1}(N_{r}+Q)}) - \frac{1}{2} \log(1 + \frac{P_{1}}{P_{2}+N_{2}}), \\ R_{2} \le \frac{1}{2} \log(1 + \frac{P_{2}(Q+N_{1}+N_{r})}{N_{1}(N_{r}+Q)}) - \frac{1}{2} \log(1 + \frac{P_{2}}{P_{1}+N_{2}}), \\ R_{1} + R_{2} \le \frac{1}{2} \log(1 + \frac{(P_{1}+P_{2})(Q+N_{1}+N_{r})}{N_{1}(N_{r}+Q)}) - \frac{1}{2} \log(1 + \frac{P_{1}+P_{2}}{N_{2}}). \end{array} \right\}$$

here Q satisfies  $Q \ge \frac{(P_1 + P_2)^2 + (P_1 + P_2)(N_r + N_1) + N_r N_1}{P_r}$ 

Proof:

Here note that  $N_1 \leq N_2$  implies  $I(X_r; Y) \geq I(X_r; Z)$ . The region  $\mathcal{G}^3$  is obtained by substituting  $X_1 \sim \mathcal{N}(0, P_1)$ ,  $X_2 \sim \mathcal{N}(0, P_2)$ ,  $X_r \sim \mathcal{N}(0, P_r)$ ,  $\hat{Y}_r = Y_r + Z_Q^1$  and  $Z_Q \sim \mathcal{N}(0, Q)$  into the region  $\mathcal{L}^3$  of Theorem 3, and using the fact that  $X_1, X_2$  and  $X_r$  are independent random variables.

Analogously,  $N_1 \ge N_2$  implies  $I(X_r; Y) \le I(X_r; Z)$ . The region  $\mathcal{G}^4$  is obtained by substituting  $X_1 \sim \mathcal{N}(0, P_1)$ ,  $X_2 \sim \mathcal{N}(0, P_2)$ ,  $X_r \sim \mathcal{N}(0, P_r)$ ,  $\hat{Y}_r = Y_r + Z_Q$  and  $Z_Q \sim \mathcal{N}(0, Q)$  into the region  $\mathcal{L}^4$  of Theorem 3, and using the fact that  $X_1, X_2$  and  $X_r$  are independent random variables. Thus, the proof of Theorem 7 is completed.

By using Theorem 4, we provide an outer bound on the secrecy capacity region of the GMARC-WT under the condition that  $N_2 \ge N_1$ , see the followings.

<sup>1</sup>Here note that  $\hat{Y}_r = Y_r + Z_Q$  is from the similar argument for the CF strategy of the Gaussian relay channel [28, pp. 402-403].

Theorem 8: For the case that  $N_2 \ge N_1$ , an outer bound  $\mathcal{R}^{gout}$  on the secrecy capacity region of the GMARC-WT is given by

$$\mathcal{R}^{gout} = \bigcup_{0 \le \alpha, \beta_1, \beta_2, \gamma \le 1} \left\{ \begin{array}{l} (R_1, R_2) : R_1 \ge 0, R_2 \ge 0, \\ R_1 \le \frac{1}{2} \log(1 + \frac{P_r(\alpha + \beta_2 - \alpha\beta_2) + \beta_2 P_1}{N_1}) - \frac{1}{2} \log(\frac{C + \gamma(P_1 + P_2 + P_r + N_2 - C)}{N_2 + P_r(\alpha + \beta_1 - \alpha\beta_1) + \beta_1 P_2}), \\ R_2 \le \frac{1}{2} \log(1 + \frac{P_r(\alpha + \beta_1 - \alpha\beta_1) + \beta_1 P_2}{N_1}) - \frac{1}{2} \log(\frac{C + \gamma(P_1 + P_2 + P_r + N_2 - C)}{N_2 + P_r(\alpha + \beta_2 - \alpha\beta_2) + \beta_2 P_1}), \\ R_1 + R_2 \le \frac{1}{2} \log(\frac{C + \gamma(P_1 + P_2 + P_r + N_1 - C)}{N_1}) - \frac{1}{2} \log(\frac{C + \gamma(P_1 + P_2 + P_r + N_2 - C)}{N_2 + \alpha P_r}), \end{array} \right), \end{array} \right\},$$

where C satisfies

$$C = \max\{N_2 + P_r(\alpha + \beta_1 - \alpha\beta_1) + \beta_1 P_2, N_2 + P_r(\alpha + \beta_2 - \alpha\beta_2) + \beta_2 P_1\}$$

Proof:

See Appendix E.

*Theorem 9:* Finally, remember that [9] provides an achievable secrecy rate region  $\mathcal{R}^{Gi}$  of the Gaussian multipleaccess wiretap channel (GMAC-WT), and it is given by

$$\mathcal{R}^{Gi} = \begin{cases} (R_1, R_2) : R_1, R_2 \ge 0, \\ R_1 \le \frac{1}{2} \log(1 + \frac{P_1}{N_1}) - \frac{1}{2} \log(1 + \frac{P_1}{N_2 + P_2}) \\ R_2 \le \frac{1}{2} \log(1 + \frac{P_2}{N_1}) - \frac{1}{2} \log(1 + \frac{P_2}{N_2 + P_1}) \\ R_1 + R_2 \le \frac{1}{2} \log(1 + \frac{P_1 + P_2}{N_1}) - \frac{1}{2} \log(1 + \frac{P_1 + P_2}{N_2}) \end{cases} \end{cases}$$

Proof:

The proof is in [9], and it is omitted here.

## B. Numerical Examples and Discussions

Letting  $P_1 = 5$ ,  $P_2 = 6$ ,  $P_r = 20$ ,  $N_1 = 2$ ,  $N_2 = 14$  and Q = 200, the following Figure 2, 3, 4 and 5 show the inner and outer bounds on the secrecy capacity region of the GMARC-WT for different values of  $N_r$ .

Compared with the achievable secrecy rate region  $\mathcal{R}^{G_i}$  of GMAC-WT, it is easy to see that the NF region ( $\mathcal{R}^{g^2}$ ) and the CF region ( $\mathcal{R}^{g^3}$ ) enhance the region  $\mathcal{R}^{G_i}$  (no relay). The CF region is always smaller than the NF region, and when  $Q \to \infty$ , the CF region tends to the NF region. For the DF region ( $\mathcal{R}^{g_1}$ ), we find that when  $N_r$  is much larger than  $N_1$ ,  $\mathcal{R}^{g_1}$  is even smaller than  $\mathcal{R}^{G_i}$  (see Figure 2). When  $N_r$  is close to  $N_1$  (still larger than  $N_1$ ),  $\mathcal{R}^{g_1}$ is larger than  $\mathcal{R}^{G_i}$ , but it is still smaller than the NF and CF regions (see Figure 3). When  $N_r$  is smaller than  $N_1$ , as we can see in Figure 4 and 5, the DF region  $\mathcal{R}^{g_1}$  is larger than the NF and CF regions.

Figure 2, 3, 4 and 5 also show that there exists a gap between the inner and outer bounds, and the gap is reduced as  $N_r$  decreases.



Fig. 2: The bounds on the secrecy capacity region of GMARC-WT for  $N_r = 5$ 



Fig. 3: The bounds on the secrecy capacity region of GMARC-WT for  $N_r = 2.3$ 

# IV. CONCLUSION

In this paper, first, we provide three inner bounds on the secrecy capacity region (achievable secrecy rate regions) of the discrete memoryless model of Figure 1. The decode-forward (DF), noise-forward (NF), and compress-forward (CF) relay strategies are used in the construction of these inner bounds. Second, we investigate the degraded discrete memoryless MARC-WT, and present an outer bound on the secrecy capacity region of this degraded case. Finally,



Fig. 4: The bounds on the secrecy capacity region of GMARC-WT for  $N_r = 1.6$ 



Fig. 5: The bounds on the secrecy capacity region of GMARC-WT for  $N_r = 0$ 

we study the Gaussian MARC-WT, and find that the NF and CF strategies help to enhance Tekin-Yener's achievable secrecy rate region of Gaussian MAC-WT. Moreover, we find that if the channel from the transmitters to the relay is less noisy than the channels from the transmitters to the legitimate receiver and the wiretapper, the achievable secrecy rate region of the DF strategy is even larger than the corresponding regions of the NF and CF strategies.

#### ACKNOWLEDGEMENT

The authors would like to thank Professor Ning Cai for his valuable suggestions to improve this paper. This work was supported by a sub-project in National Basic Research Program of China under Grant 2012CB316100 on Broadband Mobile Communications at High Speeds, the National Natural Science Foundation of China under Grant 61301121, and the Fundamental Research Funds for the Central Universities under Grant 2682014CX099.

# APPENDIX A

# **PROOF OF THEOREM 1**

In order to prove Theorem 1, we need to show that any pair  $(R_1, R_2) \in \mathcal{R}^{d_1}$  is achievable, i.e., for any  $\epsilon > 0$ , there exists a sequence of codes  $(2^{NR_1}, 2^{NR_2}, N)$  such that  $\frac{\log ||\mathcal{W}_1||}{N} = R_1, \frac{\log ||\mathcal{W}_2||}{N} = R_2, P_e \le \epsilon$  and  $\frac{1}{N}H(W_1, W_2|Z^N) \ge R_1 + R_2 - \epsilon$ . The details are as follows.

The coding scheme combines the decode-and-forward (DF) strategy of MARC [21], random binning, superposition coding, and block Markov coding techniques, see the followings.

Fix the joint probability mass function  $P_{Y,Z,Y_r|X_r,X_1,X_2}(y,z,y_r|x_r,x_1,x_2)P_{X_r|V_1,V_2}(x_r|v_1,v_2)$  $P_{X_1|V_1}(x_1|v_1)P_{X_2|V_2}(x_2|v_2)P_{V_1}(v_1)P_{V_2}(v_2)$ . For a given  $(R_1, R_2) \in \mathcal{R}^{d_1}$ , define the messages  $W_1$  and  $W_2$  taking values in the alphabets  $W_1$  and  $W_2$ , respectively, where

$$\mathcal{W}_1 = \{1, 2, ..., 2^{NR_1}\}, \ \mathcal{W}_2 = \{1, 2, ..., 2^{NR_2}\}.$$

# **Relay Code-books Construction:**

For a given  $R_1^* \ge 0$ , generate at random  $2^{N(R_1+R_1^*)}$  i.i.d. sequences  $v_1^N$  according to  $P_{V_1^N}(v_1^N) = \prod_{i=1}^N P_{V_1}(v_{1,i})$ . Index them as  $v_1^N(a_1, b_1)$ , where  $a_1 \in \{1, 2, ..., 2^{NR_1}\}$  and  $b_1 \in \{1, 2, ..., 2^{NR_1^*}\}$ . For convenience, define  $s_1 = 1$  $(a_1, b_1)$ , where  $s_1 \in \{1, 2, ..., 2^{N(R_1 + R_1^*)}\}$ .

Analogously, for a given  $R_2^* \ge 0$ , generate at random  $2^{N(R_2+R_2^*)}$  i.i.d. sequences  $v_2^N$  according to  $P_{V_2^N}(v_2^N) =$  $\prod_{i=1}^{N} P_{V_2}(v_{2,i}). \text{ Index them as } v_2^N(a_2, b_2), \text{ where } a_2 \in \{1, 2, ..., 2^{NR_2}\} \text{ and } b_2 \in \{1, 2, ..., 2^{NR_2^*}\}.$  For convenience, define  $s_2 = (a_2, b_2)$ , where  $s_2 \in \{1, 2, ..., 2^{N(R_2 + R_2^*)}\}$ .

Generate at random  $2^{N(R_1+R_1^*+R_2+R_2^*)}$  i.i.d. sequences  $x_r^N$  according to  $P_{X_r^N|V_1^N,V_2^N}(x_r^N|v_1^N,v_2^N) =$ 

 $\prod_{i=1}^{N} P_{X_{r,i}|V_{1,i},V_{2,i}}(x_{r,i}|v_{1,i},v_{2,i}). \text{ Index them as } x_r^N(s_1,s_2), \text{ where } s_1 \in \{1,2,...,2^{N(R_1+R_1^*)}\} \text{ and } s_2 \in \{1,2,...,2^{N(R_2+R_2^*)}\}.$ 

# Transmitters' Code-books Construction:

- For a given  $v_1^N(s_1)$ , generate at random  $2^{N(R_1+R_1^*)}$  i.i.d. sequences  $x_1^N(w_1, w_1^*|s_1)$  ( $w_1 \in \{1, 2, ..., 2^{NR_1}\}, w_1^* \in \{1, 2, ..., 2^{NR_1}\}$  $\{1, 2, ..., 2^{NR_1^*}\}, s_1 \in \{1, 2, ..., 2^{N(R_1 + R_1^*)}\})$  according to  $\prod_{i=1}^N P_{X_1|V_1}(x_{1,i}|v_{1,i}).$
- Analogously, for a given  $v_2^N(s_2)$ , generate at random  $2^{N(R_2+R_2^*)}$  i.i.d. sequences  $x_2^N(w_2, w_2^*|s_2)$  ( $w_2 \in$  $\{1, 2, ..., 2^{NR_2}\}, w_2^* \in \{1, 2, ..., 2^{NR_2^*}\}, s_2 \in \{1, 2, ..., 2^{N(R_2 + R_2^*)}\}) \text{ according to } \prod_{i=1}^N P_{X_2|V_2}(x_{2,i}|v_{2,i}).$

Encoding: We exploit the block Markov coding scheme, because, as argued in [23], the loss induced by this scheme is negligible as the number of blocks  $n \to \infty$ . For block i  $(1 \le i \le n)$ , encoding proceeds as follows.

First, for convenience, the messages  $w_1$ ,  $w_1^*$ ,  $w_2$ ,  $w_2^*$ ,  $s_1$  and  $s_2$  transmitted in the *i*-th block are denoted by  $w_{1,i}$ ,  $w_{1,i}^*, w_{2,i}, w_{2,i}^*, s_{1,i}$  and  $s_{2,i}$ , respectively.

# • (Channel encoders)

1) The message  $w_{1,i}^*$   $(1 \le i \le n-1)$  is randomly chosen from the set  $\{1, 2, ..., 2^{NR_1^*}\}$ . The transmitter 1 (encoder 1) sends  $x_1^N(w_{1,1}, w_{1,1}^*|1, 1)$  at the first block,  $x_1^N(w_{1,i}, w_{1,i}^*|w_{1,i-1}, w_{1,i-1}^*)$  (note that here  $s_{1,i} = (w_{1,i-1}, w_{1,i-1}^*)$ ) from block  $2 \le i \le n-1$ , and  $x_1^N(1, 1|w_{1,n-1}, w_{1,n-1}^*)$  at block n  $(s_{1,n} = (w_{1,n-1}, w_{1,n-1}^*))$ . 2) The message  $w_{2,i}^*$   $(1 \le i \le n-1)$  is randomly chosen from the set  $\{1, 2, ..., 2^{NR_2^*}\}$ . The transmitter 2 (encoder 2) sends  $x_2^N(w_{2,1}, w_{2,1}^*|1, 1)$  at the first block,  $x_2^N(w_{2,i}, w_{2,i}^*|w_{2,i-1}, w_{2,i-1}^*)$   $(s_{2,i} = (w_{2,i-1}, w_{2,i-1}^*))$  from block  $2 \le i \le n-1$ , and  $x_2^N(1, 1|w_{2,n-1}, w_{2,n-1}^*)$  at block n  $(s_{2,n} = (w_{2,n-1}, w_{2,n-1}^*))$ .

# • (Relay encoder)

The relay sends  $(v_1^N(1,1), v_2^N(1,1), x_r^N(1,1,1,1))$  at the first block, and

 $(v_1^N(\hat{s}_{1,i}), v_2^N(\hat{s}_{2,i}), x_r^N(\hat{s}_{1,i}, \hat{s}_{2,i})) \text{ from block } 2 \leq i \leq n, \text{ where } \hat{s}_{1,i} = (\hat{w}_{1,i-1}, \hat{w}_{1,i-1}^*) \text{ and } \hat{s}_{2,i} = (\hat{w}_{2,i-1}, \hat{w}_{2,i-1}^*).$ 

# Decoding: Decoding proceeds as follows.

1) (At the relay) At the end of block i  $(1 \le i \le n)$ , the relay already has an estimation of the  $s_{1,i}$  and  $s_{2,i}$  (denoted by  $\hat{s}_{1,i}$  and  $\hat{s}_{2,i}$ , respectively), and will declare that it receives  $\hat{w}_{1,i}$ ,  $\hat{w}_{1,i}^*$ ,  $\hat{w}_{2,i}$  and  $\hat{w}_{2,i}^*$  if this is the only quadruple such that  $(x_1^N(\hat{w}_{1,i}, \hat{w}_{1,i}^*|\hat{s}_{1,i}), x_2^N(\hat{w}_{2,i}, \hat{w}_{2,i}^*|\hat{s}_{2,i}), x_r^N(\hat{s}_{1,i}, \hat{s}_{2,i}), v_1(\hat{s}_{1,i}), v_2(\hat{s}_{2,i}), y_r^N(i))$  are jointly typical. Here note that  $y_r^N(i)$  indicates the output sequence  $y_r^N$  in block i,  $\hat{s}_{1,i+1} = (\hat{w}_{1,i}, \hat{w}_{1,i}^*)$  and  $\hat{s}_{2,i+1} = (\hat{w}_{2,i}, \hat{w}_{2,i}^*)$ . The indexes  $\hat{s}_{1,i+1}$  and  $\hat{s}_{2,i+1}$  will be used in the i + 1-th block.

Based on the AEP, the error probability  $Pr\{(\hat{s}_{1,i+1}, \hat{s}_{2,i+1}) \neq (s_{1,i+1}, s_{2,i+1})\}$  goes to 0 if

$$R_1 + R_1^* \le I(X_1; Y_r | X_r, V_1, V_2, X_2), \tag{A1}$$

$$R_2 + R_2^* \le I(X_2; Y_r | X_r, V_1, V_2, X_1), \tag{A2}$$

$$R_1 + R_1^* + R_2 + R_2^* \le I(X_1, X_2; Y_r | X_r, V_1, V_2).$$
(A3)

2) (At the legitimate receiver) The legitimate receiver decodes from the last block, i.e., block n. At the end of block i + 1, the legitimate receiver already has an estimation of the  $\check{w}_{1,i+1}$ ,  $\check{w}_{1,i+1}^*$ ,  $\check{w}_{2,i+1}$  and  $\check{w}_{2,i+1}^*$ , and will declare that it receives  $\check{s}_{1,i+1}$  and  $\check{s}_{2,i+1}$  if this is the only pair such that  $(x_1^N(\check{w}_{1,i+1},\check{w}_{1,i+1}^*|\check{s}_{1,i+1}), x_2^N(\check{w}_{2,i+1}, \check{w}_{2,i+1}^*|\check{s}_{2,i+1}), x_r^N(\check{s}_{1,i+1},\check{s}_{2,i+1}), v_1(\check{s}_{1,i+1}), v_2(\check{s}_{2,i+1}), y^N(i+1))$  are jointly typical. Here note that  $y^N(i+1)$  indicates the output sequence  $y^N$  in block i + 1,  $\check{s}_{1,i+1} = (\check{w}_{1,i},\check{w}_{1,i}^*)$  and  $\check{s}_{2,i+1} = (\check{w}_{2,i},\check{w}_{2,i}^*)$ .

Based on the AEP, the error probability  $Pr\{(\check{s}_{1,i+1},\check{s}_{2,i+1})\neq(s_{1,i+1},s_{2,i+1})\}$  goes to 0 if

$$R_1 + R_1^* \le I(V_1, X_r, X_1; Y | V_2, X_2) \stackrel{(a)}{=} I(X_r, X_1; Y | X_2, V_2),$$
(A4)

$$R_2 + R_2^* \le I(V_2, X_r, X_2; Y | V_1, X_1) \stackrel{(b)}{=} I(X_r, X_2; Y | X_1, V_1),$$
(A5)

$$R_1 + R_1^* + R_2 + R_2^* \le I(V_1, V_2, X_r, X_1, X_2; Y) \stackrel{(c)}{=} I(X_r, X_1, X_2; Y),$$
(A6)

where (a) is from the Markov chain  $V_1 \rightarrow (X_r, X_1, X_2, V_2) \rightarrow Y$ , (b) is from the Markov chain  $V_2 \rightarrow (X_r, X_1, X_2, V_1) \rightarrow Y$ , and (c) is from the Markov chain  $(V_1, V_2) \rightarrow (X_r, X_1, X_2) \rightarrow Y$ .

By using (A1), (A2), (A3), (A4), (A5) and (A6), it is easy to check that  $P_e \leq \epsilon$ . It remains to show that  $\Delta \geq R_1 + R_2 - \epsilon$ , see the followings.

# **Equivocation Analysis:**

Similar to the equivocation analysis of [18, proof of Theorem 2], for simplicity, we only focus on the equivocation of one block, see the followings.

$$\frac{1}{N}H(W_{1}, W_{2}|Z^{N}) = \frac{1}{N}(H(W_{1}, W_{2}, Z^{N}) - H(Z^{N}))$$

$$= \frac{1}{N}(H(W_{1}, W_{2}, Z^{N}, X_{1}^{N}, X_{2}^{N}) - H(X_{1}^{N}, X_{2}^{N}|W_{1}, W_{2}, Z^{N}) - H(Z^{N}))$$

$$\stackrel{(a)}{=} \frac{1}{N}(H(Z^{N}|X_{1}^{N}, X_{2}^{N}) + H(X_{1}^{N}) + H(X_{2}^{N}) - H(X_{1}^{N}, X_{2}^{N}|W_{1}, W_{2}, Z^{N}) - H(Z^{N}))$$

$$= \frac{1}{N}(H(X_{1}^{N}) + H(X_{2}^{N}) - I(X_{1}^{N}, X_{2}^{N}; Z^{N}) - H(X_{1}^{N}, X_{2}^{N}|W_{1}, W_{2}, Z^{N})), \quad (A7)$$

where (a) follows from  $(W_1, W_2) \to (X_1^N, X_2^N) \to Z^N$ ,  $H(W_1|X_1^N) = 0$ ,  $H(W_2|X_2^N) = 0$ , and  $X_1^N$  is independent of  $X_2^N$ .

Consider the first term of (A7), the code-book generation of  $x_1^N$  shows that the total number of  $x_1^N$  is  $2^{N(R_1+R_1^*)}$ . Thus, using the same approach as that in [7, Lemma 3], we have

$$\frac{1}{N}H(X_1^N) \ge R_1 + R_1^* - \epsilon_{1,N},\tag{A8}$$

where  $\epsilon_{1,N} \to 0$  as  $N \to \infty$ .

Analogously, the second term of (A7) is bounded by

$$\frac{1}{N}H(X_2^N) \ge R_2 + R_2^* - \epsilon_{2,N},\tag{A9}$$

where  $\epsilon_{2,N} \to 0$  as  $N \to \infty$ .

For the third term of (A7), since the channel is memoryless, and  $X_1^N$ ,  $X_2^N$ ,  $X_r^N$  are i.i.d. generated, we get

$$\frac{1}{N}I(X_1^N, X_2^N; Z^N) = I(X_1, X_2; Z).$$
(A10)

Now, we consider the last term of (A7). Given  $W_1$  and  $W_2$ , the wiretapper does joint decoding at each block. At the end of block 1, the wiretapper tries to find a unique pair  $(\tilde{w}_{1,1}^*, \tilde{w}_{2,1}^*)$  such that

 $(x_1^N(w_{1,1}, \tilde{w}_{1,1}^*|1, 1), x_2^N(w_{2,1}, \tilde{w}_{2,1}^*|1, 1), z^N(1))$  are jointly typical. At the end of block  $i \ (2 \le i \le n-1)$ , the wiretapper already has an estimation of the  $\tilde{w}_{1,i-1}^*$  and  $\tilde{w}_{2,i-1}^*$ , and thus he also get  $\tilde{s}_{1,i} = (w_{1,i-1}, \tilde{w}_{1,i-1}^*)$  and  $\tilde{s}_{2,i} = (w_{2,i-1}, \tilde{w}_{2,i-1}^*)$ . Then he tries to find a unique pair  $(\tilde{w}_{1,i}^*, \tilde{w}_{2,i}^*)$  such that  $(x_1^N(w_{1,i}, \tilde{w}_{1,i}^*|\tilde{s}_{1,i}), x_2^N(w_{2,i}, \tilde{w}_{2,i}^*|\tilde{s}_{2,i}), z^N(i))$  are jointly typical. Based on the AEP, the error probability  $Pr\{(\tilde{w}_{1,i}^*, \tilde{w}_{2,i}^*) \neq (w_{1,i}^*, w_{2,i}^*)\}$  goes to 0 if

$$R_1^* \le I(X_1; Z | X_2), \tag{A11}$$

$$R_2^* \le I(X_2; Z | X_1), \tag{A12}$$

$$R_1^* + R_2^* \le I(X_1, X_2; Z).$$
(A13)

Then based on Fanos inequality, we have

$$\frac{1}{N}H(X_1^N, X_2^N | W_1, W_2, Z^N) \le \epsilon_{3,N},$$
(A14)

where  $\epsilon_{3,N} \to 0$  as  $N \to \infty$ .

Substituting (A8), (A9), (A10) and (A14) into (A7), we have

$$\frac{1}{N}H(W_1, W_2|Z^N) \ge R_1 + R_1^* + R_2 + R_2^* - I(X_1, X_2; Z) - \epsilon_{1,N} - \epsilon_{2,N} - \epsilon_{3,N}.$$
(A15)

It is easy to see that if we let

$$R_1^* + R_2^* = I(X_1, X_2; Z), (A16)$$

and choose sufficiently large N such that  $\epsilon_{1,N} + \epsilon_{2,N} + \epsilon_{3,N} \leq \epsilon$ ,  $\Delta = \frac{1}{N}H(W_1, W_2|Z^N) \geq R_1 + R_2 - \epsilon$  is guaranteed.

Based on (A1), (A2), (A3), (A4), (A5), (A6), (A11), (A12) and (A16), the achievable region  $\mathcal{R}^{d_1}$  is obtained. The proof of Theorem 1 is completed.

## APPENDIX B

#### **PROOF OF THEOREM 2**

For Theorem 2, we only need to prove that the corner points of  $\mathcal{L}^1$  and  $\mathcal{L}^2$  are achievable, see the followings.

(Case 1) If I(X<sub>r</sub>; Y) ≥ I(X<sub>r</sub>; Z), we allow the legitimate receiver to decode x<sup>N</sup><sub>r</sub>, and the wiretapper can not decode it. For case 1, it is sufficient to show that the pair (R<sub>1</sub>, R<sub>2</sub>) ∈ L<sup>1</sup> with the condition

$$R_1 = I(X_1; Y | X_2, X_r) - I(X_1, X_r; Z) + R_r, \quad R_2 = I(X_2; Y | X_r) - I(X_2; Z | X_1, X_r)$$
(A17)

is achievable. The achievability proof of the other corner point  $(R_1 = I(X_1; Y|X_r) - I(X_1; Z|X_2, X_r), R_2 = I(X_2; Y|X_1, X_r) - I(X_2, X_r; Z) + R_r)$  follows by symmetry.

(Case 2) If I(X<sub>r</sub>; Y) ≤ I(X<sub>r</sub>; Z), we allow both the receivers to decode x<sup>N</sup><sub>r</sub>. For case 2, it is sufficient to show that the pair (R<sub>1</sub>, R<sub>2</sub>) ∈ L<sup>2</sup> with the condition

$$R_1 = I(X_1; Y|X_2, X_r) - I(X_1; Z|X_r), \quad R_2 = I(X_2; Y|X_r) - I(X_2; Z|X_1, X_r)$$
(A18)

is achievable. The achievability proof of the other corner point  $(R_1 = I(X_1; Y|X_r) - I(X_1; Z|X_2, X_r), R_2 = I(X_2; Y|X_1, X_r) - I(X_2; Z|X_r))$  follows by symmetry.

Fix the joint probability mass function  $P_{Y,Z,Y_r|X_r,X_1,X_2}(y, z, y_r|x_r, x_1, x_2)P_{X_r}(x_r)P_{X_1}(x_1)P_{X_2}(x_2)$ . Define the messages  $W_1$ ,  $W_2$  taking values in the alphabets  $W_1$ ,  $W_2$ , respectively, where

$$\mathcal{W}_1 = \{1, 2, ..., 2^{NR_1}\}, \ \mathcal{W}_2 = \{1, 2, ..., 2^{NR_2}\}.$$

Code-book Construction for the Two Cases:

- Code-book construction for case 1:
  - First, generate at random  $2^{N(R_r-\epsilon')}$  (where  $\epsilon'$  is a small positive number) i.i.d. sequences at the relay node each drawn according to  $P_{X_r^N}(x_r^N) = \prod_{i=1}^N P_{X_r}(x_{r,i})$ , index them as  $x_r^N(a)$ ,  $a \in [1, 2^{N(R_r-\epsilon')}]$ , where

$$R_r = \min\{I(X_r; Z|X_1), I(X_r; Z|X_2), I(X_r; Y)\}.$$
(A19)

Here note that

$$R_r \ge I(X_r; Z). \tag{A20}$$

- Second, generate  $2^{N(I(X_2;Y|X_r)-\epsilon')}$  i.i.d. codewords  $x_2^N$  according to  $P_{X_2}(x_2)$ , and divide them into  $2^{NR_2}$  bins. Each bin contains  $2^{N(I(X_2;Y|X_r)-\epsilon'-R_2)}$  codewords, where

$$I(X_2; Y|X_r) - \epsilon' - R_2 = I(X_2; Z|X_1, X_r) - \epsilon'.$$
(A21)

- Third, generate  $2^{N(I(X_1;Y|X_2,X_r)-\epsilon')}$  i.i.d. codewords  $x_1^N$  according to  $P_{X_1}(x_1)$ , and divide them into  $2^{NR_1}$  bins. Each bin contains  $2^{N(I(X_1;Y|X_2,X_r)-\epsilon'-R_1)}$  codewords.

# • Code-book Construction for case 2:

- Generate at random  $2^{N(R_r-\epsilon')}$  i.i.d. sequences at the relay node each drawn according to  $P_{X_r^N}(x_r^N) = \prod_{i=1}^N P_{X_r}(x_{r,i})$ , index them as  $x_r^N(a)$ ,  $a \in [1, 2^{N(R_r-\epsilon')}]$ , where

$$R_r = I(X_r; Y) \le I(X_r; Z). \tag{A22}$$

- Second, generate  $2^{N(I(X_2;Y|X_r)-\epsilon')}$  i.i.d. codewords  $x_2^N$  according to  $P_{X_2}(x_2)$ , and divide them into  $2^{NR_2}$  bins. Each bin contains  $2^{N(I(X_2;Y|X_r)-\epsilon'-R_2)}$  codewords, where

$$I(X_2; Y|X_r) - \epsilon' - R_2 = I(X_2; Z|X_1, X_r) - \epsilon'.$$
(A23)

- Third, generate  $2^{N(I(X_1;Y|X_2,X_r)-\epsilon')}$  i.i.d. codewords  $x_1^N$  according to  $P_{X_1}(x_1)$ , and divide them into  $2^{NR_1}$  bins. Each bin contains  $2^{N(I(X_1;Y|X_2,X_r)-\epsilon'-R_1)}$  codewords, where

$$I(X_1; Y|X_2, X_r) - \epsilon' - R_1 = I(X_1; Z|X_r) - \epsilon'.$$
(A24)

## **Encoding for both cases:**

The relay uniformly picks a codeword  $x_r^N(a)$  from  $[1, 2^{N(R_r - \epsilon')}]$ , and sends  $x_r^N(a)$ .

For a given confidential message  $w_2$ , randomly choose a codeword  $x_2^N$  in bin  $w_2$  to transmit. Similarly, for a given confidential message  $w_1$ , randomly choose a codeword  $x_1^N$  in bin  $w_1$  to transmit.

#### Decoding for both cases:

For a given  $y^N$ , try to find a sequence  $x_r^N(\hat{a})$  such that  $(x_r^N(\hat{a}), y^N)$  are jointly typical. If there exists a unique sequence with the index  $\hat{a}$ , put out the corresponding  $\hat{a}$ , else declare a decoding error. Based on the AEP and (A19) (or (A22)), the probability  $Pr\{\hat{a} = a\}$  goes to 1.

After decoding  $\hat{a}$ , the legitimate receiver tries to find a sequence  $x_2^N(\hat{w}_2)$  such that  $(x_2^N(\hat{w}_2), x_r^N(\hat{a}), y^N)$  are jointly typical. If there exists a unique sequence with the index  $\hat{w}_2$ , put out the corresponding  $\hat{w}_2$ , else declare a decoding error. Based on the AEP and the construction of  $x_2^N$  for both cases, the probability  $Pr\{\hat{w}_2 = w_2\}$  goes to 1.

Finally, after decoding  $\hat{a}$  and  $\hat{w}_2$ , the legitimate receiver tries to find a sequence  $x_1^N(\hat{w}_1)$  such that  $(x_1^N(\hat{w}_1), x_2^N(\hat{w}_2), x_r^N(\hat{a}), y^N)$  are jointly typical. If there exists a unique sequence with the index  $\hat{w}_1$ , put out the corresponding  $\hat{w}_1$ , else declare a decoding error. Based on the AEP and the construction of  $x_1^N$  for both cases, the probability  $Pr\{\hat{w}_1 = w_1\}$  goes to 1.

 $P_e \leq \epsilon$  is easy to be checked by using the above encoding-decoding schemes. Now, it remains to prove  $\Delta \geq R_1 + R_2 - \epsilon$  for both cases, see the followings.

# **Equivocation Analysis:**

**Proof of**  $\Delta \geq R_1 + R_2 - \epsilon$  for case 1:

$$\Delta = \frac{1}{N} H(W_1, W_2 | Z^N)$$
  
=  $\frac{1}{N} (H(W_1 | Z^N) + H(W_2 | W_1, Z^N)).$  (A25)

The first term in (A25) is bounded as follows.

$$\frac{1}{N}H(W_{1}|Z^{N}) = \frac{1}{N}(H(W_{1},Z^{N}) - H(Z^{N}))$$

$$= \frac{1}{N}(H(W_{1},Z^{N},X_{1}^{N},X_{r}^{N}) - H(X_{1}^{N},X_{r}^{N}|W_{1},Z^{N}) - H(Z^{N}))$$

$$\stackrel{(a)}{=} \frac{1}{N}(H(Z^{N}|X_{1}^{N},X_{r}^{N}) + H(X_{1}^{N}) + H(X_{r}^{N}) - H(X_{1}^{N},X_{r}^{N}|W_{1},Z^{N}) - H(Z^{N}))$$

$$= \frac{1}{N}(H(X_{1}^{N}) + H(X_{r}^{N}) - I(X_{1}^{N},X_{r}^{N};Z^{N}) - H(X_{1}^{N},X_{r}^{N}|W_{1},Z^{N})), \quad (A26)$$

where (a) follows from  $W_1 \to (X_1^N, X_r^N) \to Z^N$ ,  $H(W_1|X_1^N) = 0$  and the fact that  $X_1^N$  is independent of  $X_r^N$ .

Consider the first term in (A26), the code-book generation of  $x_1^N$  shows that the total number of  $x_1^N$  is  $2^{N(I(X_1;Y|X_2,X_r)-\epsilon')}$ . Thus, using the same approach as that in [7, Lemma 3], we have

$$\frac{1}{N}H(X_{1}^{N}) \ge I(X_{1};Y|X_{2},X_{r}) - \epsilon^{'} - \epsilon_{1,N},$$
(A27)

where  $\epsilon_{1,N} \to 0$  as  $N \to \infty$ .

For the second term in (A26), the code-book generation of  $x_r^N$  guarantees that

$$\frac{1}{N}H(X_r^N) \ge R_r - \epsilon' - \epsilon_{2,N},\tag{A28}$$

where  $\epsilon_{2,N} \to 0$  as  $N \to \infty$ .

For the third term in (A26), since the channel is memoryless, and  $X_1^N$ ,  $X_2^N$ ,  $X_r^N$  are i.i.d. generated, we get

$$\frac{1}{N}I(X_1^N, X_r^N; Z^N) = I(X_1, X_r; Z).$$
(A29)

Now, we consider the last term of (A26). Given  $w_1$ , the wiretapper can do joint decoding. Specifically, given  $z^N$  and  $w_1$ ,

$$\frac{1}{N}H(X_1^N, X_r^N | W_1, Z^N) \le \epsilon_{3,N}$$
(A30)

 $(\epsilon_{3,N} \to 0 \text{ as } N \to \infty)$  is guaranteed if  $R_r \leq I(X_r; Z|X_1)$  and  $R_r \geq I(X_r; Z)$ , and this is from the properties of AEP (similar argument is used in the proof of [18, Theorem 3]). By checking (A19) and (A20), (A30) is obtained. Substituting (A27), (A28), (A29) and (A30) into (A26), we have

$$\frac{1}{N}H(W_{1}|Z^{N}) \ge I(X_{1};Y|X_{2},X_{r}) + R_{r} - I(X_{1},X_{r};Z) - 2\epsilon^{'} - \epsilon_{1,N} - \epsilon_{2,N} - \epsilon_{3,N}.$$
(A31)

The second term in (A25) is bounded as follows.

$$\frac{1}{N}H(W_{2}|W_{1}, Z^{N}) \geq \frac{1}{N}H(W_{2}|W_{1}, Z^{N}, X_{1}^{N}, X_{r}^{N}) 
\stackrel{(1)}{=} \frac{1}{N}H(W_{2}|Z^{N}, X_{1}^{N}, X_{r}^{N}) 
= \frac{1}{N}(H(W_{2}, Z^{N}, X_{1}^{N}, X_{r}^{N}) - H(Z^{N}, X_{1}^{N}, X_{r}^{N})) 
= \frac{1}{N}(H(W_{2}, Z^{N}, X_{1}^{N}, X_{r}^{N}, X_{2}^{N}) - H(X_{2}^{N}|W_{2}, Z^{N}, X_{1}^{N}, X_{r}^{N}) - H(Z^{N}, X_{1}^{N}, X_{r}^{N})) 
\stackrel{(2)}{=} \frac{1}{N}(H(Z^{N}|X_{1}^{N}, X_{2}^{N}, X_{r}^{N}) + H(X_{r}^{N}) + H(X_{1}^{N}) + H(X_{2}^{N}) 
-H(X_{2}^{N}|W_{2}, Z^{N}, X_{1}^{N}, X_{r}^{N}) - H(Z^{N}|X_{1}^{N}, X_{r}^{N}) - H(X_{1}^{N}) - H(X_{r}^{N})) 
= \frac{1}{N}(H(X_{2}^{N}) - I(X_{2}^{N}; Z^{N}|X_{1}^{N}, X_{r}^{N}) - H(X_{2}^{N}|W_{2}, Z^{N}, X_{1}^{N}, X_{r}^{N})), \quad (A32)$$

where (1) is from the Markov chain  $W_1 \to (Z^N, X_1^N, X_r^N) \to W_2$ , and (2) is from the Markov chain  $W_2 \to (X_1^N, X_2^N, X_r^N) \to Z^N$ ,  $H(W_2|X_2^N) = 0$ , and the fact that  $X_1^N, X_2^N$  and  $X_r^N$  are independent.

Consider the first term in (A32), the code-book generation of  $x_2^N$  shows that the total number of  $x_2^N$  is  $2^{N(I(X_2;Y|X_r)-\epsilon')}$ . Thus, using the same approach as that in [7, Lemma 3], we have

$$\frac{1}{N}H(X_{2}^{N}) \ge I(X_{2};Y|X_{r}) - \epsilon' - \epsilon_{4,N},$$
(A33)

where  $\epsilon_{4,N} \to 0$  as  $N \to \infty$ .

For the second term in (A32), since the channel is memoryless, and  $X_1^N$ ,  $X_2^N$ ,  $X_r^N$  are i.i.d. generated, we get

$$\frac{1}{N}I(X_2^N; Z^N | X_1^N, X_r^N) = I(X_2; Z | X_1, X_r).$$
(A34)

Now, we consider the last term of (A32). Given  $Z^N$ ,  $X_1^N$ ,  $X_r^N$  and  $W_2$ , the total number of possible codewords of  $x_2^N$  is  $2^{N(I(X_2;Y|X_r)-\epsilon'-R_2)}$ . By using the Fano's inequality and (A21), we have

$$\frac{1}{N}H(X_2^N|W_2, Z^N, X_1^N, X_r^N) \le \epsilon_{5,N},$$
(A35)

where  $\epsilon_{5,N} \to 0$  as  $N \to \infty$ .

Substituting (A33), (A34) and (A35) into (A32), we have

$$\frac{1}{N}H(W_{2}|W_{1}, Z^{N}) \ge I(X_{2}; Y|X_{r}) - I(X_{2}; Z|X_{1}, X_{r}) - \epsilon^{'} - \epsilon_{4,N} - \epsilon_{5,N}.$$
(A36)

Substituting (A31) and (A36) into (A25), and choosing  $\epsilon'$  and sufficiently large N such that  $3\epsilon' + \epsilon_{1,N} + \epsilon_{2,N} + \epsilon_{3,N} + \epsilon_{4,N} + \epsilon_{5,N} \leq \epsilon$ ,  $\Delta \geq R_1 + R_2 - \epsilon$  for case 1 is proved.

**Proof of**  $\Delta \ge R_1 + R_2 - \epsilon$  for case 2:

$$\Delta = \frac{1}{N} H(W_1, W_2 | Z^N)$$
  
=  $\frac{1}{N} (H(W_1 | Z^N) + H(W_2 | W_1, Z^N)).$  (A37)

The first term in (A37) is bounded as follows.

$$\frac{1}{N}H(W_{1}|Z^{N}) \geq \frac{1}{N}H(W_{1}|Z^{N}, X_{r}^{N}) 
= \frac{1}{N}(H(W_{1}, Z^{N}, X_{r}^{N}) - H(Z^{N}, X_{r}^{N})) 
= \frac{1}{N}(H(W_{1}, Z^{N}, X_{1}^{N}, X_{r}^{N}) - H(X_{1}^{N}|W_{1}, Z^{N}, X_{r}^{N}) - H(Z^{N}, X_{r}^{N})) 
\stackrel{(a)}{=} \frac{1}{N}(H(Z^{N}|X_{1}^{N}, X_{r}^{N}) + H(X_{1}^{N}) + H(X_{r}^{N}) - H(X_{1}^{N}|W_{1}, Z^{N}, X_{r}^{N})) 
- H(Z^{N}|X_{r}^{N}) - H(X_{r}^{N})) 
= \frac{1}{N}(H(X_{1}^{N}) - I(X_{1}^{N}; Z^{N}|X_{r}^{N}) - H(X_{1}^{N}|W_{1}, Z^{N}, X_{r}^{N})), \quad (A38)$$

where (a) follows from  $W_1 \to (X_1^N, X_r^N) \to Z^N$ ,  $H(W_1|X_1^N) = 0$  and the fact that  $X_1^N$  is independent of  $X_r^N$ .

Consider the first term in (A38), the code-book generation of  $x_1^N$  shows that the total number of  $x_1^N$  is  $2^{N(I(X_1;Y|X_2,X_r)-\epsilon')}$ . Thus, using the same approach as that in [7, Lemma 3], we have

$$\frac{1}{N}H(X_{1}^{N}) \ge I(X_{1};Y|X_{2},X_{r}) - \epsilon' - \epsilon_{1,N},$$
(A39)

where  $\epsilon_{1,N} \to 0$  as  $N \to \infty$ .

For the second term in (A38), since the channel is memoryless, and  $X_1^N$ ,  $X_2^N$ ,  $X_r^N$  are i.i.d. generated, we get

$$\frac{1}{N}I(X_1^N;Z^N|X_r^N) = I(X_1;Z|X_r).$$
(A40)

Now, we consider the last term of (A38). Given  $Z^N$ ,  $X_r^N$  and  $W_1$ , the total number of possible codewords of  $x_1^N$  is  $2^{N(I(X_1;Y|X_2,X_r)-\epsilon'-R_1)}$ . By using the Fano's inequality and (A24), we have

$$\frac{1}{N}H(X_1^N|W_1, Z^N, X_r^N) \le \epsilon_{2,N},$$
(A41)

where  $\epsilon_{2,N} \to 0$  as  $N \to \infty$ .

Substituting (A39), (A40) and (A41) into (A38), we have

$$\frac{1}{N}H(W_1|Z^N) \ge I(X_1;Y|X_2,X_r) - I(X_1;Z|X_r) - \epsilon' - \epsilon_{1,N} - \epsilon_{2,N}.$$
(A42)

The second term in (A37) is bounded the same as that for case 1, and thus, we have

$$\frac{1}{N}H(W_{2}|W_{1}, Z^{N}) \ge I(X_{2}; Y|X_{r}) - I(X_{2}; Z|X_{1}, X_{r}) - \epsilon^{'} - \epsilon_{3,N} - \epsilon_{4,N},$$
(A43)

where  $\epsilon_{3,N}, \epsilon_{4,N} \to 0$  as  $N \to \infty$ . The proof is omitted here.

Substituting (A42) and (A43) into (A37), and choosing  $\epsilon'$  and sufficiently large N such that  $2\epsilon' + \epsilon_{1,N} + \epsilon_{2,N} + \epsilon_{3,N} + \epsilon_{4,N} \leq \epsilon$ ,  $\Delta \geq R_1 + R_2 - \epsilon$  for case 2 is proved.

The proof of Theorem 2 is completed.

#### APPENDIX C

## **PROOF OF THEOREM 3**

For Theorem 3, we only need to prove that the corner points of  $\mathcal{L}^3$  and  $\mathcal{L}^4$  are achievable, see the followings.

(Case 1) If I(X<sub>r</sub>; Y) ≥ I(X<sub>r</sub>; Z), we allow the legitimate receiver to decode x<sup>N</sup><sub>r</sub>, and the wiretapper can not decode it. For case 1, it is sufficient to show that the pair (R<sub>1</sub>, R<sub>2</sub>) ∈ L<sup>3</sup> with the condition

$$R_1 = I(X_1; Y, \dot{Y}_r | X_2, X_r) - I(X_1, X_r; Z) + R^*, \quad R_2 = I(X_2; Y, \dot{Y}_r | X_r) - I(X_2; Z | X_1, X_r)$$
(A44)

is achievable. The achievability proof of the other corner point  $(R_1 = I(X_1; Y, \hat{Y}_r | X_r) - I(X_1; Z | X_2, X_r), R_2 = I(X_2; Y, \hat{Y}_r | X_1, X_r) - I(X_2, X_r; Z) + R^*)$  follows by symmetry. Here note that  $R^*$  satisfies

$$\min\{I(X_r; Z|X_1), I(X_r; Z|X_2), I(X_r; Y)\} - R^* \ge I(Y_r; Y_r|X_r).$$
(A45)

(Case 2) If I(Y<sub>r</sub>; Ŷ<sub>r</sub>|X<sub>r</sub>) ≤ I(X<sub>r</sub>; Y) ≤ I(X<sub>r</sub>; Z), we allow both the receivers to decode x<sup>N</sup><sub>r</sub>. For case 2, it is sufficient to show that the pair (R<sub>1</sub>, R<sub>2</sub>) ∈ L<sup>4</sup> with the condition

$$R_1 = I(X_1; Y, \hat{Y}_r | X_2, X_r) - I(X_1; Z | X_r), \quad R_2 = I(X_2; Y, \hat{Y}_r | X_r) - I(X_2; Z | X_1, X_r)$$
(A46)

is achievable. The achievability proof of the other corner point  $(R_1 = I(X_1; Y, \hat{Y}_r | X_r) - I(X_1; Z | X_2, X_r), R_2 = I(X_2; Y, \hat{Y}_r | X_1, X_r) - I(X_2; Z | X_r))$  follows by symmetry.

Fix the joint probability mass function  $P_{\hat{Y}_r|Y_r,X_r}(\hat{y}_r|y_r,x_r)P_{Y,Z,Y_r|X_r,X_1,X_2}(y,z,y_r|x_r,x_1,x_2)P_{X_r}(x_r)P_{X_1}(x_1)P_{X_2}(x_2)$ . Define the messages  $W_1$ ,  $W_2$  taking values in the alphabets  $W_1$ ,  $W_2$ , respectively, where

$$\mathcal{W}_1 = \{1, 2, ..., 2^{NR_1}\}, \ \mathcal{W}_2 = \{1, 2, ..., 2^{NR_2}\}.$$

#### **Code-book Construction for the Two Cases:**

- Code-book construction for case 1:
  - First, generate at random  $2^{N(R_{r1}^* \epsilon^{'})}$  ( $\epsilon^{'}$  is a small positive number) i.i.d. sequences  $x_r^N$  at the relay node each drawn according to  $P_{X_r^N}(x_r^N) = \prod_{i=1}^N P_{X_r}(x_{r,i})$ , index them as  $x_r^N(a)$ ,  $a \in [1, 2^{N(R_{r1}^* \epsilon^{'})}]$ , where

$$R_{r1}^* = \min\{I(X_r; Z|X_1), I(X_r; Z|X_2), I(X_r; Y)\}.$$
(A47)

Here note that

$$R_{r1}^* \ge I(X_r; Z). \tag{A48}$$

For each  $x_r^N(a)$   $(a \in [1, 2^{N(R_{r_1}^* - \epsilon')}])$ , generate at random  $2^{N(R_{r_1}^* - \epsilon' - R^*)}$  i.i.d.  $\hat{y}_r^N$  according to  $P_{\hat{Y}_r^N | X_r^N}(\hat{y}_r^N | x_r^N) = \prod_{i=1}^N P_{\hat{Y}_r | X_r}(\hat{y}_{r,i} | x_{r,i})$ . Label these  $\hat{y}_r^N$  as  $\hat{y}_r^N(m, a)$ ,  $m \in [1, 2^{N(R_{r_1}^* - \epsilon' - R^*)}]$ ,  $a \in [1, 2^{N(R_{r_1}^* - \epsilon')}]$ . Equally divide  $2^{N(R_{r_1}^* - \epsilon')}$  sequences of  $x_r^N$  into  $2^{N(R_{r_1}^* - \epsilon' - R^*)}$  bins, hence there are  $2^{NR^*}$  sequences of  $x_r^N$  at each bin.

- Second, generate  $2^{N(I(X_2;Y,\hat{Y}_r|X_r)-\epsilon')}$  i.i.d. codewords  $x_2^N$  according to  $P_{X_2}(x_2)$ , and divide them into  $2^{NR_2}$  bins. Each bin contains  $2^{N(I(X_2;Y,\hat{Y}_r|X_r)-\epsilon'-R_2)}$  codewords, where

$$I(X_{2}; Y, \hat{Y}_{r} | X_{r}) - \epsilon' - R_{2} = I(X_{2}; Z | X_{1}, X_{r}) - \epsilon'.$$
(A49)

Third, generate 2<sup>N(I(X<sub>1</sub>;Y,Ŷ<sub>r</sub>|X<sub>2</sub>,X<sub>r</sub>)-ϵ'+R\*-R\*1)</sup> i.i.d. codewords x<sub>1</sub><sup>N</sup> according to P<sub>X1</sub>(x<sub>1</sub>), and divide them into 2<sup>NR1</sup> bins. Each bin contains 2<sup>N(I(X<sub>1</sub>;Y,Ŷ<sub>r</sub>|X<sub>2</sub>,X<sub>r</sub>)-ϵ'+R\*-R\*1-R1)</sup> codewords. Here note that from (A45) and (A47), we know that R\* ≤ R\*1, and thus, we have

$$I(X_1; Y, \hat{Y}_r | X_2, X_r) - \epsilon' + R^* - R_{r1}^* \le I(X_1; Y, \hat{Y}_r | X_2, X_r) - \epsilon'.$$
(A50)

In addition, by using  $R_1 = I(X_1; Y, \hat{Y}_r | X_2, X_r) - I(X_1, X_r; Z) + R^*$ , the codewords  $x_1^N$  in each bin is upper bounded by

$$I(X_{1}; Y, \hat{Y}_{r} | X_{2}, X_{r}) - \epsilon' + R^{*} - R_{r1}^{*} - R_{1}$$

$$= I(X_{1}; Y, \hat{Y}_{r} | X_{2}, X_{r}) - \epsilon' + R^{*} - R_{r1}^{*}$$

$$-(I(X_{1}; Y, \hat{Y}_{r} | X_{2}, X_{r}) - I(X_{1}, X_{r}; Z) + R^{*})$$

$$= I(X_{1}, X_{r}; Z) - R_{r1}^{*} - \epsilon'$$

$$\stackrel{(a)}{\leq} I(X_{1}, X_{r}; Z) - I(X_{r}; Z) - \epsilon'$$

$$= I(X_{1}; Z | X_{r}) - \epsilon', \qquad (A51)$$

where (a) is from (A48).

# Code-book Construction for case 2:

- First, generate at random  $2^{N(R_{r2}^*-\epsilon')}$  i.i.d. sequences  $x_r^N$  at the relay node each drawn according to  $P_{X_r^N}(x_r^N) = \prod_{i=1}^N P_{X_r}(x_{r,i})$ , index them as  $x_r^N(a)$ ,  $a \in [1, 2^{N(R_{r2}^*-\epsilon')}]$ , where

$$R_{r2}^* = I(X_r; Y) \le I(X_r; Z).$$
 (A52)

For each  $x_r^N(a)$   $(a \in [1, 2^{N(R_{r_2}^* - \epsilon^{'})}])$ , generate at random  $2^{N(R_{r_2}^* - \epsilon^{'})}$  i.i.d.  $\hat{y}_r^N$  according to  $P_{\hat{Y}_r^N|X_r^N}(\hat{y}_r^N|x_r^N) = \prod_{i=1}^N P_{\hat{Y}_r|X_r}(\hat{y}_{r,i}|x_{r,i})$ . Label these  $\hat{y}_r^N$  as  $\hat{y}_r^N(a)$ ,  $a \in [1, 2^{N(R_{r_2}^* - \epsilon^{'})}]$ .

- Second, generate  $2^{N(I(X_2;Y,\hat{Y}_r|X_r)-\epsilon')}$  i.i.d. codewords  $x_2^N$  according to  $P_{X_2}(x_2)$ , and divide them into  $2^{NR_2}$  bins. Each bin contains  $2^{N(I(X_2;Y,\hat{Y}_r|X_r)-\epsilon'-R_2)}$  codewords, where

$$I(X_{2}; Y, \hat{Y}_{r}|X_{r}) - \epsilon' - R_{2} = I(X_{2}; Z|X_{1}, X_{r}) - \epsilon'.$$
(A53)

- Third, generate  $2^{N(I(X_1;Y,\hat{Y}_r|X_2,X_r)-\epsilon')}$  i.i.d. codewords  $x_1^N$  according to  $P_{X_1}(x_1)$ , and divide them into  $2^{NR_1}$  bins. Each bin contains  $2^{N(I(X_1;Y,\hat{Y}_r|X_2,X_r)-\epsilon'-R_1)}$  codewords, where

$$I(X_1; Y, \hat{Y}_r | X_2, X_r) - \epsilon' - R_1 = I(X_1; Z | X_r) - \epsilon'.$$
(A54)

# **Encoding:**

Encoding involves the mapping of message indices to channel inputs, which are facilitated by the sequences generated above. We exploit the block Markov coding scheme, as argued in [23], the loss induced by this scheme is negligible as the number of blocks  $n \to \infty$ . For block i  $(1 \le i \le n)$ , encoding proceeds as follows.

First, for convenience, the messages  $w_1$  and  $w_2$  transmitted in the *i*-th block are denoted by  $w_{1,i}$  and  $w_{2,i}$ , respectively.  $y_r^N(i)$  and  $\hat{y}_r^N(i)$  are the  $y_r^N$  and  $\hat{y}_r^N$  for the *i*-th block, respectively.

# • Encoding for case 1:

At the end of block i  $(2 \le i \le n)$ , assume that  $(x_r^N(a_i), y_r^N(i), \hat{y}_r^N(m_i, a_i))$  are jointly typical, then we choose  $a_{i+1}$  uniformly from bin  $m_i$ , and the relay sends  $x_r^N(a_{i+1})$  at block i + 1. In the first block, the relay sends  $x_r^N(1)$ .

For a given confidential message  $w_2$ , randomly choose a codeword  $x_2^N$  in bin  $w_2$  to transmit. Similarly, for a given confidential message  $w_1$ , randomly choose a codeword  $x_1^N$  in bin  $w_1$  to transmit.

• Encoding for case 2:

In block i  $(1 \le i \le n)$ , the relay randomly choose an index  $a_i$  from  $[1, 2^{N(R_{r_2}^* - \epsilon')}]$ , and sends  $x_r^N(a_i)$  and  $\hat{y}_r^N(a_i)$ .

For a given confidential message  $w_2$ , randomly choose a codeword  $x_2^N$  in bin  $w_2$  to transmit. Similarly, for a given confidential message  $w_1$ , randomly choose a codeword  $x_1^N$  in bin  $w_1$  to transmit.

# **Decoding:**

# • Decoding for case 1:

(At the relay) At the end of block *i*, the relay already has  $a_i$ , it then decides  $m_i$  by choosing  $m_i$  such that  $(x_r^N(a_i), y_r^N(i), \hat{y}_r^N(m_i, a_i))$  are jointly typical. There exists such  $m_i$ , if

$$R_{r1}^* - R^* \ge I(Y_r; \hat{Y}_r | X_r), \tag{A55}$$

and N is sufficiently large. Choose  $a_{i+1}$  uniformly from bin  $m_i$ .

(At the legitimate receiver) The legitimate receiver does backward decoding. The decoding process starts at the last block n, the legitimate receiver decodes  $a_n$  by choosing unique  $\check{a}_n$  such that  $(x_r^N(\check{a}_n), y^N(n))$  are jointly typical. Since  $R_{r_1}^*$  satisfies (A47), the probability  $Pr\{\check{a}_n = a_n\}$  goes to 1 for sufficiently large N.

Next, the legitimate receiver moves to the block n-1. Now it already has  $\check{a}_n$ , hence we also have  $\check{m}_{n-1} = f(\check{a}_n)$ (here f is a deterministic function, which means that  $\check{m}_{n-1}$  can be determined by  $\check{a}_n$ ). It first declares that  $\check{a}_{n-1}$  is received, if  $\check{a}_{n-1}$  is the unique one such that  $(x_r^N(\check{a}_{n-1}), y^N(n-1))$  are joint typical. If (A47) is satisfied,  $\check{a}_{n-1} = a_{n-1}$  with high probability. After knowing  $\check{a}_{n-1}$ , the destination gets an estimation of  $w_{2,n-1}$  by picking the unique  $\check{w}_{2,n-1}$  such that  $(x_2^N(\check{w}_{2,n-1}), \hat{y}_r^N(\check{m}_{n-1}, \check{a}_{n-1}), y^N(n-1), x_r^N(\check{a}_{n-1}))$  are jointly typical. We will have  $\check{w}_{2,n-1} = w_{2,n-1}$  with high probability, if the codewords of  $x_2^N$  is upper bounded by  $2^{NI(X_2;Y,\hat{Y}_r|X_r)}$  and N is sufficiently large.

After decoding  $\check{w}_{2,n-1}$ , the legitimate receiver tries to find a quintuple such that

 $(x_1^N(\check{w}_{1,n-1}), x_2^N(\check{w}_{2,n-1}), \hat{y}_r^N(\check{m}_{n-1}, \check{a}_{n-1}), y^N(n-1), x_r^N(\check{a}_{n-1}))$  are jointly typical. Based on the AEP, the probability  $Pr\{\check{w}_{1,n-1} = w_{1,n-1}\}$  goes to 1 if the codewords of  $x_1^N$  is upper bounded by  $2^{NI(X_1;Y,\hat{Y}_r|X_2,X_r)}$  and N is sufficiently large.

The decoding scheme of the legitimate receiver in block i  $(1 \le i \le n-2)$  is similar to that in block n-1, and we omit it here.

• Decoding for case 2:

(At the relay) The relay does not need to decode any codeword.

(At the legitimate receiver) In block i  $(1 \le i \le n)$ , the legitimate receiver decodes  $a_i$  by choosing unique  $\check{a}_i$  such that  $(x_r^N(\check{a}_i), y^N(i))$  are jointly typical. Since  $R_{r_2}^*$  satisfies (A52), the probability  $Pr\{\check{a}_i = a_i\}$  goes to 1 for sufficiently large N.

Now since the legitimate receiver has  $\check{a}_i$ , he also knows  $\hat{y}_r^N(\check{a}_i)$ . Then he gets an estimation of  $w_{2,i}$  by picking the unique  $\check{w}_{2,i}$  such that  $(x_2^N(\check{w}_{2,i}), \hat{y}_r^N(\check{a}_i), y^N(i), x_r^N(\check{a}_i))$  are jointly typical. We will have  $\check{w}_{2,i} = w_{2,i}$  with high probability, if the codewords of  $x_2^N$  is upper bounded by  $2^{NI(X_2;Y,\hat{Y}_r|X_r)}$  and N is sufficiently large.

After decoding  $\check{w}_{2,i}$ , the legitimate receiver tries to find a quintuple such that

 $(x_1^N(\check{w}_{1,i}), x_2^N(\check{w}_{2,i}), \hat{y}_r^N(\check{a}_i), y^N(i), x_r^N(\check{a}_i))$  are jointly typical. Based on the AEP, the probability  $Pr\{\check{w}_{1,i} = w_{1,i}\}$  goes to 1 if the codewords of  $x_1^N$  is upper bounded by  $2^{NI(X_1;Y,\hat{Y}_r|X_2,X_r)}$  and N is sufficiently large.

 $P_e \leq \epsilon$  is easy to be checked by using the above encoding-decoding schemes. Now, it remains to prove  $\Delta \geq R_1 + R_2 - \epsilon$  for both cases, see the followings.

# **Equivocation Analysis:**

**Proof of**  $\Delta \ge R_1 + R_2 - \epsilon$  for case 1:

$$\Delta = \frac{1}{N} H(W_1, W_2 | Z^N)$$
  
=  $\frac{1}{N} (H(W_1 | Z^N) + H(W_2 | W_1, Z^N)).$  (A56)

The first term in (A56) is bounded as follows.

$$\frac{1}{N}H(W_{1}|Z^{N}) = \frac{1}{N}(H(W_{1}, Z^{N}) - H(Z^{N}))$$

$$= \frac{1}{N}(H(W_{1}, Z^{N}, X_{1}^{N}, X_{r}^{N}) - H(X_{1}^{N}, X_{r}^{N}|W_{1}, Z^{N}) - H(Z^{N}))$$

$$\stackrel{(a)}{=} \frac{1}{N}(H(Z^{N}|X_{1}^{N}, X_{r}^{N}) + H(X_{1}^{N}) + H(X_{r}^{N}) - H(X_{1}^{N}, X_{r}^{N}|W_{1}, Z^{N}) - H(Z^{N}))$$

$$= \frac{1}{N}(H(X_{1}^{N}) + H(X_{r}^{N}) - I(X_{1}^{N}, X_{r}^{N}; Z^{N}) - H(X_{1}^{N}, X_{r}^{N}|W_{1}, Z^{N})), \quad (A57)$$

where (a) follows from  $W_1 \to (X_1^N, X_r^N) \to Z^N$ ,  $H(W_1|X_1^N) = 0$  and the fact that  $X_1^N$  is independent of  $X_r^N$ . Consider the first term in (A57), the code-book generation of  $x_1^N$  shows that the total number of  $x_1^N$  is upper

bounded by (A51). Thus, using the same approach as that in [7, Lemma 3], we have

$$\frac{1}{N}H(X_{1}^{N}) \ge I(X_{1}; Y, \hat{Y}_{r} | X_{2}, X_{r}) + R^{*} - R_{r1}^{*} - \epsilon^{'} - \epsilon_{1,N},$$
(A58)

where  $\epsilon_{1,N} \to 0$  as  $N \to \infty$ .

For the second term in (A57), the code-book generation of  $x_r^N$  and [7, Lemma 3] guarantee that

$$\frac{1}{N}H(X_{r}^{N}) \ge R_{r1}^{*} - \epsilon^{'} - \epsilon_{2,N},$$
(A59)

where  $\epsilon_{2,N} \to 0$  as  $N \to \infty$ .

For the third term in (A57), since the channel is memoryless, and  $X_1^N$ ,  $X_2^N$ ,  $X_r^N$  are i.i.d. generated, we get

$$\frac{1}{N}I(X_1^N, X_r^N; Z^N) = I(X_1, X_r; Z).$$
(A60)

Now, we consider the last term of (A57). Given  $w_1$ , the wiretapper can do joint decoding. Specifically, given  $z^N$  and  $w_1$ ,

$$\frac{1}{N}H(X_1^N, X_r^N | W_1, Z^N) \le \epsilon_{3,N}$$
(A61)

 $(\epsilon_{3,N} \to 0 \text{ as } N \to \infty)$  is guaranteed if  $R_r \leq I(X_r; Z|X_1)$  and  $I(X_1; Y, \hat{Y}_r|X_2, X_r) - \epsilon' + R^* - R^*_{r1} - R_1 \leq I(X_1; Z|X_r)$ , and this is from the properties of AEP (similar argument is used in the proof of [18, Theorem 3]). By checking (A47) and (A51), (A61) is obtained.

Substituting (A58), (A59), (A60) and (A61) into (A57), we have

$$\frac{1}{N}H(W_{1}|Z^{N}) \ge I(X_{1};Y,\hat{Y}_{r}|X_{2},X_{r}) + R^{*} - I(X_{1},X_{r};Z) - 2\epsilon^{'} - \epsilon_{1,N} - \epsilon_{2,N} - \epsilon_{3,N}.$$
(A62)

The second term in (A56) is bounded as follows.

$$\frac{1}{N}H(W_{2}|W_{1}, Z^{N}) \geq \frac{1}{N}H(W_{2}|W_{1}, Z^{N}, X_{1}^{N}, X_{r}^{N}) 
\stackrel{(1)}{=} \frac{1}{N}H(W_{2}|Z^{N}, X_{1}^{N}, X_{r}^{N}) 
= \frac{1}{N}(H(W_{2}, Z^{N}, X_{1}^{N}, X_{r}^{N}) - H(Z^{N}, X_{1}^{N}, X_{r}^{N})) 
= \frac{1}{N}(H(W_{2}, Z^{N}, X_{1}^{N}, X_{r}^{N}, X_{2}^{N}) - H(X_{2}^{N}|W_{2}, Z^{N}, X_{1}^{N}, X_{r}^{N}) - H(Z^{N}, X_{1}^{N}, X_{r}^{N})) 
\stackrel{(2)}{=} \frac{1}{N}(H(Z^{N}|X_{1}^{N}, X_{2}^{N}, X_{r}^{N}) + H(X_{r}^{N}) + H(X_{1}^{N}) + H(X_{2}^{N}) 
- H(X_{2}^{N}|W_{2}, Z^{N}, X_{1}^{N}, X_{r}^{N}) - H(Z^{N}|X_{1}^{N}, X_{r}^{N}) - H(X_{1}^{N}) - H(X_{1}^{N})) 
= \frac{1}{N}(H(X_{2}^{N}) - I(X_{2}^{N}; Z^{N}|X_{1}^{N}, X_{r}^{N}) - H(X_{2}^{N}|W_{2}, Z^{N}, X_{1}^{N}, X_{r}^{N})), \quad (A63)$$

where (1) is from the Markov chain  $W_1 \to (Z^N, X_1^N, X_r^N) \to W_2$ , and (2) is from the Markov chain  $W_2 \to (X_1^N, X_2^N, X_r^N) \to Z^N$ ,  $H(W_2|X_2^N) = 0$ , and the fact that  $X_1^N, X_2^N$  and  $X_r^N$  are independent.

Consider the first term in (A63), using the same approach as that in [7, Lemma 3], we have

$$\frac{1}{N}H(X_{2}^{N}) \ge I(X_{2}; Y, \hat{Y}_{r}|X_{r}) - \epsilon^{'} - \epsilon_{4,N},$$
(A64)

where  $\epsilon_{4,N} \to 0$  as  $N \to \infty$ .

For the second term in (A63), since the channel is memoryless, and  $X_1^N$ ,  $X_2^N$ ,  $X_r^N$  are i.i.d. generated, we get

$$\frac{1}{N}I(X_2^N; Z^N | X_1^N, X_r^N) = I(X_2; Z | X_1, X_r).$$
(A65)

Now, we consider the last term of (A63). Given  $Z^N$ ,  $X_1^N$ ,  $X_r^N$  and  $W_2$ , the total number of possible codewords of  $x_2^N$  is  $2^{N(I(X_2;Y,\hat{Y}_r|X_r)-\epsilon'-R_2)}$ . By using the Fano's inequality and (A49), we have

$$\frac{1}{N}H(X_2^N|W_2, Z^N, X_1^N, X_r^N) \le \epsilon_{5,N}.$$
(A66)

Substituting (A64), (A65) and (A66) into (A63), we have

$$\frac{1}{N}H(W_2|W_1, Z^N) \ge I(X_2; Y, \hat{Y}_r|X_r) - I(X_2; Z|X_1, X_r) - \epsilon' - \epsilon_{4,N} - \epsilon_{5,N}.$$
(A67)

Substituting (A62) and (A67) into (A56), and choosing  $\epsilon'$  and sufficiently large N such that  $3\epsilon' + \epsilon_{1,N} + \epsilon_{2,N} + \epsilon_{3,N} + \epsilon_{4,N} + \epsilon_{5,N} \leq \epsilon$ ,  $\Delta \geq R_1 + R_2 - \epsilon$  for case 1 is proved.

Proof of  $\Delta \ge R_1 + R_2 - \epsilon$  for case 2:

$$\Delta = \frac{1}{N} H(W_1, W_2 | Z^N)$$
  
=  $\frac{1}{N} (H(W_1 | Z^N) + H(W_2 | W_1, Z^N)).$  (A68)

The first term in (A68) is bounded as follows.

$$\frac{1}{N}H(W_{1}|Z^{N}) \geq \frac{1}{N}H(W_{1}|Z^{N}, X_{r}^{N}) 
= \frac{1}{N}(H(W_{1}, Z^{N}, X_{r}^{N}) - H(Z^{N}, X_{r}^{N})) 
= \frac{1}{N}(H(W_{1}, Z^{N}, X_{1}^{N}, X_{r}^{N}) - H(X_{1}^{N}|W_{1}, Z^{N}, X_{r}^{N}) - H(Z^{N}, X_{r}^{N})) 
\stackrel{(a)}{=} \frac{1}{N}(H(Z^{N}|X_{1}^{N}, X_{r}^{N}) + H(X_{1}^{N}) + H(X_{r}^{N}) - H(X_{1}^{N}|W_{1}, Z^{N}, X_{r}^{N}) 
- H(Z^{N}|X_{r}^{N}) - H(X_{r}^{N})) 
= \frac{1}{N}(H(X_{1}^{N}) - I(X_{1}^{N}; Z^{N}|X_{r}^{N}) - H(X_{1}^{N}|W_{1}, Z^{N}, X_{r}^{N})), \quad (A69)$$

where (a) follows from  $W_1 \to (X_1^N, X_r^N) \to Z^N$ ,  $H(W_1|X_1^N) = 0$  and the fact that  $X_1^N$  is independent of  $X_r^N$ .

Consider the first term in (A69), the code-book generation of  $x_1^N$  shows that the total number of  $x_1^N$  is  $2^{N(I(X_1;Y,\hat{Y}_r|X_2,X_r)-\epsilon')}$ . Thus, using the same approach as that in [7, Lemma 3], we have

$$\frac{1}{N}H(X_{1}^{N}) \ge I(X_{1}; Y, \hat{Y}_{r} | X_{2}, X_{r}) - \epsilon' - \epsilon_{1,N},$$
(A70)

where  $\epsilon_{1,N} \to 0$  as  $N \to \infty$ .

For the second term in (A69), since the channel is memoryless, and  $X_1^N$ ,  $X_2^N$ ,  $X_r^N$  are i.i.d. generated, we get

$$\frac{1}{N}I(X_1^N; Z^N | X_r^N) = I(X_1; Z | X_r).$$
(A71)

Now, we consider the last term of (A69). Given  $Z^N$ ,  $X_r^N$  and  $W_1$ , the total number of possible codewords of  $x_1^N$  is  $2^{N(I(X_1;Y,\hat{Y}_r|X_2,X_r)-\epsilon'-R_1)}$ . By using the Fano's inequality and (A54), we have

$$\frac{1}{N}H(X_1^N|W_1, Z^N, X_r^N) \le \epsilon_{2,N},$$
(A72)

where  $\epsilon_{2,N} \to 0$  as  $N \to \infty$ .

Substituting (A70), (A71) and (A72) into (A69), we have

$$\frac{1}{N}H(W_1|Z^N) \ge I(X_1; Y, \hat{Y}_r|X_2, X_r) - I(X_1; Z|X_r) - \epsilon' - \epsilon_{1,N} - \epsilon_{2,N}.$$
(A73)

The second term in (A68) is bounded the same as that for case 1, and thus, we have

$$\lim_{N \to \infty} \frac{1}{N} H(W_2|W_1, Z^N) \ge I(X_2; Y, \hat{Y}_r | X_r) - I(X_2; Z | X_1, X_r) - \epsilon' - \epsilon_{3,N} - \epsilon_{4,N}.$$
(A74)

The proof is omitted here.

Substituting (A73) and (A74) into (A68), and choosing  $\epsilon'$  and sufficiently large N such that  $2\epsilon' + \epsilon_{1,N} + \epsilon_{2,N} + \epsilon_{3,N} + \epsilon_{4,N} \le \epsilon$ ,  $\Delta \ge R_1 + R_2 - \epsilon$  for case 2 is proved.

The proof of Theorem 3 is completed.

# APPENDIX D

# **PROOF OF THEOREM 4**

In this section, we prove Theorem 4: all the achievable secrecy pairs  $(R_1, R_2)$  of the degraded discrete memoryless MARC-WT are contained in the set  $\mathcal{R}^{ddo}$ . We will prove the inequalities of Theorem 4 in the remainder of this section.

(Proof of  $R_1 \leq I(X_1, X_r; Y|X_2, U) - I(X_1; Z|U)$ ):

$$\begin{split} R_{1} - \epsilon &= \frac{1}{N} H(W_{1}) - \epsilon^{\binom{1}{2}} \frac{1}{N} H(W_{1}|Z^{N}) \\ &\stackrel{(2)}{\leq} \frac{1}{N} (H(W_{1}|Z^{N}) - H(W_{1}|Z^{N}, W_{2}, Y^{N}, X_{2}^{N}) + \delta(P_{e})) \\ &\stackrel{(3)}{\equiv} \frac{1}{N} (H(W_{1}|Z^{N}) - H(W_{1}|Z^{N}, Y^{N}, X_{2}^{N}) + \delta(P_{e})) \\ &= \frac{1}{N} (I(W_{1};Y^{N}, X_{2}^{N}|Z^{N}) - H(W_{1}, X_{2}^{N}, W_{1}, X_{1}^{N}) + \delta(P_{e})) \\ &\stackrel{(4)}{\leq} \frac{1}{N} (H(Y^{N}, X_{2}^{N}|Z^{N}) - H(Y^{N}, X_{2}^{N}|Z^{N}, W_{1}, X_{1}^{N}) + \delta(P_{e})) \\ &\stackrel{(4)}{=} \frac{1}{N} (I(Y^{N}, X_{2}^{N}|Z^{N}) - H(Y^{N}, X_{2}^{N}|Z^{N}, X_{1}^{N}) + \delta(P_{e})) \\ &= \frac{1}{N} (I(Y^{N}, X_{2}^{N}; X_{1}^{N}|Z^{N}) - H(Y^{N}, X_{2}^{N}|Z^{N}, X_{1}^{N}) + \delta(P_{e})) \\ &\stackrel{(5)}{=} \frac{1}{N} (H(X_{1}^{N}|Z^{N}) - H(X_{1}^{N}|Z^{N}, Y^{N}, X_{2}^{N}) - H(X_{1}^{N}) + H(X_{1}^{N}|X_{2}^{N}) + \delta(P_{e})) \\ &\stackrel{(6)}{=} \frac{1}{N} (I(X_{1}^{N}, X_{r}^{N}; Y^{N}|X_{2}^{N}) - I(X_{1}^{N}; Z^{N}) + \delta(P_{e})) \\ &\leq \frac{1}{N} (I(X_{1}^{N}, X_{r}^{N}; Y^{N}|X_{2}^{N}) - I(X_{1}^{N}; Z^{N}) + \delta(P_{e})) \\ &= \frac{1}{N} \sum_{i=1}^{N} (H(Y_{i}|Y^{i-1}, X_{2}^{N}) - H(Y_{i}|X_{1,i}, X_{2,i}, X_{r,i}, Z^{i-1}) - H(Z_{i}|Z^{i-1}, X_{1}^{N})) + \frac{\delta(P_{e})}{N} \\ &\stackrel{(7)}{=} \frac{1}{N} \sum_{i=1}^{N} (H(Y_{i}|Y^{i-1}, X_{2}^{N}, Z^{i-1}) - H(Y_{i}|X_{1,i}, X_{2,i}, X_{r,i}, Z^{i-1}) - H(Z_{i}|Z^{i-1}) + H(Z_{i}|Z^{i-1}, X_{1,i})) + \frac{\delta(P_{e})}{N} \\ &\leq \frac{1}{N} \sum_{i=1}^{N} (H(Y_{i}|X_{2,i}, Z^{i-1}) - H(Y_{i}|X_{1,i}, X_{2,i}, X_{r,i}, Z^{i-1}) - H(Z_{i}|Z^{i-1}, X_{1,i})) + \frac{\delta(P_{e})}{N} \\ &\stackrel{(8)}{=} \frac{1}{N} \sum_{i=1}^{N} (H(Y_{i}|X_{2,i}, Z^{i-1}, J = i) - H(Y_{i}|X_{1,i}, X_{2,i}, X_{r,i}, Z^{i-1}, J = i) - H(Z_{i}|Z^{i-1}, J = i) \\ \end{aligned}$$

$$+H(Z_{i}|Z^{i-1}, X_{1,i}, J = i)) + \frac{\delta(P_{e})}{N}$$

$$\stackrel{(9)}{=} H(Y_{J}|X_{2,J}, Z^{J-1}, J) - H(Y_{J}|X_{1,J}, X_{2,J}, X_{r,J}, Z^{J-1}, J) - H(Z_{J}|Z^{J-1}, J) + H(Z_{J}|Z^{J-1}, X_{1,J}, J) + \frac{\delta(P_{e})}{N}$$

$$\stackrel{(10)}{=} I(X_{1}, X_{r}; Y|X_{2}, U) - I(X_{1}; Z|U) + \frac{\delta(P_{e})}{N},$$
(A75)

where (1) is from the fact that the secrecy requirement on the full message set also ensures the secrecy of individual message (see (2.3)), (2) is from the Fanos inequality, (3) is from  $H(W_2|X_2^N) = 0$ , (4) is from  $H(W_1|X_1^N) = 0$ , (5) and (6) are from the fact that the wiretap channel is degraded, which implies the Markov chain  $X_1^N \to (X_2^N, Y^N) \to (X_2^N, Y^N)$  $Z^N$ , and from the fact that  $X_1^N$  is independent of  $X_2^N$ , (7) is from the Markov chains  $Y_i \to (Y^{i-1}, X_2^N) \to Z^{i-1}$ and  $Y_i \rightarrow (X_{1,i}, X_{2,i}, X_{r,i}) \rightarrow Z^{i-1}$  (these Markov chains are also from the fact that the wiretap channel is degraded), (8) is from J is a random variable (uniformly distributed over  $\{1, 2, ..., N\}$ ), and it is independent of  $X_1^N$ ,  $X_2^N$ ,  $X_r^N$ ,  $Y^N$  and  $Z^N$ , (9) is from J is uniformly distributed over  $\{1, 2, ..., N\}$ , and (10) is from the definitions that  $X_1 \triangleq X_{1,J}, X_2 \triangleq X_{2,J}, X_r \triangleq X_{r,J}, Y \triangleq Y_J, Z \triangleq Z_J$  and  $U \triangleq (Z^{J-1}, J)$ .

By using  $P_e \leq \epsilon$  and letting  $\epsilon \to 0$ ,  $R_1 \leq I(X_1, X_r; Y | X_2, U) - I(X_1; Z | U)$  is proved.

(**Proof of**  $R_2 \leq I(X_2, X_r; Y | X_1, U) - I(X_2; Z | U)$ ):

(

The proof is analogous to the proof of  $R_1 \leq I(X_1, X_r; Y | X_2, U) - I(X_1; Z | U)$ , and it is omitted here. **Proof of**  $R_1 + R_2 \leq I(X_1, X_2, X_r; Y|U) - I(X_1, X_2; Z|U)$ :

$$\begin{split} R_{1} + R_{2} - \epsilon & \stackrel{(4)}{\leq} & \Delta = \frac{1}{N} H(W_{1}, W_{2} | Z^{N}) \\ & \stackrel{(2)}{\leq} & \frac{1}{N} (H(W_{1}, W_{2} | Z^{N}) + \delta(P_{e}) - H(W_{1}, W_{2} | Y^{N}, Z^{N})) \\ & \leq & \frac{1}{N} (H(Y^{N} | Z^{N}) - H(Y^{N} | Z^{N}, W_{1}, W_{2}, X_{1}^{N}, X_{2}^{N}) + \delta(P_{e})) \\ & \stackrel{(3)}{=} & \frac{1}{N} (H(Y^{N} | Z^{N}) - H(Y^{N} | Z^{N}, X_{1}^{N}, X_{2}^{N}) + \delta(P_{e})) \\ & = & \frac{1}{N} (I(X_{1}^{N}, X_{2}^{N}; Y^{N}) - I(X_{1}^{N}, X_{2}^{N}; Z^{N}) + \delta(P_{e})) \\ & \leq & \frac{1}{N} (I(X_{1}^{N}, X_{2}^{N}, X_{r}^{N}; Y^{N}) - I(X_{1}^{N}, X_{2}^{N}; Z^{N}) + \delta(P_{e})) \\ & \stackrel{(4)}{=} & \frac{1}{N} \sum_{i=1}^{N} (H(Y_{i} | Y^{i-1}) - H(Y_{i} | X_{1,i}, X_{2,i}, X_{r,i}, Z^{i-1}) - H(Z_{i} | Z^{i-1}) + H(Z_{i} | X_{1,i}, X_{2,i}, Z^{i-1})) + \frac{\delta(P_{e})}{N} \\ & \stackrel{(6)}{\leq} & \frac{1}{N} \sum_{i=1}^{N} (H(Y_{i} | Z^{i-1}) - H(Y_{i} | X_{1,i}, X_{2,i}, X_{r,i}, Z^{i-1}) - H(Z_{i} | Z^{i-1}) + H(Z_{i} | X_{1,i}, X_{2,i}, Z^{i-1})) + \frac{\delta(P_{e})}{N} \\ & \stackrel{(6)}{=} & \frac{1}{N} \sum_{i=1}^{N} (H(Y_{i} | Z^{i-1}, J = i) - H(Y_{i} | X_{1,i}, X_{2,i}, X_{r,i}, Z^{i-1}, J = i) \\ & -H(Z_{i} | Z^{i-1}, J ) - H(Y_{j} | X_{1,j}, X_{2,j}, Z^{j-1}, J) \\ & -H(Z_{j} | Z^{J-1}, J) - H(Y_{j} | X_{1,j}, X_{2,j}, Z^{J-1}, J) + \frac{\delta(P_{e})}{N} \\ \end{array}$$

$$\stackrel{(8)}{\leq} I(X_1, X_2, X_r; Y|U) - I(X_1, X_2; Z|U) + \frac{\delta(\epsilon)}{N},$$
(A76)

where (1) is from (2.2), (2) is from the Fanos inequality, (3) is from  $(W_1, W_2) \rightarrow (X_1^N, X_2^N, Z^N) \rightarrow Y^N$ , (4) is from  $Y_i \rightarrow (X_{1,i}, X_{2,i}, X_{r,i}) \rightarrow Z^{i-1}$ , (5) is from  $Y_i \rightarrow Y^{i-1} \rightarrow Z^{i-1}$ , (6) is from J is a random variable (uniformly distributed over  $\{1, 2, ..., N\}$ ), and it is independent of  $X_1^N, X_2^N, X_r^N, Y^N$  and  $Z^N$ , (7) is from J is uniformly distributed over  $\{1, 2, ..., N\}$ , and (8) is from the definitions that  $X_1 \triangleq X_{1,J}, X_2 \triangleq X_{2,J}, X_r \triangleq X_{r,J},$  $Y \triangleq Y_J, Z \triangleq Z_J$  and  $U \triangleq (Z^{J-1}, J)$ , and the fact that  $P_e \leq \epsilon$ .

Letting  $\epsilon \to 0$ ,  $R_1 + R_2 \leq I(X_1, X_2, X_r; Y|U) - I(X_1, X_2; Z|U)$  is proved.

The proof of Theorem 4 is completed.

## APPENDIX E

#### **PROOF OF THEOREM 8**

Since  $N_2 \ge N_1$ , the GMARC-WT reduces to a kind of degraded MARC-WT with the Markov chain  $(X_1, X_2, X_r, Y_r) \rightarrow Y \rightarrow Z$ , and thus the outer bound  $\mathcal{R}^{gout}$  can be obtained from Theorem 4. The details are as follows.

From (A75), we know that

$$R_{1} \leq \frac{1}{N} \sum_{i=1}^{N} (h(Y_{i}|X_{2,i}, Z^{i-1}) - h(Y_{i}|X_{1,i}, X_{2,i}, X_{r,i}, Z^{i-1}) - h(Z_{i}|Z^{i-1}) + h(Z_{i}|Z^{i-1}, X_{1,i})) + \frac{\delta(P_{e})}{N}.$$
(A77)

Analogously,

$$R_{2} \leq \frac{1}{N} \sum_{i=1}^{N} (h(Y_{i}|X_{1,i}, Z^{i-1}) - h(Y_{i}|X_{1,i}, X_{2,i}, X_{r,i}, Z^{i-1}) - h(Z_{i}|Z^{i-1}) + h(Z_{i}|Z^{i-1}, X_{2,i})) + \frac{\delta(P_{e})}{N}.$$
(A78)

From (A76), we have

$$R_{1} + R_{2} \leq \frac{1}{N} \sum_{i=1}^{N} (h(Y_{i}|Z^{i-1}) - h(Y_{i}|X_{1,i}, X_{2,i}, X_{r,i}, Z^{i-1}) - h(Z_{i}|Z^{i-1}) + h(Z_{i}|X_{1,i}, X_{2,i}, Z^{i-1})) + \frac{\delta(P_{e})}{N}).$$
(A79)

It remains to bound the conditional entropies in (A77), (A78) and (A79), see the followings. First note that

$$\frac{1}{N} \sum_{i=1}^{N} h(Z_i | X_{1,i}, X_{2,i}, Z^{i-1}) \leq \frac{1}{N} \sum_{i=1}^{N} h(Z_i | X_{1,i}, X_{2,i}) \\
\stackrel{(1)}{\leq} \frac{1}{N} \sum_{i=1}^{N} h(Z_{2,i} + X_{r,i}) \\
\leq \frac{1}{N} \sum_{i=1}^{N} \frac{1}{2} \log 2\pi e(E[X_{r,i}^2] + N_2)$$

32

$$\stackrel{(2)}{\leq} \frac{1}{2} \log 2\pi e(\frac{1}{N} \sum_{i=1}^{N} E[X_{r,i}^2] + N_2)$$

$$\leq \frac{1}{2} \log 2\pi e(P_r + N_2),$$
(A80)

where (1) is from  $Z_i = X_{1,i} + X_{2,i} + X_{r,i} + Z_{2,i}$ , and (2) is from Jensen's inequality.

On the other hand,

$$\frac{1}{N} \sum_{i=1}^{N} h(Z_i | X_{1,i}, X_{2,i}, Z^{i-1}) \geq \frac{1}{N} \sum_{i=1}^{N} h(Z_i | X_{1,i}, X_{2,i}, X_{r,i}, Z^{i-1}) \\
\stackrel{(a)}{=} \frac{1}{N} \sum_{i=1}^{N} h(Z_i | X_{1,i}, X_{2,i}, X_{r,i}) \\
= \frac{1}{N} \sum_{i=1}^{N} h(Z_{2,i}) \\
= \frac{1}{N} \sum_{i=1}^{N} \frac{1}{2} \log 2\pi e N_2 = \frac{1}{2} \log 2\pi e N_2,$$
(A81)

where (a) is from the Markov chain  $Z^{i-1} \to (X_{1,i}, X_{2,i}, X_{r,i}) \to Z_i$ .

Combining (A80) and (A81), we establish that there exists some  $\alpha \in [0,1]$  such that

$$\frac{1}{N}\sum_{i=1}^{N}h(Z_i|X_{1,i}, X_{2,i}, Z^{i-1}) = \frac{1}{2}\log 2\pi e(\alpha P_r + N_2).$$
(A82)

Second, since

$$\frac{1}{N} \sum_{i=1}^{N} h(Z_i | X_{1,i}, Z^{i-1}) \geq \frac{1}{N} \sum_{i=1}^{N} h(Z_i | X_{1,i}, X_{2,i}, Z^{i-1}) \\
= \frac{1}{2} \log 2\pi e(\alpha P_r + N_2),$$
(A83)

and

$$\frac{1}{N} \sum_{i=1}^{N} h(Z_i | X_{1,i}, Z^{i-1}) \leq \frac{1}{N} \sum_{i=1}^{N} h(Z_i | X_{1,i}) \\
\leq \frac{1}{N} \sum_{i=1}^{N} h(Z_{2,i} + X_{2,i} + X_{r,i}) \\
\leq \frac{1}{N} \sum_{i=1}^{N} \frac{1}{2} \log 2\pi e(E[X_{r,i}^2] + E[X_{2,i}^2] + N_2) \\
\leq \frac{1}{2} \log 2\pi e(\frac{1}{N} \sum_{i=1}^{N} E[X_{r,i}^2] + \frac{1}{N} \sum_{i=1}^{N} E[X_{2,i}^2] + N_2) \\
\leq \frac{1}{2} \log 2\pi e(P_r + P_2 + N_2),$$
(A84)

we establish that there exists some  $\beta_1 \in [0,1]$  such that

$$\frac{1}{N} \sum_{i=1}^{N} h(Z_i | X_{1,i}, Z^{i-1}) = \frac{1}{2} \log 2\pi e(\alpha P_r + N_2 + \beta_1 (P_r + P_2 + N_2 - \alpha P_r - N_2))$$
  
=  $\frac{1}{2} \log 2\pi e(N_2 + P_r(\alpha + \beta_1 - \alpha\beta_1) + \beta_1 P_2).$  (A85)

Third, analogously, there exists some  $\beta_2 \in [0,1]$  such that

$$\frac{1}{N}\sum_{i=1}^{N}h(Z_i|X_{2,i}, Z^{i-1}) = \frac{1}{2}\log 2\pi e(N_2 + P_r(\alpha + \beta_2 - \alpha\beta_2) + \beta_2 P_1).$$
(A86)

Fourth, since

$$\frac{1}{N} \sum_{i=1}^{N} h(Z_i | Z^{i-1}) \geq \frac{1}{N} \sum_{i=1}^{N} h(Z_i | X_{1,i}, Z^{i-1}) \\
= \frac{1}{2} \log 2\pi e(N_2 + P_r(\alpha + \beta_1 - \alpha\beta_1) + \beta_1 P_2),$$
(A87)

$$\frac{1}{N} \sum_{i=1}^{N} h(Z_i | Z^{i-1}) \geq \frac{1}{N} \sum_{i=1}^{N} h(Z_i | X_{2,i}, Z^{i-1}) \\ = \frac{1}{2} \log 2\pi e (N_2 + P_r(\alpha + \beta_2 - \alpha\beta_2) + \beta_2 P_1)$$
(A88)

and

$$\frac{1}{N} \sum_{i=1}^{N} h(Z_i | Z^{i-1}) \leq \frac{1}{N} \sum_{i=1}^{N} h(Z_i) \\
= \frac{1}{N} \sum_{i=1}^{N} h(Z_{2,i} + X_{1,i} + X_{2,i} + X_{r,i}) \\
\leq \frac{1}{N} \sum_{i=1}^{N} \frac{1}{2} \log 2\pi e(E[X_{r,i}^2] + E[X_{1,i}^2] + E[X_{2,i}^2] + N_2) \\
\leq \frac{1}{2} \log 2\pi e(\frac{1}{N} \sum_{i=1}^{N} E[X_{r,i}^2] + \frac{1}{N} \sum_{i=1}^{N} E[X_{1,i}^2] + \frac{1}{N} \sum_{i=1}^{N} E[X_{2,i}^2] + N_2) \\
\leq \frac{1}{2} \log 2\pi e(P_r + P_1 + P_2 + N_2), \quad (A89)$$

there exists some  $\gamma \in [0,1]$  such that

$$\frac{1}{N}\sum_{i=1}^{N}h(Z_i|Z^{i-1}) = \frac{1}{2}\log 2\pi e(C + \gamma(P_r + P_1 + P_2 + N_2 - C)),$$
(A90)

where C is given by

$$C = \max\{N_2 + P_r(\alpha + \beta_1 - \alpha\beta_1) + \beta_1 P_2, N_2 + P_r(\alpha + \beta_2 - \alpha\beta_2) + \beta_2 P_1\}.$$
 (A91)

Fifth, by using the entropy power inequality, we have

$$2^{2h(Z_{i}|X_{1,i},Z^{i-1})} \stackrel{(1)}{=} 2^{2h(Y_{i}+Z'_{2,i}|X_{1,i},Z^{i-1})} \\ \stackrel{(2)}{\geq} 2^{2h(Y_{i}|X_{1,i},Z^{i-1})} + 2^{2h(Z'_{2,i}|X_{1,i},Z^{i-1})} \\ \stackrel{(3)}{=} 2^{2h(Y_{i}|X_{1,i},Z^{i-1})} + 2^{2h(Z'_{2,i})},$$
(A92)

where (1) is from the definition that  $Z'_{2,i} = Z_{2,i} - Z_{1,i}$ , (2) is from the entropy power inequality, and (3) is from  $Z'_{2,i}$  is independent of  $X_{1,i}$  and  $Z^{i-1}$ .

Substituting  $h(Z'_{2,i}) = \frac{1}{2} \log 2\pi e(N_2 - N_1)$  and (A85) into (A92), and using Jensen's inequality, we have

$$\frac{1}{N}\sum_{i=1}^{N}h(Y_i|X_{1,i}, Z^{i-1}) \le \frac{1}{2}\log 2\pi e(P_r(\alpha + \beta_1 - \alpha\beta_1) + \beta_1 P_2 + N_1).$$
(A93)

Analogously, we have

$$\frac{1}{N}\sum_{i=1}^{N}h(Y_i|X_{2,i}, Z^{i-1}) \le \frac{1}{2}\log 2\pi e(P_r(\alpha + \beta_2 - \alpha\beta_2) + \beta_2 P_1 + N_1),$$
(A94)

and

$$\frac{1}{N}\sum_{i=1}^{N}h(Y_i|Z^{i-1}) \le \frac{1}{2}\log 2\pi e(C + \gamma(P_r + P_1 + P_2 + N_1 - C)).$$
(A95)

Finally, note that

$$h(Y_i|X_{1,i}, X_{2,i}, X_{r,i}, Z^{i-1}) = h(Z_{1,i}|X_{1,i}, X_{2,i}, X_{r,i}, Z^{i-1})$$

$$\stackrel{(1)}{=} h(Z_{1,i}) = \frac{1}{2}\log 2\pi e N_1,$$
(A96)

where (1) is from  $Z_{1,i}$  is independent of  $X_{1,i}$ ,  $X_{2,i}$ ,  $X_{r,i}$  and  $Z^{i-1}$ .

Substituting (A82), (A85), (A86), (A90), (A93), (A94), (A95) and (A96) into (A77), (A78) and (A79), using the fact that  $P_e \leq \epsilon$  and letting  $\epsilon \to 0$ , Theorem 8 is proved.

#### REFERENCES

- [1] A. D. Wyner, "The wire-tap channel," The Bell System Technical Journal, vol. 54, no. 8, pp. 1355-1387, 1975.
- [2] S. K. Leung-Yan-Cheong, M. E. Hellman, "The Gaussian wire-tap channel," IEEE Trans Inf Theory, vol. IT-24, no. 4, pp. 451-456, July 1978.
- [3] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans Inf Theory*, vol. IT-24, no. 3, pp. 339-348, May 1978.
- [4] J. Körner and K. Marton, "General broadcast channels with degraded message sets," *IEEE Trans Inf Theory*, vol. IT-23, no. 1, pp. 60-64, January 1977.
- [5] R. Liu, I. Maric, P. Spasojevic and R.D Yates, "Discrete memoryless interference and broadcast channels with confidential messages: secrecy rate regions," *IEEE Trans Inf Theory*, vol. IT-54, no. 6, pp. 2493-2507, Jun. 2008.
- [6] J. Xu, Y. Cao, and B. Chen, "Capacity bounds for broadcast channels with confidential messages," *IEEE Trans Inf Theory*, vol. IT-55, no. 6, pp. 4529-4542. 2009.
- [7] Y. Liang and H. V. Poor, "Multiple-access channels with confidential messages," *IEEE Trans Inf Theory*, vol. IT-54, no. 3, pp. 976-1002, Mar. 2008.
- [8] E. Tekin and A. Yener, "The Gaussian multiple access wire-tap channel," *IEEE Trans Inf Theory*, vol. IT-54, no. 12, pp. 5747-5755, Dec. 2008.
- [9] E. Tekin and A. Yener, "The general Gaussian multiple access and two-way wire-tap channels: Achievable rates and cooperative jamming," IEEE Trans Inf Theory, vol. IT-54, no. 6, pp. 2735-2751, June 2008.
- [10] E. Ekrem and S. Ulukus, "On the secrecy of multiple access wiretap channel," in Proc. Annual Allerton Conf. on Communications, Control and Computing, Monticello, IL, Sept. 2008.
- [11] R. Bassily and S. Ulukus, "A New Achievable Ergodic Secrecy Rate Region for the Fading Multiple Access Wiretap Channel," in Proc. Annual Allerton Conf. on Communications, Control and Computing, Monticello, IL, Sept. 2009.
- [12] M. Wiese and H. Boche, "An Achievable Region for the Wiretap Multiple-Access Channel with Common Message," Proceedings of 2012 IEEE International Symposium on Information Theory, 2012.

- [13] M. Wiese and H. Boche, "Strong secrecy for multiple access channels," Information Theory, Combinatorics, and Search Theory, Springer Berlin Heidelberg, pp. 71-122, 2013.
- [14] M. H. Yassaee and M. R. Aref, "Multiple access wiretap channels with strong secrecy," Proceedings of IEEE Information Theory Workshop, 2010.
- [15] X. He, A. Khisti, and A. Yener, "MIMO Multiple Access Channel With an Arbitrarily Varying Eavesdropper: Secrecy Degrees of Freedom," IEEE Trans Inf Theory, vol. IT-59, no. 8, pp. 4733-4745, 2013.
- [16] P. Xu, Z. Ding, and X. Dai, "Rate Regions for Multiple Access Channel With Conference and Secrecy Constraints," *IEEE Trans. Inf. Forensics and Security*, vol. 8, no. 12, pp. 1961-1974, 2013.
- [17] Y. Liang, A. Somekh-Baruch, H. V. Poor, S. Shamai, and S. Verdu, "Capacity of cognitive interference channels with and without secrecy," IEEE Trans. Inf. Theory, vol. IT-55, pp. 604-619, 2009.
- [18] L. Lai and H. El Gamal, "The relay-eavesdropper channel: cooperation for secrecy," *IEEE Trans Inf Theory*, vol. IT-54, no. 9, pp. 4005C4019, Sep. 2008.
- [19] Y. Oohama, "Coding for relay channels with confidential messages," in *Proceedings of IEEE Information Theory Workshop*, Australia, 2001.
- [20] E. Ekrem and S. Ulukus, "Secrecy in cooperative relay broadcast channels," IEEE Trans Inf Theory, vol. IT-57, pp. 137-155, 2011.
- [21] G. Kramer, M. Gastpar and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans Inf Theory*, vol. IT-51, pp. 3037-3063, 2005.
- [22] Y. Liang, H. V. Poor and S. Shamai, "Secure communication over fading channels," *IEEE Trans Inf Theory*, vol. IT-54, pp. 2470-2492, 2008.
- [23] T. M. Cover and A. El Gamal, "Capacity theorems for the relay channel," IEEE Trans Inf Theory, vol. IT-25, pp. 572-584, 1979.
- [24] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," IEEE Trans Inf Theory, vol. IT-25, pp. 306-311, 1979.
- [25] A. A. El Gamal and E. C. van der Meulen, "A proof of Martons coding theorems for the discrete memoryless broadcast channel," IEEE Trans Inf Theory, vol. IT-27, pp. 120-122, 1981.
- [26] L. Sankaranarayanan, G. Kramer and N. B. Mandayam, "Capacity theorems for the multiple-access relay channel," Proceedings of Allerton Conference on Communications, Control and Computing, 2004.
- [27] X. Tang, R. Liu, P. P. Spasojevic and H. V. Poor, "Interference assisted secret communication," *IEEE Trans Inf Theory*, vol. IT-57, pp. 3153-3167, 2011.
- [28] A. El Gamal and Y. H. Kim, Network information theory. Cambridge University Press, 2011.